

```

\usepackage{tabularray}
\usepackage{lineno}

\begin{document}

\modulolinenumbers[1]
\title{Geographic patterns of upward shifts in treeline vegetation across western North America,
1984-2017}

% \Author[affil]{given_name}{surname}

\Author[1]{gresham.joanna@gmail.com}{Joanna L.}{Corimanya} %% correspondence author
\Author[2,3]{Daniel}{Jiménez-García}
\Author[444]{Xingong}{Li}
\Author[1]{A. Townsend}{Peterson}

\affil[1]{Biodiversity Institute, University of Kansas, Lawrence, Kansas, USA, USA}
\affil[2]{Laboratorio de Biodiversidad, Centro de Agroecología y Ambiente, Instituto de Ciencias,
Benemérita Universidad Autónoma de Puebla, Puebla, Mexico}
\affil[3]{Laboratorio Nacional CONAHCYT de Biología del Cambio Climático, Mexico, Mexico}
\affil[4]{Department of Geography, University of Kansas, Lawrence, Kansas, USA}
\affil[4]{Department of Geography, University of Kansas, Lawrence, Kansas, USA}
%% The [] brackets identify the author with the corresponding affiliation. 1, 2, 3, etc. should be
inserted.

%% If an author is
dec{"originatingScript":"m2","payload":{"guid":"18b1454c-2578-4f50-8643-bb3652bec330f163d4",
", "muid":"7ba93509-301d-4860-8c4f-7372723a85b5ae67c8", "sid":"2a3b8aaf-416c-422d-80d6-5
492290ab7249cb585"}}{"originatingScript":"m2","payload":{"guid":"18b1454c-2578-4f50-8643-b
b3652bec330f163d4", "muid":"7ba93509-301d-4860-8c4f-7372723a85b5ae67c8", "sid":"2486161
4-596c-47a8-93e2-ae2baa2678a3be4b64"}}e{"originatingScript":"m2","payload":{"guid":"18b145
4c-2578-4f50-8643-bb3652bec330f163d4", "muid":"7ba93509-301d-4860-8c4f-7372723a85b5ae
67c8", "sid":"66a106ae-caf4-4f06-9631-141038251fc9e018d5"}}ased, please mark the
respective author name(s) with a dagger, e.g. "\Author[2,$\dag$]{Anton}{Smith}", and add a
further "\affil[$\dag$]{deceased, 1 July 2019}".

%% If authors contributed equally, please mark the respective author names with an asterisk,
e.g. "\Author[2,*]{Anton}{Smith}" and "\Author[3,*]{Bradley}{Miller}" and add a further affiliation:
"\affil[*]{These authors contributed equally to this work.}".

\runningtitle{TEXT}

```

\runningauthor{TEXT}

\received{}

\pubdiscuss{} %% only important for two-stage journals

\revised{}

\accepted{}

\published{}

%% These dates will be inserted by Copernicus Publications during the typesetting process.

\firstpage{1}

\maketitle

\begin{abstract}

Previous research has shown that (1) treelines are shifting upward in elevation on high mountain peaks worldwide, and (2) the rate of the upward shift appears to have increased markedly in recent decades. Because treeline elevational shift is a process manifested over broad scales of space and time, a particular challenge has been that of obtaining a broad-enough view of patterns of treeline shift to permit inferences about geographic and environmental patterns. What is more, intensive studies of treelines have been concentrated in North Temperate regions, such that little information is available about treeline shift patterns at lower latitudes. We have attempted to address this challenge by analyzing a long time series of vegetation indices derived from Landsat imagery obtained and analyzed via Google Earth Engine from the 1980s to the present. We sampled vegetation indices at points spaced every 100 m along 100 km transects radiating out from 115 high peaks across western North America (Canada to Central America), which means that we are sampling approximately every second or third pixel in the corresponding Landsat images. Considerable data preparation was necessary, including ending transects <2 km into closed forest, identifying current treelines via reference to Google Earth imagery, and consideration only of up to <1 km above treeline. Patterns that emerged were—as is well known—that treelines are generally higher at lower latitudes, but—previously unknown—that the magnitude of treeline shifts is nonrandomly distributed with respect to latitude, longitude, and their interaction. This analysis, via a broad-scale view of treeline shifts over almost 40 years and a geographic span of more than 40° of latitude, demonstrates that climate change effects and consequent treeline shifts are most dramatic in tropical regions where few or no detailed treeline studies have been and are being conducted.

\end{abstract}

\section*{Keywords}

treeline, elevational shift, climate change, long-term studies

\introduction %% \introduction[modified heading if necessary]

%hey this is to learn how to cite~\citep{cud17}.

