*Geophysical Research Letters*

Supporting Information for

**Deforestation-driven increases in shallow clouds are greatest in drier, low-aerosol regions of Southeast Asia**

Gabrielle R. Leung1, Leah D. Grant1, and Susan C. van den Heever1

1 Department of Atmospheric Science, Colorado State University, Fort Collins, CO

**Contents of this file**

Text S1

Figure S1

Figure S2

**Introduction**

This supporting information provides a more detailed description of the methods used and supplementary figures.

Text S1. Detailed methods

S1.1 Data sources

We used atmospheric data from the Moderate Resolution Imaging Spectrometer (MODIS) instrument onboard the Terra and Aqua satellites. Terra has an overpass time of 10:30AM/PM, while Aqua has an overpass time of 1:30AM/PM. The four MODIS products utilized were: cloud fraction (CF), cloud top height (CTH), aerosol optical depth (AOD), and precipitable water (PWAT). The cloud retrievals were available at ~1km horizontal resolution during both daytime and nighttime swaths (Platnick et al., 2017). We only use pixels where the cloud mask is confidently or probably cloudy. The aerosol data are available at ~3km horizontal resolution during daytime swaths using the Dark Target retrieval (Levy et al., 2013), and the moisture data are available at ~3km horizontal resolution during daytime swaths using the near-infrared retrieval (Kaufman & Gao, 1992). We obtained Level 2 swaths for the full time period available (Terra: 2001-2020, Aqua: 2003-2020), and then reprojected the data onto a common grid (with a grid spacing of 0.008º or ~1km at the equator) to calculate the annual averages for each overpass time.

In addition, we used forest cover data from the University of Maryland Global Forest Cover (GFC) dataset (Hansen et al., 2013). The GFC data were derived from 30m-resolution Landsat imagery, which is used to provide both the percent of land covered by forest in 2000 and the year in which a deforestation event occurred. We resampled the GFC data and reprojected it onto the same grid as the MODIS data to obtain annual forest cover at 1km spatial resolution for the entire Southeast Asia region (**Fig. 1**). By comparing year-to-year differences in forest cover for a given pixel, the forest loss could then be determined and deforestation events could be identified.

S1.2 “Difference-in-differences” statistical analysis

In order to understand the impact of deforestation on cloud properties, we needed to separate changes driven by the forest loss itself from changes driven by other sources of interannual variability using the “difference-in-differences” method (Crompton et al., 2021). First, across the two decades of data examined here, we identified any pixels which had undergone a deforestation event, defined as a >50% loss in forest cover from one year to the next. This definition intentionally excludes gradual changes in forest cover and focuses the analysis on dramatic changes in forest cover that are most likely to impact the cloud field on an annual timescale. It should also be noted that the changes in the cloud field are only considered in the years directly before and after the deforestation event (i.e., a given pixel can only be deforested one time), and thus exclude potential impacts from forest regrowth. Deforested pixels and any pixel within a 10km radius from them (**Figure S1**) were defined as *assessment pixels* that may potentially experience deforestation-driven changes in cloud properties. Such a change in a cloud property *C* for an assessment pixel during the year *t* is given by:

which is the temporal change in the given cloud property across the year before and after the deforestation event. We tested the sensitivity of these results to varying the radius for defining an assessment pixel between 10 and 25km and found there were no qualitative changes in our results.

Secondly, for each assessment pixel, we identified a matching set of *control pixels* that experienced similar synoptic weather forcing as the assessment pixel but did not directly experience the effects of deforestation. Control pixels were between 10-25km from a deforested pixel (i.e., not assessment pixels) and consisted of intact forest (>90% forest cover remaining and <2% accumulated forest loss since 2000). We tested the sensitivity of these results to varying distance definitions (up to 5km closer or further) and found there were no qualitative changes in our results. For the same deforestation event described above, the corresponding change in *C* for a control pixel is similarly written as:

Thirdly, we calculated the deforestation-induced changes in the cloud field. Both the control and assessment pixels will have similar changes in *C* driven by large-scale interannual variability, but only the assessment pixels will have a change in *C* driven by deforestation. The effect of deforestation 𝜀 is thus given by the “difference-in-differences”:

