NBER WORKING PAPER SERIES

ESTIMATING THE SPATIAL AMPLIFICATION OF DAMAGE CAUSED BY DEGRADATION IN THE AMAZON

Rafael Araujo
Juliano Assunção
Marina Hirota
José A. Scheinkman

Working Paper 31550
http://www.nber.org/papers/w31550

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
August 2023

MH thanks Serrapilheira Institute (grant number Serra-1709-18983). RA and JA thank Ilan Gleiser and Evan Bollig from AWS for comments on an earlier version of this project. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Estimating the Spatial Amplification of Damage Caused by Degradation in the Amazon
Rafael Araujo, Juliano Assunção, Marina Hirota, and José A. Scheinkman
NBER Working Paper No. 31550
August 2023
JEL No. C23,Q54,Q57

ABSTRACT

The Amazon rainforests have been undergoing unprecedented levels of human-induced disturbances. In addition to local impacts, such changes are likely to cascade following the eastern-western atmospheric flow generated by trade winds. We propose a model of spatial and temporal interactions created by this flow to estimate the spread of local disturbances to downwind locations along atmospheric trajectories. The spatial component captures cascading effects propagated by neighboring regions while the temporal component captures persistence. All these network effects can be described by a single matrix, acting as a spatial multiplier that amplifies local disturbances. This matrix can be used to easily map where the damage of an initial forest disturbance is amplified and propagated to. We identify regions that are likely to cause the largest impact throughout the basin, and those that are the most vulnerable to shocks caused by remote deforestation. On average, the presence of cascading effects mediated by winds doubles the impact of an initial damage. However, there is heterogeneity in this impact. While damage in some regions does not propagate, in others amplification may reach 250%.

Rafael Araujo
Fundação Getulio Vargas' Sao Paulo
School of Economics (FGV EESP)
Rua Itapeva, 474
01302-000, São Paulo/SP
Brazil
carlquist.rafael@gmail.com

Juliano Assunção
Departamento de Economia
PUC-Rio
Rua Marquês de São Vicente, 225/F210
22453-900 - Rio de Janeiro/RJ
Brazil
and Climate Policy Initiative
juliano@econ.puc-rio.br

Marina Hirota
Federal University of Santa Catarina
Florianópolis, SC
Brazil
marinahirota@gmail.com

José A. Scheinkman
Department of Economics
Columbia University
New York, NY 10027
and NBER
js3317@columbia.edu

1 Introduction

The Amazon rainforests[1] play a crucial role in global climate regulation due to its significant carbon storage capacity (Baccini et al., 2012; Mitchard, 2018). Releasing the carbon stored in this vast ecosystem would jeopardize our ability to limit global warming to 1.5°C (Armstrong McKay et al., 2022). There are concerns regarding the forest's resilience (Boulton et al., 2022), with certain areas transitioning from a carbon sink to a carbon source (Gatti et al., 2021). There are also studies pointing to a potential tipping point, when extensive portions of the forest undergo a shift towards low-density vegetation (Shukla et al., 1990; Wunderling et al., 2022).

Monitoring the local to basin-wide extension of deforestation and degradation impacts is essential to measure the extent to which local tree mortality caused by disturbances, such as extreme drought events or deforestation, can affect other regions and potentially trigger cascading impacts and further degradation. Existing literature on this question has pursued two main avenues of research. One approach employs large-scale calibrated models that connect forest dynamics and climate to construct forward-looking scenarios of forest resilience (Shukla et al., 1990; Nobre et al., 1991; Cox et al., 2004; Sampaio et al., 2007; Malhi et al., 2009; Hirota et al., 2011; Boisier et al., 2015; Zemp et al., 2017; Wunderling et al., 2022; Staal et al., 2023). Another strand of research focuses on measuring resilience indicators, such as changes in remotely sensed vegetation status (e.g., VOD and NDVI) (Verbesselt et al., 2016; Boulton et al., 2022), net carbon emissions (Phillips et al., 2009; Doughty et al., 2015; Gatti et al., 2021; Qin et al., 2021), and rainfall regimes (Salati et al., 1979; Panisset et al., 2018; Staal et al., 2018; Ciemer et al., 2019; Smith et al., 2023).

Although these approaches have provided invaluable knowledge on the dynamics of the forest, complementary insights can be gained through the application of reduced-form econometric methods as in (Diebold and Rudebusch, 2022; Diebold et al., 2022), which are earlier examples of the use of econometric methods in climate science. By abstracting from the intricate underlying mechanisms, these methods allow us to use a data-driven approach to estimate causal effects. In particular, we can make explicit the hypothesis that is necessary for statistical identification, that is, the necessary hypothesis to claim that we estimate causal effects rather than merely a correlation. Additionally, a reduced-form representation of forest dynamics allows for the establishment of linear relationships, enhancing the interpretation of the estimates, simplifying computations and statistical inference (see e.g. Wooldridge (2010); Arellano (2003)).

[1] We write the plural to emphasize the heterogeneity of the forest

Based on such econometric methods, we propose a model of the interactions across time and space that estimates local to basin-wide forest degradation within the Amazon rainforests. We use the model to assess how the Leaf Area Index (LAI) at a location is influenced by its own past LAI values (locally) and by LAI values of upwind locations (remotely) as determined by atmospheric trajectories. The spatial component of the model captures the cascading effects propagated by local disturbances to their neighbors while the temporal component captures the persistence of local disturbances. Of particular interest is the fact that all these feedback effects can be described by a single matrix, which acts as a spatial multiplier that amplifies disturbances occurring locally. By leveraging this matrix, we can identify locations that promote the largest basin-wide impacts as well the most vulnerable spots to remote disturbances.

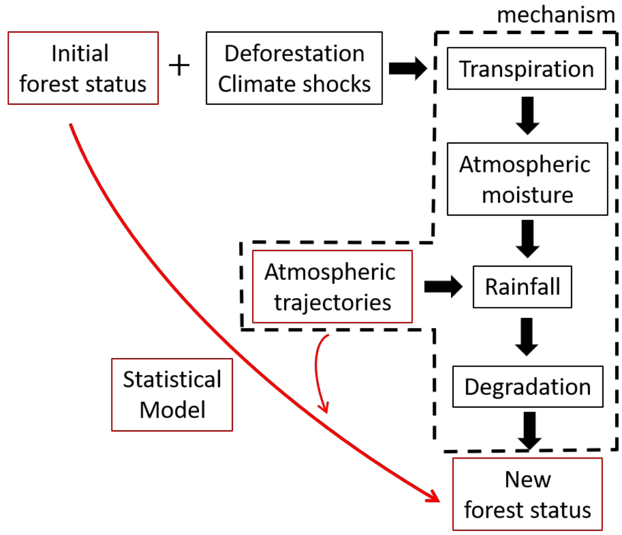
2 Model

Trees play a crucial role in the water cycle by absorbing water from the soil through the roots and releasing it back to the atmosphere through transpiration (Salati et al., 1979). This transpired water contributes to the moisture content in the air and is transported by atmospheric circulation contributing to downwind rainfall (Spracklen et al., 2012). Deforestation and other disturbances reduce overall transpiration, leading to a decrease in atmospheric moisture and rainfall levels for downwind regions. Less rainfall contributes to increasing forest degradation and subsequent CO_2 emissions (Qin et al., 2021) (see Figure 1). In our statistical model, we abstract from the transpiration mechanism to estimate how forest status of a region is affected by its past status, the *entire* configuration of forest status, and atmospheric circulation (Figure 1).

The model has temporal and spatial dimensions describing the dynamics of the forest. Let y_{it} denote the state of forest in pixel i and period t. In our application, we measure y_{it} as the Leaf Area Index (LAI) (Figure 2A). We represent the state of the entire forest in time t by a vector Y_t with entries y_{it}.

The time dimension follows a first-order auto-regressive process describing how a pixel of forest at t is affected by its previous state at $t-1$. On the spatial dimension, we allow for a high-order structure to capture the process of atmospheric circulation, through spatial lags. The spatial lags represent how the state of the forest in pixel i relates to the state of the forest in pixel j. Let $w_{ij,t} \in \{0, 1\}$ equals one if pixel j has a direct effect on pixel i via atmospheric circulation in period t. We make two identifying assumptions: that $w_{ii,t} = 0$ and that atmospheric circulation is unidirectional ($w_{ij,t}.w_{ji,t} = 0$).