The upper elevational limits of forests in mountain systems represent a fascinating and dramatic manifestation of distributional limitation at the species and community levels. As such, treeline phenomena have seen extensive analysis and discussion in the ecological literature: they are an important manifestation of the geographic ecology of ecosystems, and likely reflect important climate-related controls ~\citep{kullman98}. Numerous studies have been developed that aim to understand factors driving the location and possible shifts in treelines, with the general conclusion that treelines are determined by complex suites of factors ~\citep{cud17, korner98, holtmeier05, irl16, Grafius12, kienle23}. Whereas some researchers have concluded that treeline position can be distilled down to simple rules regarding seasonal mean ground temperatures ~\citep{korner04}, others have argued that treeline drivers are considerably more multidimensional and complex ~\citep{korner14, zhao15}. ~~In this study, we adopt Körner's (~\citeyear{korner2012} definition of elevational treeline, i.e., the uppermost elevation on a mountain slope at which upright woody plants (trees > 2m tall) can maintain a self-sustaining population. Above that limit, insufficient warmth (a too short or too cold growing season) prohibits the regular recruitment and survival of true tree forms, even if isolated, krummholz-like individuals may sporadically occur.~~ In this study, we adopt Körner's (~\citeyear{korner2012} definition of elevational treeline, i.e., the uppermost elevation on a mountain slope at which upright woody plants (trees > 2m tall) can maintain a self-sustaining population. Above that limit, insufficient warmth (a too short or too cold growing season) prohibits the regular recruitment and survival of true tree forms, even if isolated, krummholz-like individuals may sporadically occur.

Clearly, considerable complexity is involved in any attempt to characterize treeline phenomena. However, dendroecological approaches offer the useful possibility of obtaining establishment ages on an individual-tree basis across broad stands of trees at or near treelines ~\citep{grant11}. When treelines change, a key challenge is that of considering treeline shifts (e.g., elevational advance upward with warming climate) versus densification (e.g., sparse forest or scattered trees near treeline filling in with more trees, regardless of whether the upper limit of the trees changes or not) ~\citep{shi22}. Finally, treeline is a highly scale-dependent phenomenon, such that all of its qualities vary in importance and effect at different spatial extents and resolutions ~\citep{holtmeier17}.

From early in the discussions about the possibility that global climates would warm with increasing greenhouse gas concentrations ~\citep{lamarche84, john02}, the expectation has been that treelines would advance up mountain slopes as climatic controls relax at extreme elevations. Empirical evidence has been mixed, however, with some studies documenting what appears to be very rapid treeline advance ~\citep{peterson22}, and others finding no evidence

of overall tendency to change ~\citep{beloiu22}. One broad analysis found that treeline advance was faster in subarctic regions than in temperate regions ~\citep{lu21}, and another found that treelines experiencing stronger winter warming and with diffuse treeline forms were more likely to advance ~\citep{he23}.

Nonetheless, most of these previous broad-scale analyses of patterns of treeline advance in the face of warming climates have been based on datasets with strong inherent biases and significant gaps. That is, in largest part, treeline studies have been conducted in the North Temperate zone: examples of such biased analyses are many ~\citep{shi22,zhao15,korner98,lu21}. A few analyses have achieved a somewhat better balance of representation of treelines in the Tropics and in the Southern Hemisphere ~\citep{he23,hansson23,kienle23}. The concern, of course, is that such information gaps and biases in what information is available may blind researchers and their analyses to very real and important patterns in the global occurrence of the phenomenon of treeline advance.

Here, we address these important knowledge gaps about treeline dynamics in the face of warming climates globally over the past several decades. We assess the null hypothesis that magnitude of alpine treeline shifts is not related to a series of geographic features, such as latitude, longitude, distance to coast, and others. Specifically, to be able to assess treeline shifts on a continent-wide basis, we use a long time series of remote-sensing data to seek patterns in the magnitude of alpine treeline shifts across 115 high peaks scattered across western North America, ranging from Central America to southern Canada. We use vegetation index trends along transects radiating out from each peak in eight cardinal and sub-cardinal directions; the vegetation index approach has the advantage of “seeing” vegetative mass generally, in effect integrating over both treeline advance and densification of sparse, near-treeline forests ~\citep{feuillet20}. Of course, these broad-scale analyses are not a substitute for more detailed, field-based analyses, nor should vegetation-index-based assessments replace more fine-grained inspections of the actual geometry of treelines. Still, the result is a novel dataset from which we have derived several intriguing insights about geographic patterns in the magnitude of treeline elevational shifts.

\section{Methods}

\subsection{Mountain peak characterization}

Our aim was to characterize temporal changes in vegetation mass on a set of mountains that covered western North America. To that end, we chose to follow a comprehensive summary of high mountains worldwide ~\citep{peaklistUltraProminencesPage}, which is based on an effort to identify all mountains worldwide with at least a 1500 m prominence; the authors of that compendium (called the Ultras Project) researched all summits on Earth that meet this criterion, finding 1524 such peaks. From this worldwide dataset, we extracted the 354 mountain peaks located in North America (Panama to the Arctic). We used the coordinates of each peak in this dataset as a centerpoint, and plotted 8 transects in each of the cardinal and sub-cardinal directions extending out from that centerpoint (points were plotted and distances measured in meters using the WGS84 Special Mercator for Web Applications (~~EPSG:3857 datum: sphere,~~

~~delta WGS84: 0 0 0, ellipsoid: sphere, major s ax: 6378137.000, minor s ax: 6378137.000, origin long: 0, origin lat :0, origin X: 0, origin Y: 0, scale fac: 1.0, units: m, parameters: 0EPSG:3857)~~ projection to assure consistent distances among sampling stations. Transects were each initially 100 km long, with sampling stations every 100 m, so each transect initially included 1000 sampling stations.