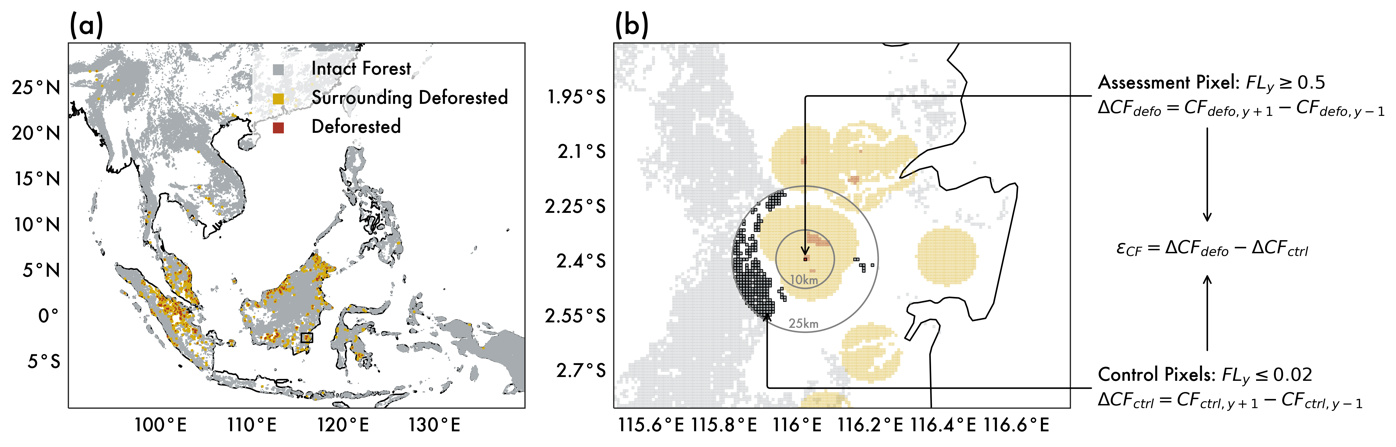
which we use as a metric of deforestation impacts on cloud properties.

Finally, we aggregated the millions of qualifying pixels across the entire region and time period as a function of forest loss. As the mean forest loss within a 1km radius of an assessment point increases (i.e., as its surface properties becomes more different from the control points), we would expect a magnification in the cloud response if these changes are being driven in a monotonic way by the deforestation itself. We calculated the mean forest loss within a 1km radius of each assessment pixel and grouped the pixels into 10 forest loss bins (ranging from 0 or no forest loss to 1 or total forest loss). For each bin, we calculated the bootstrapped estimate (sample size of 100 repeated 5000 times) of the mean and the 25th and 75th percent confidence intervals.

S1.3 Environmental analysis

To explore the environmental modulation of land-convection interactions, we further calculated the deforestation effect on cloud properties for different environmental subgroups. Specifically, we contrasted the impacts of high versus low precipitable water (PWAT) and aerosol optical depth (AOD). For each assessment point, we represented the nearby environment by calculating the mean value of PWAT and AOD within a 10km radius in the year that the deforestation event occurred. We tested the sensitivity of our results to other averaging radii (1km, 5km, and 10km) and other time periods (during the year of the deforestation event, the year prior, the year after, and the mean across all three years) for defining the environmental parameters, and found the trends presented in this paper are not sensitive to the averaging radii or time period.

Once these environmental parameters were calculated for each assessment pixel, we divided them according to their percentile values across all assessment pixels for the entire time period. In doing so, we separated the “low” and “high” categories of each environmental variable by taking the bottom and top quartiles, respectively. We then repeated the difference-in-differences calculation for the low and high categories separately.

**Figure S1.** Illustration of the “difference-in-differences” method for year 2004. (a) shows all forest pixels. Gray pixels are the potential control sample, defined as intact forest points (with total prior forest loss, FL2000-2004<0.02) that are more than 10km away from a deforested point. Red pixels are deforested areas (FL2004>0.5). Yellow pixels are those which do not satisfy the deforestation criteria but are less than 10km from a point which does. Together, the “deforested” and “surrounding deforested” pixels make up the assessment points. (b) shows a closer view of the boxed region in (a), with a single assessment pixel and corresponding control pixels highlighted in black outline. The gray range circles show the definition of the control pixels as being between 10 and 25km from the assessment pixel they are controlling for. The calculation of epsilon (deforestation effect) is described by the inset equations.

A map of different countries/regions

Description automatically generatedFigure S2. Spatial distribution of environmental parameters tested in the study. Spatial distribution of points assigned to the low/high (a) precipitable water (PWAT) and (b) aerosol optical depth (AOD) quartiles. Gray contours indicate elevations of 500, 1000, and 2000 m ASL to show topography.