Figure 1: **Mechanism and Model.**



Note: This figure illustrates the mechanism behind changes in forest vegetation and degradation. An initial forest status is perturbed by shocks (deforestation or droughts, for example). The shock will change the forest status, changing transpiration and thus atmospheric moisture. Atmospheric moisture is transported by atmospheric trajectories affecting the amount of rainfall in downwind regions. The change in downwind rainfall affects degradation processes that determine a new forest status. The statistical model abstracts from the transpiration mechanism and uses variations in atmospheric trajectory to estimate the dynamics of the forests.

We denote by W_t the climate-adjacency matrix with entries $(w_{ij,t})$, that is, it identifies pixels i and j that are direct geographical neighbors *and* are connected through atmospheric circulation in period t. The connection is established by back-trajectories built from observed wind data. We can then define $W_t^{[2]}$, the matrix of second-degree neighbors and, analogously, $W_t^{[k]}$ as the matrix of kth-degree neighbors. $W_t^{[k]}$ identifies pixels i and j that are k-th degree geographical neighbors *and* are connected through atmospheric circulation at t. See Figure S1 in the Appendix for an illustration of the climate-adjacency matrix. While W_t identifies the effects of direct neighbors, $W_t^{[2]}$ refers to the neighbors of the direct neighbors (second-degree neighbors), and so on.[2]

This notation allows us to specify the evolution of forest status as a spatial dynamic panel:

$$Y_t = \alpha Y_{t-1} + \sum_{k=1}^{K} \beta_k W_t^{[k]} Y_t + X\gamma + \varepsilon_t \tag{1}$$

where K denotes a maximum neighboring degree and $X\gamma$ is a set of pixel fixed-effects. The state of pixel i at time t (y_{it}) depends on its own state at time $t-1$ ($y_{i,t-1}$) and on the state of all pixels that affect i as first-degree neighbors ($W_t^{[1]}Y_t$), second-degree neighbors ($W_t^{[2]}Y_t$)

[2]To simplify formulae in what follows, we set $W^{[1]} := W$

and so on. This specification allows for effects that vary with the neighboring degree, i.e., a neighbor pixel has an effect of β_1 and the neighbor's neighbor pixel has an effect β_2. This specification does not nest Wunderling et al. (2022)'s approach, since in their paper the equivalent matrix $\beta_j W_t^{[j]} Y_t$ is treated as data. Instead, we parameterize the heterogeneity to create a more parsimonious model that can be estimated.

Identification. The main advantage of using a formulation like (1) is that it allows for a clear discussion of identification threats, that is, a discussion of whether potential confounding variables could potentially invalidate our estimates. If there are components in the error term ε_t that are correlated with $W_t^{[k]} Y_t$, then an Ordinary Least Squares estimator for (1) will render biased estimates for β_k. For example, if a road is built on the vicinity of neighboring pixels i and j, causing degradation on both, this effect will be wrongly captured by β_1 as being due to changes in climate caused by the upwind pixel degradation. To diminish concerns of this type of problem we adopt two procedures.

First, we employ an instrumental variable approach, which is widely used in economics (Angrist and Krueger, 2001) and political science (Sovey and Green, 2011). In an instrumental variable approach, we can further restrict the data variation used to identify the parameters of the model, disentangling the causal relationship from confounding factors. To implement such estimator we select variables (instruments) that (1) correlate with the explanatory variable and (2) affect the outcome variable solely through its influence on the explanatory variable.

In particular, we are concerned with confounding determinants of deforestation, as in the spatial lag term ($\beta_k W_t^{[k]} Y_t$) the variable Y_t is likely affected by anthropogenic processes. To deal with this, we isolate the variation in $W_t^{[k]} Y_t$ that is the result of only changes in wind. To build an instrument, we keep constant the forest status at the first period of data (Y_0) so that only the matrices W_t vary with time. Under the hypothesis that the wind pattern of interest is exogenous to local anthropogenic processes,[3] the term $W_t^{[k]} Y_0$ will not be correlated with ε_t but will correlate with $W_t^{[j]} Y_t$, capturing only the variation that is due to short-run changes in wind direction and velocity. We then use the variation of $W_t^{[j]} Y_0$ as an instrument for $W_t^{[j]} Y_t$.

Second, we allow for pixel fixed effects, which accounts for time-invariant characteristics, such as soil properties or proximity to water.

The importance of the assumption of unidirectionality. In our identification, we

[3] As we explain further in the data section, the altitude relevant for these processes is close to 3,000 meters above the forest canopy

make use of a very unique characteristic in our setting - that the matrix $W_t^{[k]}$ does not have bidirected effects, that is, $w_{ij,t}.w_{ji,t} = 0$. If pixel i affects pixel j at time t then pixel j does not affect pixel i at time t. The absence of this feedback effect allows us to avoid the simultaneity problem commonly present in Vector Autoregressive (see Stock and Watson (2001) for a review), Spatial Autoregressive (see Anselin (2010) for a review), and peer effects models (Manski, 1993). A possible exception to this diagonal argument is of potentially competing local mechanisms, which we discuss next.

Local effects. With (1) we aim to capture the effect that deforestation and other disturbances have on forest status in short and long range scenarios, that is, deforestation and other disturbances affecting forest status of regions that are hundreds or even thousands of kilometers away from the original disturbance. But, deforestation may also have local effects via other mechanisms (Cohn et al., 2019; Leite-Filho et al., 2021; Smith et al., 2023). Our identification strategy leverages our instrumental variable approach as a possible solution to this problem. By using $W_t^{[k]}Y_0$ as an instrument, we isolate only the variation that can be explained by wind variations.

Counterfactuals. In counterfactual exercises we map the forest's vegetation dynamics given a vector of shocks/disturbances (deforestation or a climate shock, for example). To isolate the effects of the shocks we fix the $W_t^{[k]}$ matrix to its sample average, that is, we define $W^{[k]} = \frac{1}{T}\sum_t W_t^{[k]}$, where T denotes the time-length of the data. For this exercise to be of interest, we must assume that the process of atmospheric transport is at least mean-ergodic, that is, the sample average ($W^{[k]}$) is representative of future patterns of atmospheric circulation. Finally, to ease the presentation of the next expressions we add a notation for the inverse matrix of the spatial lags: $\Omega := \left(I - \sum_{k=1}^{K} \beta_k W^{[k]}\right)^{-1}$. The expression of Ω summarizes in a single matrix all the feedback effects of the spatial dimension. A useful interpretation is provided by writing Ω as the Neumann series of matrices. For example, with only one spatial lag we would have that $(I - \beta_1 W)^{-1} = I + \beta_1 W + \beta_1^2 W^2 + \beta_1^3 W^3 + ...$ Hence a shock in a pixel will affect its first-degree neighbors by β_1, but its first-degree neighbors are also first-degree neighbors of other pixels, which affect it by β_1^2 and so on. The matrix Ω accounts for all of this chain of effects for an arbitrary number of spatial lags (K).

With this notation, we can write the evolution of the state of the forest as a function of a initial state and shocks. For example, for period t we have:

$$\left(I - \sum_{k=1}^{K} \beta_k W^{[k]}\right) Y_t = \alpha Y_{t-1} + X\gamma + \epsilon_t$$

which we can rewrite as:

$$Y_t = \alpha\Omega Y_{t-1} + \Omega X\gamma + \Omega\epsilon_t$$

We can substitute past values of Y_t iteratively for any arbitrary Δ time window, describing a general formulation with the following expression:

$$Y_{t+\Delta} = \alpha^{\Delta+1}\Omega^{\Delta+1}Y_{t-1} + \sum_{i=0}^{\Delta}\alpha^i\Omega^{i+1}X\gamma + \sum_{i=0}^{\Delta}\alpha^i\Omega^{i+1}\epsilon_{t+\Delta-i} \tag{2}$$

Notice that this formulation cannot be used for estimation purposes, since we do not observe the shocks $\epsilon_t, ..., \epsilon_{t+\Delta}$. But, after estimating the parameters, this expression helps us write the impulse response function ($\phi(\Delta)$) that returns the effect that a shock in time t has on the forest status at time $t + \Delta$. Given our panel data structure, there are many ways to define the output of the impulse response function. Here we define it as total change in the forest status over all pixels. Formally, let $\mathbb{I}_{row}$ denote a row vector of ones, then we define the impulse response function as:

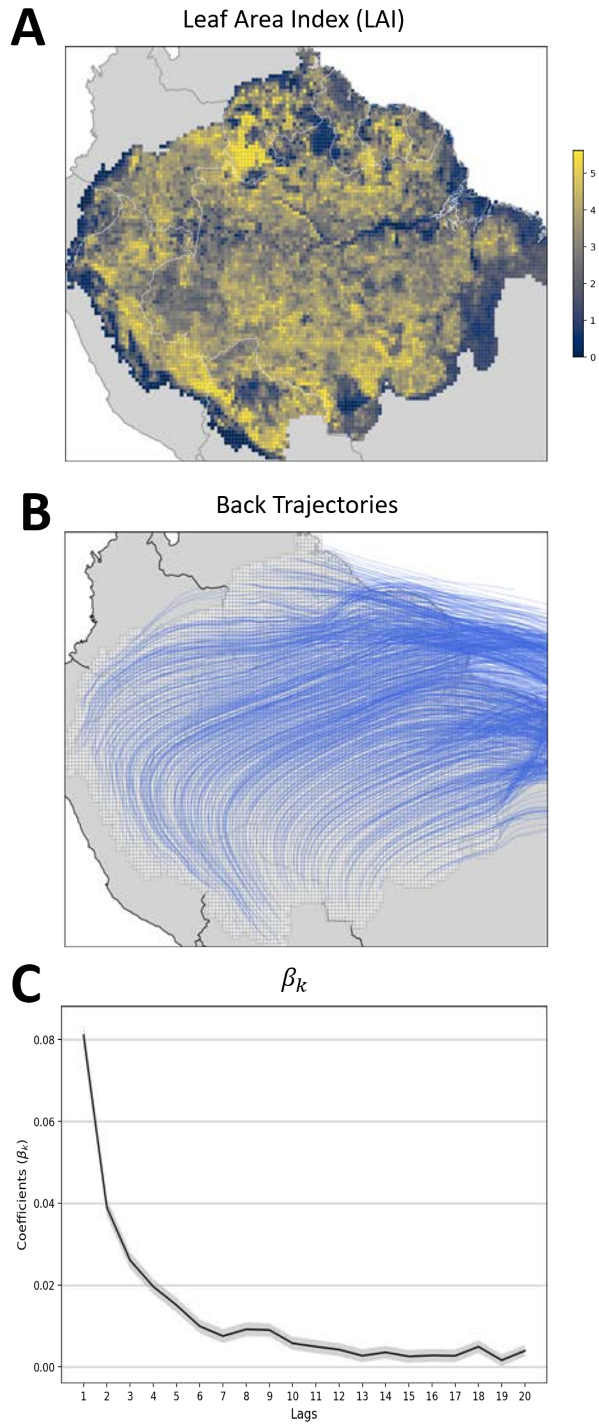
$$\phi(t+\Delta) = \alpha^{\Delta}\mathbb{I}_{row}\left(\Omega^{\Delta+1}\epsilon_t\right) \tag{3}$$

From (3), the effect that a vector of shocks (ϵ_t) will have on $\phi(t + \Delta)$ is defined by the temporal persistence coefficient (α) and the spatial feedback matrix (Ω). Thus the effect of a specific shock ϵ_i – the shock in pixel i – has a linear effect on the impulse response function. The strength of this linear effect is measured by the sum of the values in the columns of the matrix $\Omega^{\Delta+1}$. Summing the values on each row of $\Omega^{\Delta+1}$ gives a measure of the potential of a pixel to *affect* the overall Amazon system. Thus, the sum of the values in each row of the spatial feedback matrix can be interpreted as the *index of criticality* of a pixel. Analogously, summing the values on each column of $\Omega^{\Delta+1}$ gives a measure of the potential of a pixel *to be affected* by the overall Amazon system. Thus, the sum of the values in each column of the spatial feedback matrix can be interpreted as an *index of exposure* of each pixel.

Estimation

Since (1) constitutes a dynamic panel, we cannot employ a standard fixed effects estimator but we can take first-differences to remove the fixed effect and use lagged values of Y_t as an instrument for the model in differences. In dynamic panels, the standard demeaning

Figure 2: **Data and Estimates.**



Note: Data showing (A) the average leaf area index for the year 1985 (LAI) representing local forest status, and (B) atmospheric transport, calculated as the 5-day back trajectories for a sample of our pixels. (C) Estimated spatial lags (β_k) for each k between 1 and 20. The shaded gray area shows the 95% confidence interval with standard errors clustered at the pixel level.

procedure creates new independent variables that are correlated with the error term. This can bias the estimates of β_k. Taking the first difference is an alternative strategy to remove a fixed effect. This estimation corresponds to the strategies proposed by Anderson and Hsiao (1981) and Arellano and Bond (1991) for dynamic panels. Taking first differences:

$$\Delta Y_t = \alpha \Delta Y_{t-1} + \sum_{k=1}^{K} \beta_k \left(\Delta W_t^{[k]} Y_t \right) + \tilde{\epsilon}_t \tag{4}$$

We use the following moment conditions to identify the parameters ((5) and (6)). This estimator is just-identified, since the number of moment conditions matches the number of estimated parameters. Notice we use the lagged level variable Y_{t-2} as the instrument instead of the lagged difference (Arellano, 1989; Phillips and Han, 2015). We could over-identify the model, that is, have more moments than parameters, by using further lags ($Y_{t-3}, Y_{t-4}...$). The advantage would be a more efficient estimator with smaller standard errors. However, given the size of our sample, our standard errors would not improve much, furthermore, an over-identified model would require numerical optimization, which is not necessary in a just-identified model. The numerical optimization could also introduce unnecessary numerical noise to our estimates.

$$\mathbb{E}\left[\tilde{\epsilon}_t Y_{t-2}\right] = 0 \tag{5}$$

$$\mathbb{E}\left[\tilde{\epsilon}_t \Delta W_t^{[j]} Y_0\right] = 0, \tag{6}$$

A note on dynamics

The coefficient α captures a partial temporal persistence of $y_{i,t}$. The interpretation of this coefficient is partial because all the effects of the state of the forest in period t are already being captured by the remaining terms in (1), that is, α captures the temporal persistence effect after partialling out (or controlling for) the effect of the current state of the forest, which also depends on past states of the forest. Therefore, the low persistence temporal effect we estimate ($\alpha = 0.22$) does not necessarily mean that the forest has a low persistence on its status.

On aggregate shocks

We could include time fixed effects on our estimation, which could be either a month-year or only year fixed effect to capture aggregate shocks, such as EL Nino e La Nina phenomena. Nonetheless, it is likely that such effects change slowly in a way that the moment conditions ($\mathbb{E}\left[\tilde{\epsilon}_t \Delta\gamma_t\right] = 0$) would be numerically very close to zero. This is so because we already take the first difference ($\Delta\gamma_t$) before estimating the model. For persistent aggregate shocks that last for months, the first difference would already be partialling out most of their effects.

On controls

To avoid bias estimates in our reduced-form model, it is important not to control for variables that mediate the effect we want to study. Such variables are commonly known as "bad controls" (Wooldridge, 2005; Angrist and Pischke, 2009) or "mediators" (Cinelli et al., 2022). In our setting, examples of such mediators include precipitation and fire, both of which play significant roles in the dynamics of the Amazon system (Staver et al., 2011; Touboul et al., 2018; Wunderling et al., 2022). A potential causal chain of effects involves: (a) a decrease in upwind exposure to LAI leads to (b) a decrease in precipitation downwind and (c) an increase in fire vulnerability downwind, thus (d) decreasing downwind LAI. Including precipitation and/or fire as controls would keep important mechanisms constant or block their path of causality, thereby absorbing the necessary variation to estimate the model. In this example, instead of breaking down the mechanisms in the causal chain of effects $(a) \to (b) \to (c) \to (d)$ we estimate directly $(a) \to (d)$.

3 Data

The climate-adjacency matrix W_t is built from two matrices, G and C_t. G is a matrix that describes immediate geographical neighbors, $G_{ij} = 1$ if pixels i and j are immediate neighbors, otherwise $G_{ij} = 0$. Notice that G is constant over time. The matrix C_t is built with data on back-trajectories of atmospheric transportation and describes atmospheric circulation: $C_{ij,t} = 1$, if at time t a trajectory of atmospheric circulation goes through j before reaching i. Otherwise, $C_{ij,t} = 0$.