We excluded from analysis all mountains that were forested to the peak, or that showed signs of anthropogenic modification at or around the peak ~~upon visual inspection of the region in Google Earth~~ upon visual inspection of the region in Google Earth. We also excluded peaks for which treelines were not associated clearly with the upper slopes of the peak, but rather were lower, extending just a bit up the valley walls ~~and thus likely represent latitudinal treelines as opposed to altitudinal treelines~~ and thus likely represent latitudinal treelines as opposed to altitudinal treelines; such low treelines were particularly common in central and northern Canada and Alaska, such that northern peaks were excluded. Given that, in eastern North America, only one peak (Mt. Washington, in New Hampshire) met our criteria, to avoid including a genuine spatial outlier in our analyses, we omitted that peak from analysis, thus focusing our analyses on the high peaks of western North America. At the end of this process, from the initial database, we had 120 peaks remaining as a basis for our analyses (\textbf{Figure \ref{fig:1}}).

```
\setcounter{figure}{0}
\begin{figure}[t]
\includegraphics[width=17cm]{figure01.png}
\caption{The 120 high mountain peaks analyzed in this study. Triangles represent individual mountain peaks used in our analysis. The 'X' symbol is Mt. Washington, which was removed from the dataset prior to analysis. This map was constructed using QGIS ver 3.38.2. ESRI physical basemap was used to create the map.}
\label{fig:1}
\end{figure}
\clearpage
```

In Google Earth Engine, we overlaid the transect sampling points on imagery from Landsat (1984--2017), and associated the values of the normalized difference vegetation index (NDVI) with each sampling point in the transect dataset. For this analysis, we focused on early (1984--1988) and late time periods (2013--2017) within the timespan of the Landsat dataset. We used NDVI data from the annual Landsat collection (\textnormal{Landsat/LT5_L1T_ANNUAL_NDVI}, \textnormal{Landsat/LE7_L1T_ANNUAL_NDVI}, and \textnormal{Landsat/LE8_L1T_ANNUAL_NDVI}) in Google Earth Engine. We used the pre-processed LANDSAT_LT5_L1T_ANNUAL_NDVI collection in Google Earth Engine, which provides annual NDVI composites derived from Level-1 terrain-corrected Landsat 5 reflectance images (including cloud masking and quality assurance). Unfortunately, this collection is now deprecated in GEE. To ensure full transparency, our scripts for reproducing the NDVI computation from the original Landsat reflectance data are publicly available ~~link to scripts~~ on Github. Detailed information about the original dataset can be found in the Earth Engine Data Catalogue ~\citep{USGS17}: Calibration procedures and validation

methods for this collection are described by Chander et al., ~\citeyear{chander2009}-. We generated a composite for each year from the available Landsat images, and extracted NDVI values for each year via a mean reducer. We then inspected each transect of each peak individually by overlaying the point data on the Google Satellite fine-resolution data product, using the GIS capabilities of QGIS (version 3.2).

A key step was that of choosing the sampling station on each transect that corresponded to treeline, as follows. Descending from each peak (using the Google Satellite data layer in QGIS) along each transect, we identified the sampling station that most closely approximated the upper elevational limit of forest. That is, we ignored single, isolated trees, but rather identified the elevation at which forest became continuous, albeit in some cases sparse. For this sampling station, we set the field `TreesBegin` in the data table characterizing peaks to 1. ~~To focus attention on the region just above and just below treeline, we ended the transect after up to 20 additional sampling stations descending from the peak beyond treeline into the forest; however, we retained fewer than 20 sampling stations when any anthropogenic effects were noticeable upon visual inspection, or when the straight-line transect reached a valley bottom and began to ascend again. All further sampling stations beyond this point were removed from the dataset.~~

\subsection{Data refinement}

~~Values of NDVI and elevation were assigned to each sampling station via GIS overlay (“extract value to point”) operations. Values of NDVI and elevation were assigned to each sampling station via GIS overlay (“extract value to point”) operations. All subsequent data preparation was done in R (version 4.4.1) and QGIS (version 3.38.2). We cleaned the data that had been exported from Google Earth Engine by removing any “NA” and any missing values. We averaged the yearly NDVI values over the two time periods (1984-1988 and 2013-2017) to provide “before and after” comparisons that would be more immune to measurement error or other sources of year-to-year variation. random effects and error in NDVI measurements (e.g., from partial cloud cover). measurement error or other sources of year-to-year variation.~~

Our next goal was to calculate regression equations for individual mountains, slopes, and time periods, characterizing the negative-sloped relationship between elevation and ~~NDVI~~ ~~vegetation mass~~ `NDVI`. To this end, we transformed the data into a hierarchical nested list of lists; the dataset included 120 mountain peaks, each of which had 1-8 transects. Each transect had the two averaged year groups of NDVI data, for a total of 1848 distinct combinations of peak, transect, and year group; some transects were removed entirely based on the criteria listed above (~~Section 2.4.2~~ `Section 2.4.2`). In our analyses, we included only NDVI measurements from “stations” that were in relatively close proximity to treeline. That is, we included at least the last 10 stations. If twice the number of stations after the manually identified treeline to the transect edge (i.e., the furthest measured station downslope) plus one (to explicitly account for the station representing treeline itself) exceeded 10, we used this greater number of stations instead. This approach ensured we captured sufficient data from both sides of the treeline and minimized ~~the effect of terrain variability~~ `geographic noise`, ~~the effect of terrain~~

variability ~~from sources from~~ sources such as small bare peaks, increasing the probability of detecting the true relationship.

We modeled the NDVI-elevation relationship ~~with NDVI as the response variable and elevation as the predictor variable~~ with NDVI as the response variable and elevation as the predictor variable to find ~~the~~ the best type of regression equation and, ultimately, the best approximation to the ~~true~~ true relationship between these variables (~~\textbf{Figure \ref{fig:2}}; see below~~) (~~\textbf{Figure \ref{fig:2}}; see below~~); these models allowed us to associate NDVI and treeline elevation for calculation of our final response variables: ~~change in elevation and change in NDVI~~: change in elevation and change in NDVI. (~~\textbf{Figure \ref{fig:2}}; see below~~). To this end, we calculated three types of regressions on each data frame (linear, reciprocal-linear, and reciprocal-quadratic) to assess which model shape best describes the NDVI-elevation relationship. The three models were compared via the Akaike Information Criterion (AIC; \citealt{Akaike2011}) for each peak, transect, and time period. As all 1848 of these ~~NDVI-elevation relationships~~ NDVI elevation relationships were best described by a linear model we retained only linear regression equations for subsequent analyses. We excluded transects for which the regression equation was not statistically significant (~~regressions were considered significant at $\alpha = 0.05$~~) or for which the regression slope was positive. ~~We used $\alpha = 0.05$ as the threshold for statistical significance in all regressions.~~ We used $\alpha = 0.05$ as the threshold for statistical significance in all regressions. ~~This criterion removed 688 of 1848 transects, leaving 1160 transects for analysis. Finally, since our goal was to create temporal comparisons between the two time periods~~ between the two time periods, we also removed any transects for which regressions for either time period did not meet ~~the~~ the criteria ~~outlined above~~ outlined above; this filter removed another 202 transects from analysis. The final dataset thus included 958 transects on 115 peaks.

\begin{figure}[t]

\includegraphics[width=17cm]{figure02.png}

\caption{Map showing continentwide patterns of regression slopes relating NDVI to elevation for each peak, averaged across the 1-8 transects available for each peak, for the 2013-2017 time period. Red circles represent a positive slope (excluded from final analysis), and yellow circles represent a negative slope. The size of the circles coincide with the ~~magnitude of the~~ magnitude of the absolute value of each slope calculation. The 'X' symbol is Mt. Washington, which was removed from the dataset prior to analysis. This map was constructed using QGIS ver 3.38.2. ESRI physical basemap was used to create the map.}