Formally, to compute $W_t^{[k]}$, we first compute

$$\hat{W}_t^{[k]} = \left(\mathbb{I}(G^k > 0) - \sum_{j=1}^{k-1} \mathbb{I}(G^j > 0) \right) \odot C_t$$

and then change the diagonal elements of $\hat{W}_t^{[k]}$ to zero. Here $\odot$ denotes element-wise multiplication. Figure S1 in the Appendix illustrates these matrices.

To build the back-trajectories we use monthly three-dimensional wind data (direction and speed) from 1985 to 2013 Copernicus (2017). For each pixel of 0.25° resolution in the Amazon, we trace back the trajectory that arrives at that pixel at 800 hPa (Spracklen et al., 2012) for 5 days, which is enough time for back-trajectories to reach the ocean.

To measure the forest status (Y_t), we use monthly Leaf Area Index (LAI) data (Xiao et al., 2016) at the resolution of 0.05°. The LAI is a measurement that quantifies the amount of leaf surface area relative to the ground area. It provides information about the density and extent of vegetation cover. A higher LAI value indicates a denser canopy with more leaves, which can result in increased photosynthesis, carbon sequestration, biomass production, and overall productivity. To make both LAI and atmospheric trajectories data compatible, we reduce LAI's data resolution applying an average kernel. In the Appendix, Figure S3, we show the results for different ways of assigning the LAI data to match the wind data. Our results remain stable under these different sampling strategies. Figure 2 shows the spatial distribution of LAI in the Amazon in the beginning of our sample.

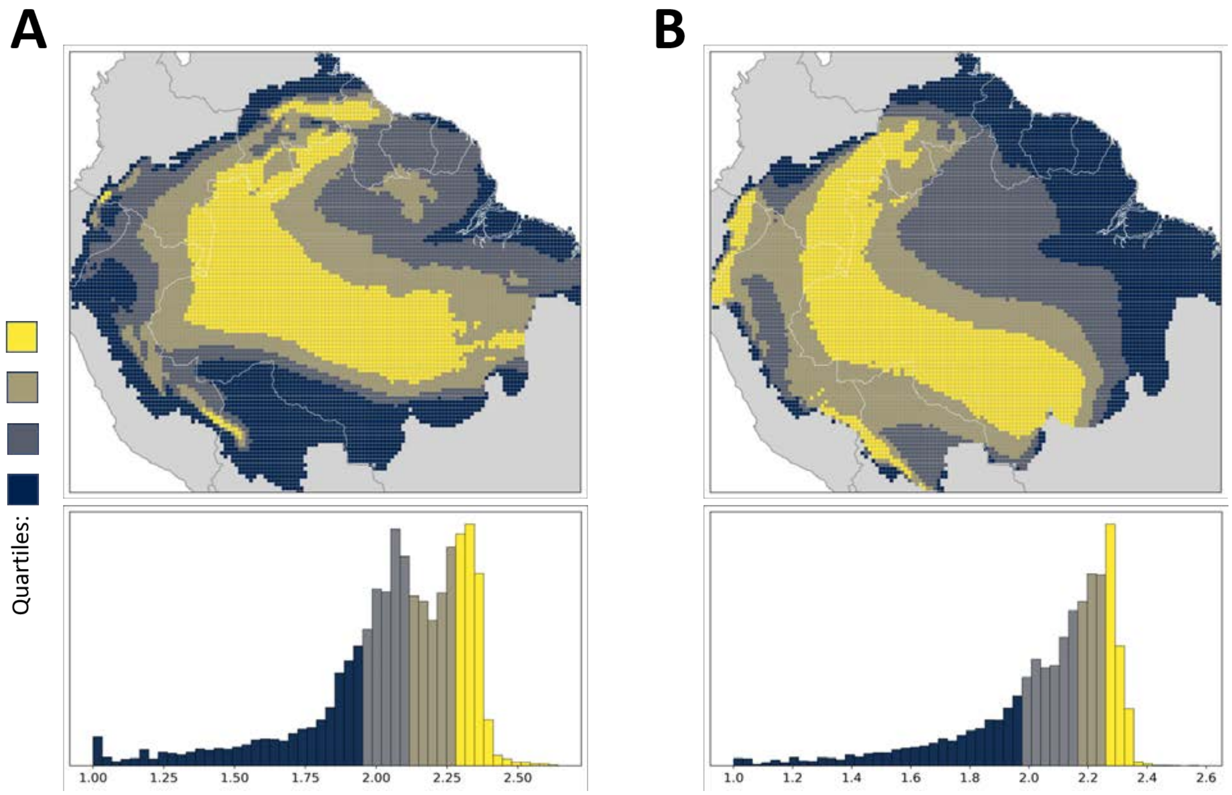
In the simulation exercise, we make use of deforestation data. To identify pixels that were most deforested between 2001 and 2022, we selected pixels with more than 50% of deforested area during that period, according to the annual data in (Hansen et al., 2013).

4 Results

Estimation. We estimate (1) with 20 lags ($K = 20$). We find a temporal persistence coefficient (α) of 0.22 (standard error is 0.001; p-value $< 0.001\%$). This persistence coefficient is higher than the average persistence of individual effects computed by Boulton et al. (2022), but within the bounds of one standard deviation of these individual coefficients. Figure 2C shows the estimated β_k's.

In Figure 2C we can only have a partial interpretation of the coefficients. The coefficients β_k gives the direct effect of a neighbor of degree k on a pixel of interest. Nonetheless, the k-th neighbor of a pixel is also the first, second or k-th neighbor of other pixels. This compounded effect, which accounts for feedback effects in all pixels, can only be captured by the inverse matrix of the spatial lags: $\Omega = \left(I - \sum_{k=1}^{K} \beta_k W^{[k]}\right)^{-1}$. Yet, a partial interpretation is also useful. Summing the β_k's, we get a direct cumulative effect of 0.26. Thus, the persistence coefficient ($\alpha = 0.22$) and the direct spatial lag effect ($\sum_k \beta_k = 0.26$) create a

Figure 3: **Index of Criticality (A) and Index of Exposure (B).**



Note: (A) the total effect that a shock in a pixel has accounting for all the spatial feedbacks. Each entry $\Omega(i, j)$ represents the effect that pixel j has on pixel i. For each pixel, represented by a column of Ω, the map shows the sum of values in that column. This sum measures the potential of that pixel to *affect* the entire forest (B) the total effect that a pixel *receives* accounting for all the spatial feedbacks. For each pixel, represented by a row of Ω, the map shows the sum of values in that row, this sum measures the potential of that pixel to be *affected* by the entire forest. The maps are colored according to the quartiles of the effects, that is, it shows the distribution from the 25% smallest multiplier effects (blue) to the 25% largest ones (yellow). The histogram below each map shows the distribution of the multiplier effects.

compounded effect of 0.48, that is, 48% of a variation in the current state of a pixel can be explained by temporal persistence plus direct upwind variations. The direct spatial lag effect ($\sum_k \beta_k = 0.26$) can be closely compared with Wunderling et al. (2022)'s results where the network effect is the tipping reason in 25% percent of the Amazon basin. All results are displayed in Table S1 in the Appendix.

Specifications with more than 20 lags ($K > 20$) show no significant results for $k > 20$. Figure S2 in the Appendix shows the result of estimating (1) with 60 lags, for which β_k with $k > 20$ is not significant at the 5% significance level.