\label{fig:2}

\end{figure}

\clearpage

The goal in these analyses was to calculate change in treeline elevation for use as a response variable in continent-wide models. ~~Alpine treeline position represents a bioclimatic threshold: trees cannot form self-sustaining closed canopy stands above it because low temperatures and a short growing season limit carbon assimilation and wood formation~~ \citep{korner2012, holtmeier09}. In turn, shifts in treeline elevation over time serve as a direct indicator of how local thermal regimes and associated growing season lengths are changing on the landscape

~~~\citep{korner2012,holtmeier05}. By modeling the change in treeline elevation, we capture how climate warming and other environmental drivers are pushing the arboreal “limit” upslope. Alpine treeline position represents a bioclimatic threshold: trees cannot form self-sustaining closed canopy stands above it because low temperatures and a short growing season limit carbon assimilation and wood formation ~\citep{korner2012, holtmeier09}. In turn, shifts in treeline elevation over time serve as a direct indicator of how local thermal regimes and associated growing-season lengths are changing on the landscape ~\citep{korner2012,holtmeier05}. By modeling the change in treeline elevation, we capture how climate warming and other environmental drivers are pushing the arboreal “limit” upslope. To this end, we inserted the elevations at our manually selected treeline position into the 2013-2017 NDVI ~~linear-linear~~ regression equations to calculate the NDVI values ~~associated with manifested~~ associated with at ~~present-day-present day~~ treeline in the recent time period. We then inserted that calculated NDVI value into the 1984-1988 regression equations to obtain an estimate of treeline elevation (i.e., we sought the elevation with the same 1984-1988 NDVI value as present-day treeline on that slope of that mountain; \textbf{Figure \ref{fig:3}}). Finally, we subtracted the 1984-1988 elevation values from the 2013-2017 elevation values to obtain an estimate of the change in treeline elevation over the broad temporal span of this study.~~

\begin{figure}[t]

\includegraphics[width=12cm]{figure03.jpg}

\caption{Example of a high mountain (Cerro de la Malinche, Tlaxcala, Mexico) and inferences deriving from it regarding position of treeline through time. Top panel: View of the mountain in Google Earth, with 8 transects radiating out from the peak in cardinal and subcardinal directions. White dots indicate stations at which NDVI values were sampled through time; purple stars indicate the position of treeline identified visually. Bottom panel: dark red points and lines show the NDVI-elevation relationship in the 1980s; blue points and lines show the same relationship in the 2010s. In one example (northward transect), the elevation of treeline observed for 2013-2017 (3960 m) was used to identify a treeline NDVI threshold (0.3135), which was in turn used to identify a likely elevation (3448 m) of the same NDVI level for 1980s conditions. Background of top panel is from Google Earth©.}

\label{fig:3}

\end{figure}

\clearpage

We also calculated a second, simpler response variable, which was simply change in NDVI at the 2013-2017 treeline. ~~In high-elevation contexts, an upward trend in NDVI within elevation bands at the present day treeline signals increased tree recruitment, shrub encroachment, or earlier green-up ~\citep{harsch09,rupp01}. Thus, by computing the change in NDVI over our study period, we capture a functional or “greenness” dimension of treeline dynamics that complements the structural dimension (change in elevation). In other words, even before trees form closed canopies, small shrubs or seedlings may begin to photosynthesize more vigorously, which will manifest as change in NDVI.~~ In high-elevation contexts, an upward trend in NDVI within elevation bands at the present day treeline signals increased tree recruitment, shrub encroachment, or earlier green-up ~\citep{harsch09,rupp01}. Thus, by computing the change in NDVI over our study period, we capture a functional or “greenness” dimension of treeline



dynamics that complements the structural dimension (change in elevation). In other words, even before trees form closed canopies, small shrubs or seedlings may begin to photosynthesize more vigorously, which will manifest as change in NDVI. To this end, we inserted the manually located 2017 treeline elevation into the two regression equations for that mountain and slope. This resulted in NDVI values at a particular elevation (i.e., recent treeline) for 2013-2017 and 1984-1988 for each peak and direction. We subtracted the 1984-1988 values from the 2013-2017 values to obtain the change in treeline NDVI. A more positive value for change in NDVI indicates an increase in NDVI between 2013-2017 and 1984-1988.

Finally, we assembled a suite of independent variables that may be of interest as possible drivers of variation in rates of treeline shift. We included (1) the number of stations in the transect below treeline (as a potential confounding factor), (2) cardinal direction of the transect, (3) latitude, and (4) longitude, all of which could be derived from the original data about each transect and peak in the analysis. We also calculated (5) the distance to the closest coastline in meters, based on the coastline corresponding to official maritime boundaries ~\cite{flanders19}. We built a raster file that contained the distance to the closest coastline for each 1.53 km ( $\sim 2.5'$  pixels). We then added these distance values to the data table for the transect sampling points using the point sampling tool in QGIS.

#### \subsubsection{Model selection}

To understand which of the above independent variables likely drives variation in rate of treeline elevational shifts, we used an iterative stepwise model selection process. We selected the model that best describes western North American geographic treeline elevational shift patterns using AIC. We explored ~~two~~<sup>three</sup> statistical models to ensure that the final model would ~~best explain geographic variation in treeline dynamics~~<sup>be robust to spatial autocorrelation</sup>~~best explain geographic variation in treeline dynamics~~. First, we built 1868 linear mixed models, each of which contained a random effect of 'Peak ID' to account for variability in local landscape characteristics. Second, we constructed 1868 spatial mixed models ~~using the R package 'spaMM'~~<sup>using the R package 'spaMM'</sup> in which we specified Matérn random effects to account for spatial autocorrelation by capturing the spatially structured variation in treeline elevation that is not explained by the fixed effects ~\cite{spamm}. These models were fitted using restricted maximum likelihood.

~~We attempted a third model set using principal coordinates of neighbor matrices (PCNM), which generates spatial eigenvectors based on geographic distances between sampling locations ~\cite{BORGARD200251}. We used the first two PCNM axes as fixed effects, instead of latitude and longitude, to explain spatially structured variation in treeline elevation change, with a random intercept for each peak. However, the inherent clustering of our data owing to very similar latitude and longitude values for the sampling points associated with individual peaks, and the fact that PCNM creates matrices on a row by row basis, led the weights across all principal coordinates to have a value of one. In the end, results under this approach were not interpretable with regards to geographic patterns, which was the aim of this study, so we did not pursue this approach further.~~

For the first two model sets (total 32 models), the response variable was the change in treeline elevation between the two time periods. We produced a second array of models, parallel to the first, in which we used change in treeline NDVI as the response variable. All other model characteristics were the same as for the models based on change in treeline elevation.

For all of the models described above, the fixed effects were different combinations of the independent variables: distance to coast, number of stations after treeline, cardinal direction of slope, latitude, and longitude, as well as the interaction between latitude and longitude. The models ranged in complexity, but we always included latitude and longitude. We compared all 32 models in an AIC table, as the response variable was constant and all models were fit by REML. We assessed significance by checking whether the 95% confidence interval of each fixed effect overlapped zero ~~~\citep{browne1979}~~. We considered results for which confidence intervals did not overlap zero to be significant. ~~Our dataset construction and analysis steps are summarized in a diagram for clarity below ( \textbf{Figure \ref{fig:4}})~~ Our dataset construction and analysis steps are summarized in a diagram for clarity below ( \textbf{Figure \ref{fig:4}})

```
\begin{figure}[t]
\includegraphics[width=17cm]{figure04.png}
\caption{Panel (a) diagrams the steps taken to 1. Characterize mountains, 2. Clean the data in preparation for analysis, and 3. Select models. In panel (b), the hierarchical structure of our dataset is conceptually illustrated.}
\label{fig:4}
\end{figure}
\clearpage
```

```
\begin{figure}[t]
\includegraphics[width=17cm]{figure04.png}
\caption{Panel (a) diagrams the steps taken to 1. Characterize mountains, 2. Clean the data in preparation for analysis, and 3. Select models. In panel (b), the hierarchical structure of our dataset is conceptually illustrated.}
\label{fig:4}
\end{figure}
\clearpage
```

## \section{Results}

### \subsection{Generalities about Treelines}

Treeline locations were non-random in a number of ways. On average, across all mountain peaks in our analyses, treeline was located at 2433 m. However, treeline position varied systematically, in that a significant relationship existed between treeline and latitude: tropical treelines averaged 3177 m, whereas temperate-zone treelines were lower, at 2244 m. As such,

all subsequent analyses in this study needed to be conditioned on the geographic complexity underlying the phenomenon of treeline.

#### \subsection{Change in Treeline Elevation}

Treelines have been changing, even over the relatively short, 30-40-year timespan of this study. Indeed, treeline shifts among the western North American peaks in this study had a mean overall shift of 20.2 m upslope, though the mean absolute shift (positive or negative) was 240 m. The distribution of change values ranged from 165 m downslope to 127 m upslope.

For the multivariate models relating change in treeline elevation to environmental drivers, we calculated the best-fit models for the linear mixed models and spatial mixed models using AIC and the coefficient of determination ( $R^2$ ). We calculated the marginal and conditional  $R^2$  values for linear mixed models and a pseudo  $R^2$  value for the spatial mixed models. The best linear mixed model included number of stations after treeline, direction of transect moving away from the peak, latitude, longitude, and the interaction between latitude and longitude as fixed effects, with mountain peak name as a random intercept (\textbf{Table \ref{tab:tableA1}}). From our candidate set of spatial mixed models, the best fit included only latitude as a fixed effect, with a Matèrn random effect structure (\textbf{Table \ref{tab:tableA2}}). When comparing all models and the two best fitting models from the linear and spatial analyses, the spatial mixed model was best overall (\textbf{Tables \ref{tab:tableA3} \& \ref{tab:table1}}).

```
\begin{table}[ht]
  \input{best_aic_elevation.tex}
  \vspace{0.25cm}
  \caption{AIC table comparing the best linear mixed model and the best spatial mixed model
from their respective comparisons, which had change in treeline elevation as the response
variable. There were 2 models in this comparison.}
  \label{tab:table1}
\end{table}
```

The best spatial mixed model, which was also the best model overall, showed that change in treeline was not significantly related to the only fixed effect, latitude ( ~~$P_{\text{pseudo-}R^2} = 0.4512$~~ ) ( $P_{\text{pseudo-}R^2} = 0.4512$ ). This model was fit using a Gaussian random effect with a Matèrn correlation structure. The smoothness parameter ( $\nu$ ) was estimated at 0.398, indicating a moderate degree of spatial continuity in treeline elevation changes. The range parameter ( $\rho$ ) was 0.00466, suggesting that spatial correlation between observations declines sharply over very short distances. The variance of the spatial random effect ( $\lambda$ ) was estimated at 3,651,000, highlighting substantial spatial variation in the data. The residual variance ( $\phi$ ) was 64,159, representing unexplained variability after accounting for spatial effects (\textbf{Table \ref{tab:table2}}).