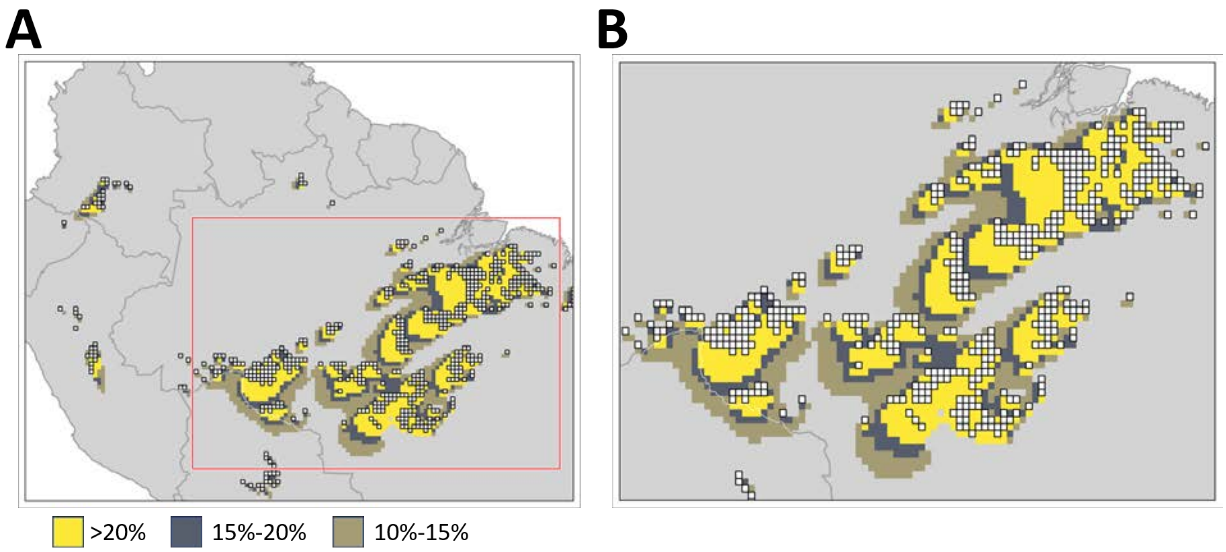
Counterfactual. With the estimated coefficients we compute the matrix Ω. We show the results for: (1) summing the values on each column of Ω, which measures the potential that each pixel has to *affect* the entire forest; (2) summing the values on each row of Ω, which measures the potential that each pixel has *to be affected* by the entire forest (Figure 3 for $\Delta = 1$). The multiplier effect, caused by the cascading effects of atmospheric trajectories, means that a shock at a pixel has an effect on a range of different pixels in the forest, causing an overall effect computed as the initial shock times a multiplier effect (Figure 3A). This multiplier effect is the sum of values in the column corresponding to the pixel where the initial shock occurred, that is, our index of criticality (Figure 3A). Looking at the other side of the coin, given a homogeneous shock at all pixels of the forest, pixels are likely to be affected heterogeneously. This overall effect is captured by a multiplier effect of the aggregate shock, that is, our index of exposure (Figure 3B). This multiplier effect is the sum of the values of the row corresponding to each pixel, that is, the values depicted in Figure 3B.

The matrix Ω is more complex than the results in Figure 3, since it gives the specific effect of one pixel on another, that is, the entry $\Omega(i, j)$ gives the effect that pixel j has on pixel i. For policy making purposes, we perform a targeted exercise, considering degradation caused by the pixels that were mostly deforested between 2001 and 2022 (Hansen et al., 2013). Thus, we select the pixels that had more than 50% of its area deforested (white squares in Fig. 4) and map the areas that would suffer indirect degradation (coloured squares in Fig. 4). The estimated multiplier effects allow us to measure this indirect degradation effect as a proportion of the decrease in LAI caused by upwind deforestation (Figure 4). To do this, for each pixel that is affected, we plot the sum of the multiplier effects corresponding to the most deforested pixels. These measures of indirect degradation are highly heterogeneous, showing the importance of mapping affected areas to efficiently allocate public policies of conservation and restoration.

Given the estimated temporal persistence parameter (α) of 0.22, the effect of a shock in t in forest status at period $t + \Delta$ decreases fast as Δ increases, faster than the multiplier effect of $\Omega^{\Delta+1}$ increases. That is, the magnitude of the contemporary cascading effects of forest disturbances dominates any past effects, or the spatial component of the model is more relevant than the temporal component. In Figures S3 and S4 in the Appendix we show the spatial evolution of $\Omega^{\Delta+1}$ for different Δ's. The magnitudes decrease sharply and the set of the most important pixels move constantly to north-northeastern region, following the inverse trajectory of trade winds.

It is important to acknowledge the limitations of our research. We do not account for other spillover effects of forest disturbance that are likely to amplify our estimates, such as border effects and local temperature effects. By design, the effects we estimate are only the ones mediated by atmospheric transport. Thus, our results represent a conservative estimate of

Figure 4: **Deforestation and Cascading Effects**



Note: These maps show: the pixels that were most deforested between 2001 and 2022 (the white squares in A and B) (Hansen et al., 2013). We selected pixels that deforested more than 50%. Together those pixels represent 32% of the total deforestation in the Amazon; we then map the areas that are affected by the deforestation in the white pixels (the values on the columns of Ω corresponding to the white pixels). To facilitate visualization we breakdown the affected pixels in bins (10%-15%, 15%-20%, and >20%) where each bin represents the share of the change in LAI in the white pixels that will be propagated. Map B shows a zoomed in version of map A, in the arch of deforestation in Brazil.

the extent to which local forest disturbances contribute to remote degradation.

Nonetheless, our model of the Amazon system dynamics provides a practical tool for policy-making. By utilizing the Ω matrix, we can map both the anticipated (ex-ante) and observed (ex-post) degradation effects resulting from forest disturbances, including climate shocks and deforestation. This methodology is valuable for conducting cost-benefit analyses of projects that contribute to deforestation, such as the construction of transportation infrastructure, as well as for assessing the effectiveness of forest protection policies, such as the establishment of conservation units.

5 Conclusion

We present a novel approach using reduced-form econometric methods to assess the dynamics of forest degradation in the Amazon rainforests. We study the interactions, both locally and remotely, of the Leaf Area Index (LAI) across the entire Amazon basin. Our data-driven model accounts for temporal persistence and also uses data on wind-trajectories to capture spatial cascading effects of disturbances that affect the resilience of the Amazon ecosystem.

Our approach allows us to estimate causal effects, adding to the literature of large-scale calibrated models and measurements of forest resilience indicators.

The results of our study reveal significant implications for the conservation and restoration of the Amazon rainforests. We find that, on average, the presence of cascading effects mediated by atmospheric circulation in the Amazon doubles the impact of an initial damage. As there is significant heterogeneity in this multiplier effect, we identify critical areas that act as the highest amplifiers of disturbances, propagating impacts to neighboring areas and potentially leading to basin-wide consequences, thus highlighting areas that require special attention in conservation policies and that could be prioritized in restoration policies. Additionally, we pinpoint locations that are particularly vulnerable to remote disturbances. This research enhances our understanding of the complex processes occurring within the Amazon rainforests, enabling policymakers to make informed decisions in safeguarding the forest's crucial role in the lives of the local communities, global climate regulation, and biodiversity preservation.

References

Anderson, Theodore Wilbur and Cheng Hsiao, "Estimation of dynamic models with error components," *Journal of the American statistical Association*, 1981, *76* (375), 598–606.

Angrist, JD and JS Pischke, "Mostly Harmless Econometrics: An Empiricist's Guide p. 373," 2009.

Angrist, Joshua D and Alan B Krueger, "Instrumental variables and the search for identification: From supply and demand to natural experiments," *Journal of Economic perspectives*, 2001, *15* (4), 69–85.

Anselin, Luc, "Thirty years of spatial econometrics," *Papers in regional science*, 2010, *89* (1), 3–25.

Arellano, Manuel, "A note on the Anderson-Hsiao estimator for panel data," *Economics letters*, 1989, *31* (4), 337–341.

_ , *Panel data econometrics*, OUP Oxford, 2003.

_ **and Stephen Bond**, "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations," *The review of economic studies*, 1991, *58* (2), 277–297.

Baccini, AGSJ, SJ Goetz, WS Walker, NT Laporte, Mindy Sun, Damien Sulla-Menashe, Joe Hackler, PSA Beck, Ralph Dubayah, MA Friedl et al., "Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps," *Nature climate change*, 2012, *2* (3), 182–185.

Boisier, Juan P, Philippe Ciais, Agnès Ducharne, and Matthieu Guimberteau, "Projected strengthening of Amazonian dry season by constrained climate model simulations," *Nature Climate Change*, 2015, *5* (7), 656–660.

Boulton, Chris A, Timothy M Lenton, and Niklas Boers, "Pronounced loss of Amazon rainforest resilience since the early 2000s," *Nature Climate Change*, 2022, *12* (3), 271–278.

Ciemer, Catrin, Niklas Boers, Marina Hirota, Jürgen Kurths, Finn Müller-Hansen, Rafael S Oliveira, and Ricarda Winkelmann, "Higher resilience to climatic disturbances in tropical vegetation exposed to more variable rainfall," *Nature Geoscience*, 2019, *12* (3), 174–179.