```
% m17_spamm, best spatial mixed model with elevation
\begin{table}[ht]
\input{model_summary_table_elevation_m17_spamm.tex}
\hspace{0.25cm}
\caption{Model summary of the top spatial mixed model. Fixed and random effect outputs are
shown. The response variable for this model was the change in treeline elevation. Significance
would bewould be denoted by bold text and was assessed by observing whether or not the
confidence interval overlapped zero. This model found no significant relationships. This model
found no significant relationships. }
\label{tab:table2}
\end{table}
\clearpage
```

The less optimal best linear mixed model can be explored as well: it showed a significant relationship between change in treeline and latitude, longitude, and the interaction between latitude and longitude ( ~~$Conditional R^2 = 0.6887$ ,  $Marginal R^2 = 0.2815$~~ ) ( $Conditional R^2 = 0.6887$ ,  $Marginal R^2 = 0.2815$ ). Change in treeline elevation was significantly higher at lower values of latitude ( $\beta = -100.6$ , 95% CI = [-155.1, -46.29],  ~~$Conditional R^2 = 0.6887$ ,  $Marginal R^2 = 0.2815$~~ , \textbf{Table \ref{tab:table3}; Figure \ref{fig:545}a}). The relationship between change in treeline elevation and the interaction between latitude and longitude were also significantly negative ( $\beta = -0.8418$ , 95% CI = [-1.345, -0.3413], \textbf{Table \ref{tab:table3}}): as longitude increases eastward, effects of latitude on treeline shift become more negative, suggesting a complex spatial relationship between these geographic variables and treeline dynamics. Longitude alone also had a significant positive relationship with change in treeline elevation ( $\beta = 36.21$ , 95% CI = [13.00, 59.52], \textbf{Table \ref{tab:table3}; Figure \ref{fig:545}b}). This result indicates that mountain treelines further east in North America (farther from the Pacific Coast) have more drastic temporal changes in their treeline elevations compared to the more western mountain treelines in our study.

```
\begin{table}[ht]
\input{model_summary_table_elevation_m8.tex}
\hspace{0.25cm}
\caption{Model summary of the top linear mixed model. Fixed and random effect outputs are
shown. The response variable for this model was the change in treeline elevation. Significance
is denoted by bold text and was assessed by observing whether or not the confidence interval
overlapped zero.}
\label{tab:table3}
\end{table}
\clearpage
```

```

\begin{figure}[t]
\includegraphics[width=15cm]{figure0545.png}
\caption{Summary of univariate relationships between treeline elevational shifts and latitude and longitude. Panel (a) shows latitude on the x-axis, while panel (b) shows longitude on the x-axis. Regression lines for both panels are denoted in black. Note that the interaction term between these two independent variables is also statistically significant.}
\label{fig:545}
\end{figure}
\clearpage

```

#### \subsection{Change in Treeline NDVI}

As with the previous response variable, we fit a series of linear mixed models and spatial mixed models with a Måtern random effect structure for change in treeline NDVI as a response variable, and compared the resulting models via AIC, both individually and in totality. The top linear mixed model (~~Marginal  $R^2$  = 0.3300, Conditional  $R^2$  = 0.7349~~) (Marginal  $R^2$  = 0.3300, Conditional  $R^2$  = 0.7349) and the top spatial mixed model had only latitude as predictor variables when compared only to models of their respective type (\textbf{Tables \ref{tab:tableA4} \& \ref{tab:tableA5}}). The best-fitting model when comparing all linear and spatial mixed models and when comparing the top models from the spatial mixed model and linear mixed model AIC tables was the spatial mixed model with the fixed effect of latitude (~~Pseudo- $R^2$  = 0.6067~~) Pseudo- $R^2$  = 0.6067) \textbf{Tables \ref{tab:tableA6} \& \ref{tab:table4}}).

```

\begin{table}[ht] % Force the table to appear exactly here
\input{best_aic_ndvi.tex}
\vspace{0.25cm}
\caption{AIC table comparing the best linear mixed model and the best spatial mixed model from their respective comparisons, which had change in treeline NDVI as the response variable. There were 2 models in this comparison.}
\label{tab:table4}
\end{table}

```

The best-fit linear mixed model revealed that change in treeline NDVI was significantly related only to latitude ( $\beta$  = -0.003303, 95% CI = [-0.004121, -0.002484], \textbf{Table \ref{tab:table5}, Figure \ref{fig:656}}). The negative slope of this relationship indicates that change in NDVI is greater at lower latitudes, indicating more greenness in the Tropics and Subtropics in more recent years.