Cinelli, Carlos, Andrew Forney, and Judea Pearl, "A crash course in good and bad controls," *Sociological Methods & Research*, 2022, p. 00491241221099552.

Cohn, Avery S, Nishan Bhattarai, Jake Campolo, Octavia Crompton, David Dralle, John Duncan, and Sally Thompson, "Forest loss in Brazil increases maximum temperatures within 50 km," *Environmental Research Letters*, 2019, *14* (8), 084047.

Copernicus, Climate Change Service, "ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate," 2017.

Cox, Peter M, RA Betts, Matthew Collins, Phil P Harris, Chris Huntingford, and CD Jones, "Amazonian forest dieback under climate-carbon cycle projections for the 21st century," *Theoretical and applied climatology*, 2004, *78*, 137–156.

Diebold, Francis X and Glenn D Rudebusch, "Probability assessments of an ice-free Arctic: Comparing statistical and climate model projections," *Journal of Econometrics*, 2022, *231* (2), 520–534.

_ , _ , Maximilian Göbel, Philippe Goulet Coulombe, and Boyuan Zhang, "When Will Arctic Sea Ice Disappear? Projections of Area, Extent, Thickness, and Volume," Technical Report, National Bureau of Economic Research 2022.

Doughty, Christopher E, DB Metcalfe, CAJ Girardin, F Farfán Amézquita, D Galiano Cabrera, W Huaraca Huasco, JE Silva-Espejo, A Araujo-Murakami, MC Da Costa, W Rocha et al., "Drought impact on forest carbon dynamics and fluxes in Amazonia," *Nature*, 2015, *519* (7541), 78–82.

Gatti, Luciana V, Luana S Basso, John B Miller, Manuel Gloor, Lucas Gatti Domingues, Henrique LG Cassol, Graciela Tejada, Luiz EOC Aragão, Carlos Nobre, Wouter Peters et al., "Amazonia as a carbon source linked to deforestation and climate change," *Nature*, 2021, *595* (7867), 388–393.

Hansen, Matthew C, Peter V Potapov, Rebecca Moore, Matt Hancher, Svetlana A Turubanova, Alexandra Tyukavina, David Thau, Stephen V Stehman, Scott J Goetz, Thomas R Loveland et al., "High-resolution global maps of 21st-century forest cover change," *science*, 2013, *342* (6160), 850–853.

Hirota, Marina, Milena Holmgren, Egbert H Van Nes, and Marten Scheffer, "Global resilience of tropical forest and savanna to critical transitions," *Science*, 2011, *334* (6053), 232–235.

Leite-Filho, Argemiro Teixeira, Britaldo Silveira Soares-Filho, Juliana Leroy Davis, Gabriel Medeiros Abrahão, and Jan Börner, "Deforestation reduces rainfall and agricultural revenues in the Brazilian Amazon," *Nature Communications*, 2021, *12* (1), 2591.

Malhi, Yadvinder, Luiz EOC Aragão, David Galbraith, Chris Huntingford, Rosie Fisher, Przemyslaw Zelazowski, Stephen Sitch, Carol McSweeney, and Patrick Meir, "Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest," *Proceedings of the National Academy of Sciences*, 2009, *106* (49), 20610–20615.

Manski, Charles F, "Identification of endogenous social effects: The reflection problem," *The review of economic studies*, 1993, *60* (3), 531–542.

McKay, David I Armstrong, Arie Staal, Jesse F Abrams, Ricarda Winkelmann, Boris Sakschewski, Sina Loriani, Ingo Fetzer, Sarah E Cornell, Johan Rockström, and Timothy M Lenton, "Exceeding 1.5 C global warming could trigger multiple climate tipping points," *Science*, 2022, *377* (6611), eabn7950.

Mitchard, Edward TA, "The tropical forest carbon cycle and climate change," *Nature*, 2018, *559* (7715), 527–534.

Nobre, Carlos A, Piers J Sellers, and Jagadish Shukla, "Amazonian deforestation and regional climate change," *Journal of climate*, 1991, *4* (10), 957–988.

Panisset, Jéssica S, Renata Libonati, Célia Marina P Gouveia, Fausto Machado-Silva, Daniela A França, José Ricardo A França, and Leonardo F Peres, "Contrasting patterns of the extreme drought episodes of 2005, 2010 and 2015 in the Amazon Basin," *International Journal of Climatology*, 2018, *38* (2), 1096–1104.

Phillips, Oliver L, Luiz EOC Aragão, Simon L Lewis, Joshua B Fisher, Jon Lloyd, Gabriela López-González, Yadvinder Malhi, Abel Monteagudo, Julie Peacock, Carlos A Quesada et al., "Drought sensitivity of the Amazon rainforest," *Science*, 2009, *323* (5919), 1344–1347.

Phillips, Peter CB and Chirok Han, "The true limit distributions of the Anderson–Hsiao IV estimators in panel autoregression," *Economics Letters*, 2015, *127*, 89–92.

Qin, Yuanwei, Xiangming Xiao, Jean-Pierre Wigneron, Philippe Ciais, Martin Brandt, Lei Fan, Xiaojun Li, Sean Crowell, Xiaocui Wu, Russell Doughty et al., "Carbon loss from forest degradation exceeds that from deforestation in the Brazilian Amazon," *Nature Climate Change*, 2021, *11* (5), 442–448.

Salati, Eneas, Attilio Dall'Olio, Eiichi Matsui, and Joel R Gat, "Recycling of water in the Amazon basin: an isotopic study," *Water resources research*, 1979, *15* (5), 1250–1258.

Sampaio, Gilvan, Carlos Nobre, Marcos Heil Costa, Prakki Satyamurty, Britaldo Silveira Soares-Filho, and Manoel Cardoso, "Regional climate change over eastern Amazonia caused by pasture and soybean cropland expansion," *Geophysical Research Letters*, 2007, *34* (17).

Shukla, Jagadish, Carlos Nobre, and Piers Sellers, "Amazon deforestation and climate change," *Science*, 1990, *247* (4948), 1322–1325.

Smith, C, JCA Baker, and DV Spracklen, "Tropical deforestation causes large reductions in observed precipitation," *Nature*, 2023, *615* (7951), 270–275.

Sovey, Allison J and Donald P Green, "Instrumental variables estimation in political science: A readers' guide," *American Journal of Political Science*, 2011, *55* (1), 188–200.

Spracklen, Dominick V, Steve R Arnold, and CM Taylor, "Observations of increased tropical rainfall preceded by air passage over forests," *Nature*, 2012, *489* (7415), 282–285.

Staal, Arie, Gerbrand Koren, Graciela Tejada, and Luciana V Gatti, "Moisture origins of the Amazon carbon source region," *Environmental Research Letters*, 2023, *18* (4), 044027.

_ , Obbe A Tuinenburg, Joyce HC Bosmans, Milena Holmgren, Egbert H van Nes, Marten Scheffer, Delphine Clara Zemp, and Stefan C Dekker, "Forest-rainfall cascades buffer against drought across the Amazon," *Nature Climate Change*, 2018, *8* (6), 539–543.

Staver, A Carla, Sally Archibald, and Simon A Levin, "The global extent and determinants of savanna and forest as alternative biome states," *science*, 2011, *334* (6053), 230–232.

Stock, James H and Mark W Watson, "Vector autoregressions," *Journal of Economic perspectives*, 2001, *15* (4), 101–115.

Touboul, Jonathan David, Ann Carla Staver, and Simon Asher Levin, "On the complex dynamics of savanna landscapes," *Proceedings of the National Academy of Sciences*, 2018, *115* (7), E1336–E1345.

Verbesselt, Jan, Nikolaus Umlauf, Marina Hirota, Milena Holmgren, Egbert H Van Nes, Martin Herold, Achim Zeileis, and Marten Scheffer, "Remotely sensed resilience of tropical forests," *Nature Climate Change*, 2016, *6* (11), 1028–1031.

Wooldridge, Jeffrey M, "Violating ignorability of treatment by controlling for too many factors," *Econometric Theory*, 2005, *21* (5), 1026–1028.

_ , *Econometric analysis of cross section and panel data*, MIT press, 2010.

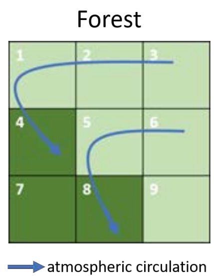
Wunderling, Nico, Arie Staal, Boris Sakschewski, Marina Hirota, Obbe A Tuinenburg, Jonathan F Donges, Henrique MJ Barbosa, and Ricarda Winkelmann, "Recurrent droughts increase risk of cascading tipping events by outpacing adaptive capacities in the Amazon rainforest," *Proceedings of the National Academy of Sciences*, 2022, *119* (32), e2120777119.

Xiao, Zhiqiang, Shunlin Liang, Tongtong Wang, and Bo Jiang, "Retrieval of leaf area index (LAI) and fraction of absorbed photosynthetically active radiation (FAPAR) from VIIRS time-series data," *Remote Sensing*, 2016, *8* (4), 351.

Zemp, Delphine Clara, Carl-Friedrich Schleussner, Henrique MJ Barbosa, Marina Hirota, Vincent Montade, Gilvan Sampaio, Arie Staal, Lan Wang-Erlandsson, and Anja Rammig, "Self-amplified Amazon forest loss due to vegetation-atmosphere feedbacks," *Nature communications*, 2017, *8* (1), 14681.

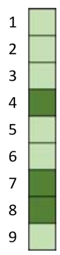
1 Appendix

Figure 1.1: Matrices that summarize the forest.



$$C_t = \begin{array}{c|ccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 8 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \qquad G = \begin{array}{c|ccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 4 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 5 & 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 \\ 6 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 & 1 \\ 7 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 \\ 8 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 1 \\ 9 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 0 \end{array}$$

$Y_t =$



$$W_t = \begin{array}{c|ccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 8 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \qquad W_t^{[2]} = \begin{array}{c|ccccccccc} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ \hline 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 6 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}$$

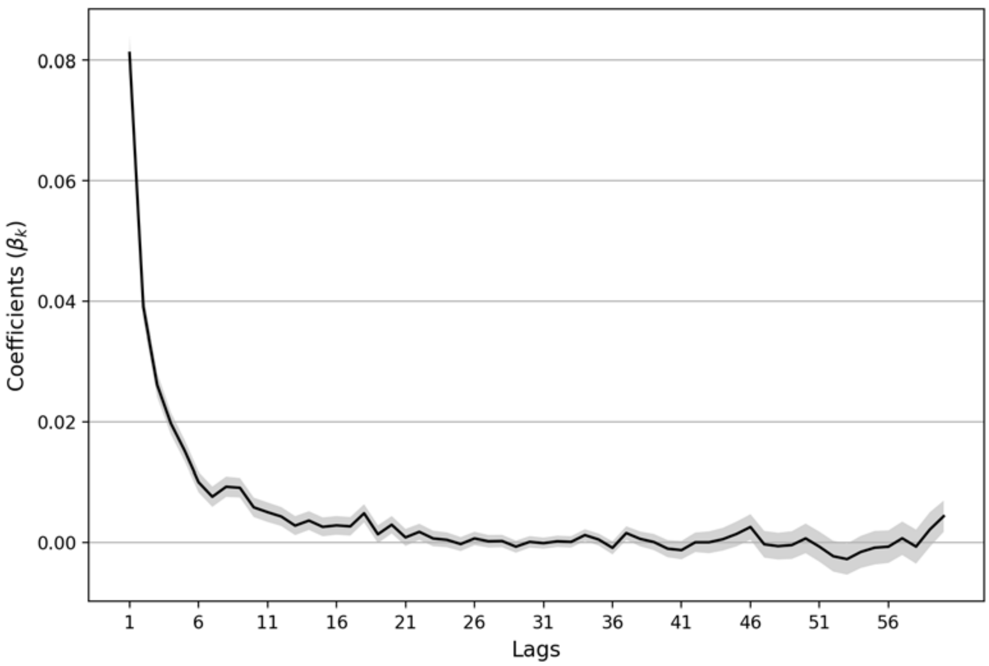
Note: This figure illustrates the matrices used to build the data. We start with the LAI data that represents the forest (the tones of green) and the atmospheric circulation (the blue arrows); our Y_t vector is a vectorized version of the forest LAI matrix; C_t is a matrix that shows for each pixel (row) which pixels (column) affect it. The G matrix describes whether pixels (i, j) are immediate geographical neighbors; $W_t := G \odot C_t$, where $\odot$ denotes the Hadamard or component-wise product, defines the first-degree climate neighbors; $W_t^{[2]}$ shows second-degree climate neighbors.

Table 1.1: Estimation results.

variable	coefficient	standard error
α	0.2200	(0.0019)
β_1	0.0811	(0.0014)
β_2	0.0391	(0.0011)
β_3	0.0262	(0.001)
β_4	0.0197	(0.0009)
β_5	0.0152	(0.0009)
β_6	0.0100	(0.0009)
β_7	0.0076	(0.0009)
β_8	0.0093	(0.0009)
β_9	0.0091	(0.0008)
β_{10}	0.0058	(0.0008)
β_{11}	0.0050	(0.0008)
β_{12}	0.0043	(0.0008)
β_{13}	0.0028	(0.0008)
β_{14}	0.0036	(0.0008)
β_{15}	0.0026	(0.0008)
β_{16}	0.0029	(0.0008)
β_{17}	0.0028	(0.0008)
β_{18}	0.0050	(0.0008)
β_{19}	0.0017	(0.0008)
β_{20}	0.0040	(0.0007)
constant	0.9605	(0.0462)

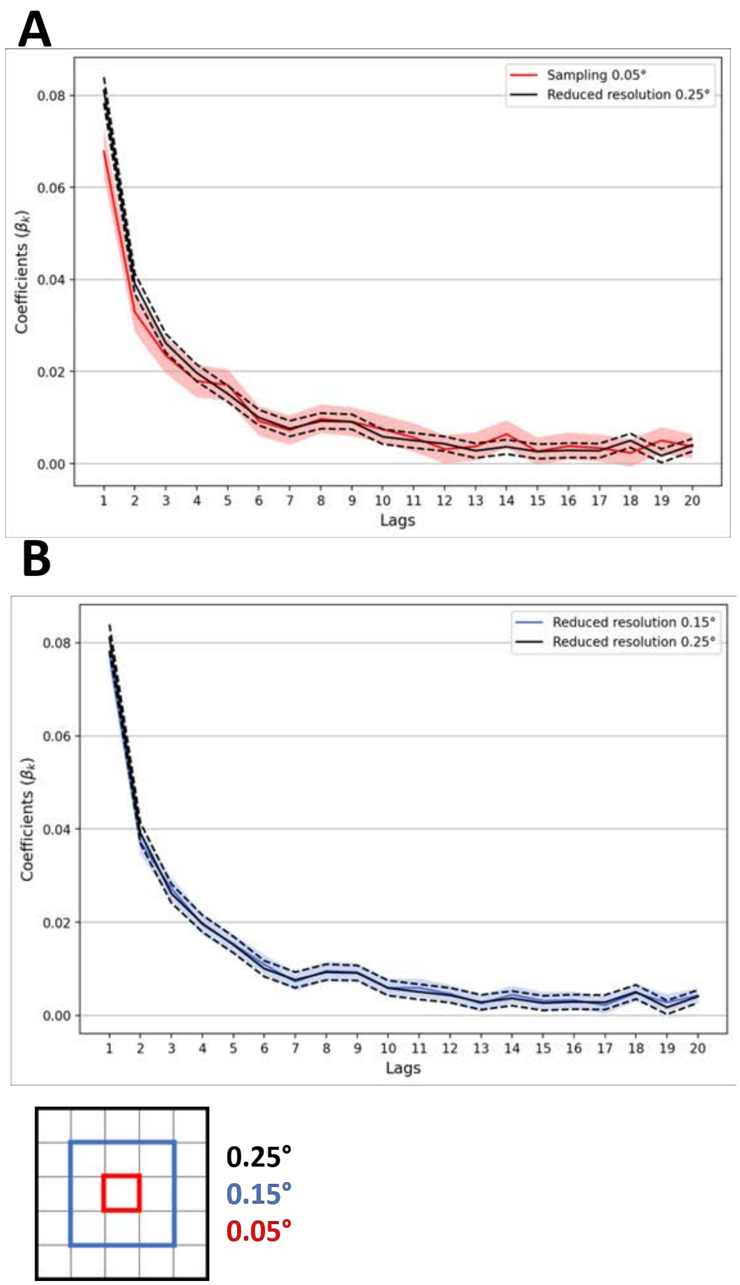
Note: This table shows the estimated coefficients of Expression (4). Standard errors in parentheses are clustered at the pixel level. Number of observations: 3,300,494. Number of clusters: 9,539.

Figure 1.2: Estimated spatial lags (β_k).



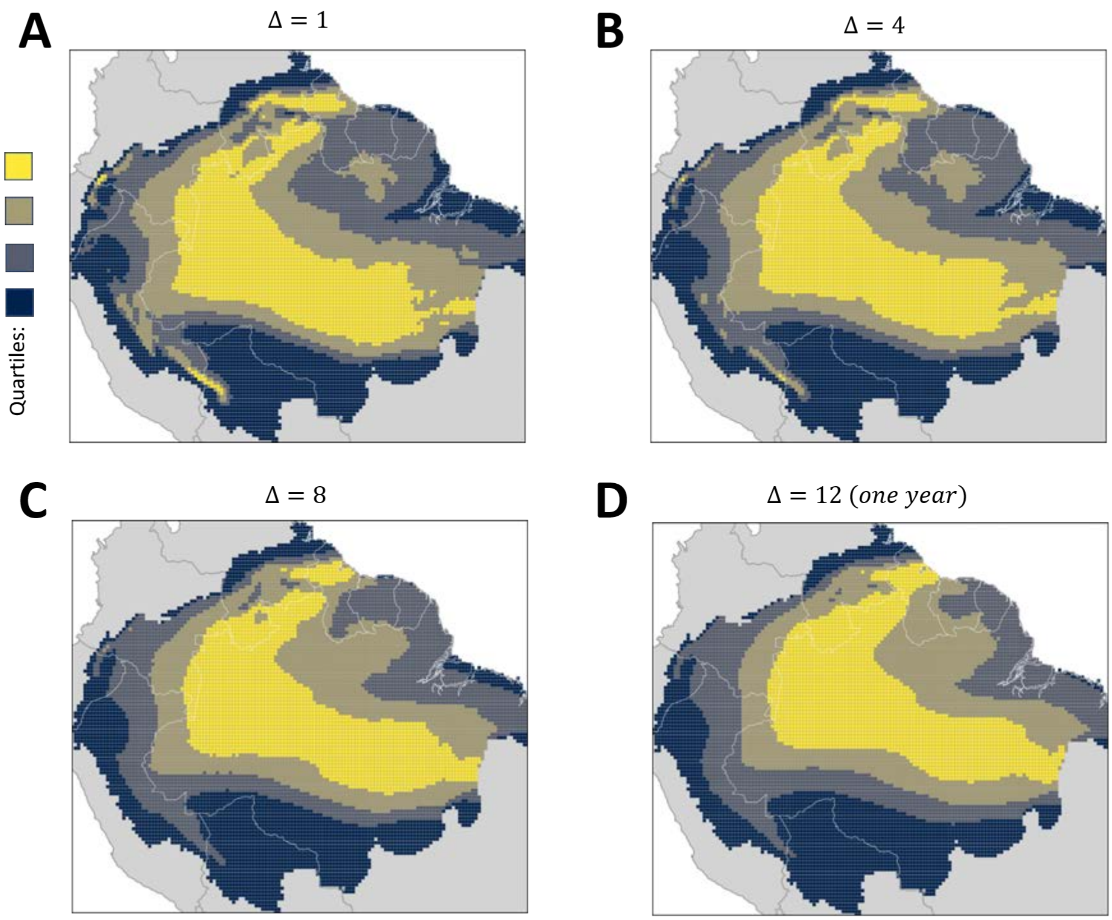
Note: This figure shows the estimated spatial lags (β_k) for each k between 1 and 60. The shaded gray area shows the 95% confidence interval with standard errors clustered at the pixel level.

Figure 1.3: Estimated spatial lags (β_k).



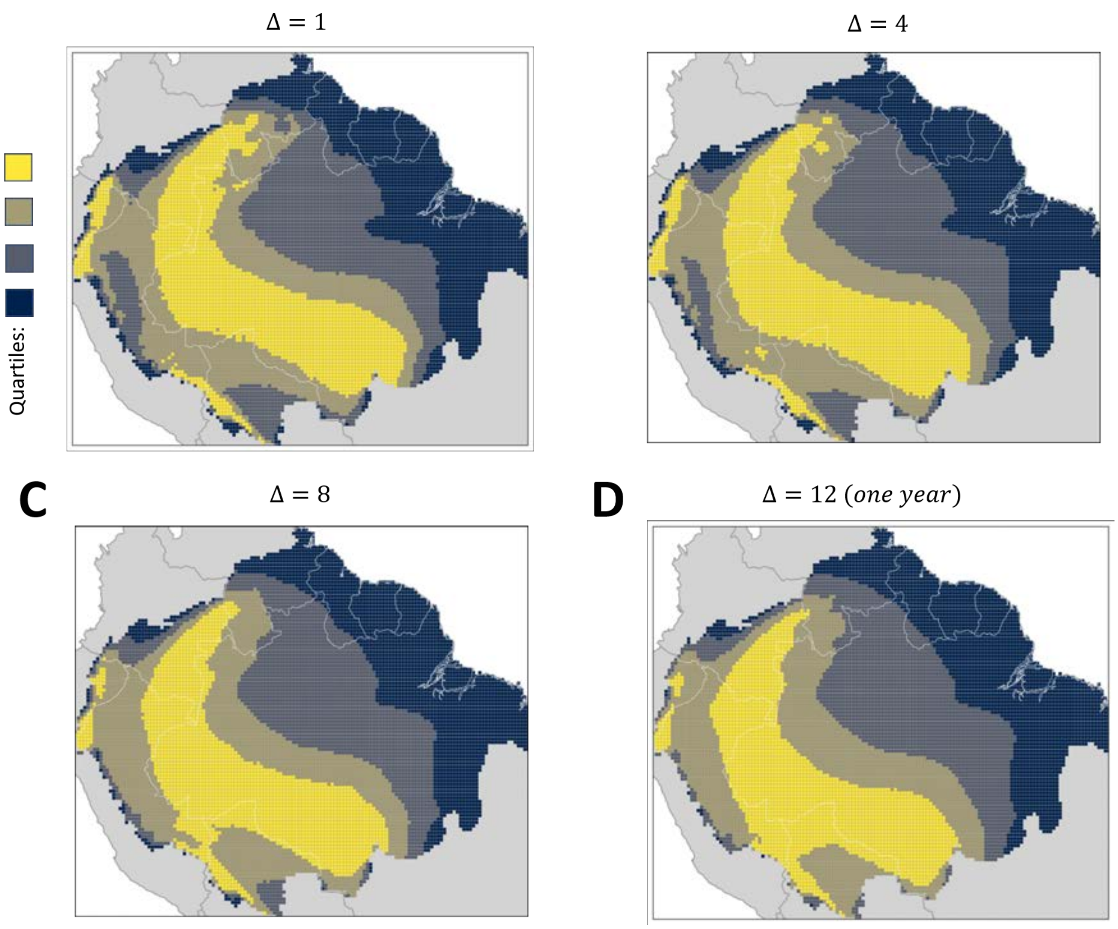
Note: This figure shows the estimated spatial lags (β_k) for each k between 1 and 20. In this exercise, instead of reducing the resolution of the LAI data to match the resolution of the wind data, we assign the LAI value (A) at the centroid of the grid (B) at the center of the grid with a reduced resolution of 0.15°. The results is shown by the red (A) and blue (B) lines, with the 95% confidence interval with standard errors clustered at the pixel level shown in shaded color. For comparison, the black line and dotted lines show the estimates and 95% confidence interval of Figure 2C, our preferred specification. Sampling by the centroid (A) results in an estimate of α of 0.075 (std. 0.001). The reduced resolution of 0.15° results in an estimate of α of 0.166 (std. 0.001).

Figure 1.4: Index of Criticality.



Note: This figure shows the total effect that a shock at a pixel would have, accounting for spatial feedback, for different time windows. As noted before, the decay of the spatial dimension α^{Δ} dominates the effect. Maximum values for each scenario are A:2.64; B:0.07; C:0.001, D:10^{-5}.

Figure 1.5: Index of Exposure.



Note: This figure shows the total effect that a pixel receives given a shock at every pixel in the forest, accounting for spatial feedback for different time windows. As noted before, the decay of the spatial dimension α^{Δ} dominates the effect. Maximum values for each scenario are A:2.57; B:0.07; C:0.001, D:10^{-5}.