```

\begin{table}[ht]
\input{model_summary_table_NDVI_m17}

```

\vspace{0.25cm}

\caption{Model summary of the top linear mixed model. Fixed and random effect outputs are shown. The response variable for this model was the change in treeline NDVI. Significance is denoted by bold text and was assessed by observing whether or not the confidence interval overlapped zero.}

\label{tab:table5}

\end{table}

\clearpage

Among the set of spatial mixed models, the top model concurred with the top linear mixed model. Latitude was again significantly negatively related to change in treeline NDVI ( $\beta = -0.003852$ , 95% CI =  $[-0.005047, -0.002705]$ , \textbf{Table \ref{tab:table6}}, Figure \ref{fig:656}), and no other variables had significant effects. The negative slope underlines the linkage between lower latitudes and more intense treeline movement. This model was the best performing overall out of all models with change in NDVI as the response variable that we tested. The smoothness parameter ( $\nu$ ) was estimated at 0.259, indicating moderate spatial continuity in the data. The range parameter ( $\rho$ ) was 1.006, suggesting that spatial correlation between observations diminishes rapidly over very short distances. The variance of the spatial random effect ( $\lambda$ ) was estimated at 0.00298, reflecting residual spatial variability in the data. The residual variance ( $\phi$ ) was estimated at 0.00107, representing the remaining variability not explained by the spatial random effect (\textbf{Table \ref{tab:table6}}).

\begin{table}[ht]

\input{model\_summary\_table\_NDVI\_m17\_spamm}

\vspace{0.25cm}

\caption{Model summary of the top spatial mixed model. Fixed and random effect outputs are shown. The response variable for this model was the change in treeline NDVI. Significance is denoted by bold text and was assessed by observing whether or not the confidence interval overlapped zero.}

\label{tab:table6}

\end{table}

\clearpage

\begin{figure}[t]

\includegraphics[width = 15cm]{figure0656.png}

\caption{Summary of the univariate relationship between NDVI at manually identified 2017 treeline elevation. Change in NDVI on the y-axis represents 2017 NDVI - 1984 NDVI. Latitude, which was significant in the models with change in treeline NDVI as the response variable, is shown on the x-axis. The regression line from the linear model of change in treeline NDVI and latitude is denoted by the black line.}

\label{fig:656}

\end{figure}

\clearpage



\conclusions[Discussion] %% \conclusions[modified heading if necessary]

\subsection{Overview}

This study represents a first broad-scope view of spatial patterns of temporal shifts in treeline elevation across a region. In that sense, it is novel, but has been limited by a significant number of data-related challenges: e.g., the necessity of eliminating the northernmost set of high peaks because treelines were not uniquely associated with individual peaks, as well as the removal of a number of peaks from consideration owing to positive slopes in the regression models relating NDVI to elevation. These complications point out the nascent nature of the endeavor and the need for quite a bit more exploration and experimentation ~~with effective methodologies~~.

Our results underlined some previous results, such as treelines occurring at higher elevations in the Tropics and Subtropics, and at lower elevations at higher latitudes ~\citep{korner98}. ~~This broad pattern makes sense, of course, if one thinks of the conditions present at the highest elevations—they are at the extremes of what is survivable for upright trees ~\citep{korner2021}. If treeline is set at least in part by hard physiological limits, and given global climate patterns and how they vary with latitude ~\citep{peterson2016}, then high-latitude treelines would necessarily be lower in elevation.~~

~~This broad pattern makes sense, of course, if one thinks of the conditions present at the highest elevations—they are at the extremes of what is survivable for upright trees ~\citep{korner2021}. If treeline is set at least in part by hard physiological limits, and given global climate patterns and how they vary with latitude ~\citep{peterson2016}, then high-latitude treelines would necessarily be lower in elevation.~~

More importantly and more novel, however, our results show clear associations between magnitude of treeline shift and latitude, such that tropical treelines have shifted upward faster than higher-latitude treelines in recent decades ~\citep{jimenezgarcia21}. ~~Such an effect has not been appreciated or reported previously, at least to our knowledge, but may relate to the greater physiological flexibility that may characterize tropical treelines: that is, high-latitude treelines may be fixed in elevation very clearly by hard physiological limits related to freeze tolerance ~\citep{korner2021}.~~ Such an effect has not been appreciated or reported previously, at least to our knowledge, but may relate to the greater physiological flexibility that may characterize tropical treelines: that is, high-latitude treelines may be fixed in elevation very clearly by hard physiological limits related to freeze tolerance ~\citep{korner2021}. This focus of treeline mobility in the tropical zone, unfortunately, coincides with significant knowledge gaps, given that the great majority of detailed studies of treelines and their dynamics ~~has~~ <sup>was</sup> been conducted on peaks at higher latitudes ~\citep{shi22,zhao15,korner98,lu21}.

Our results were suggestive of further effects, related to longitude and perhaps distance to coastlines; proximity to ocean has been underlined in past studies as important in determining treeline elevations at least ~\citep{hansson23}. That is, although we included a variable summarizing geographic distance to coastline, it did not have any significant effect in the best models. Rather, in some of the models that ranked among the best, effects of longitude were

indeed substantial. We suspect that this lack of clear effect of distance to coastlines may be related to the relatively minor representation of peaks close to coastlines in our dataset. ~~Only, we believe, further representation of peaks near to and far from coastlines will allow us to discern such effects; for this reason, we are in the process of expanding this study to all high peaks on Earth.~~

#### \subsection{Limitations}

The deepest concern regarding the analyses presented herein is, of course, the relatively short time span covered by the Landsat imagery ~~that we analyzed, with our analyses with our analyses~~ spanning just a bit more than three decades. This time span is, of course, what is available from remote-sensing data streams, as Landsat is among the deepest-time remote-sensing data sources available anywhere. ~~Even our relatively short time span of Landsat data, however, does cross the use of multiple sensors to produce the imagery, which may introduce noise into the analyses that we present herein ~\citep{vogelmann2016}.~~ Even our relatively short time span of Landsat data, however, does cross the use of multiple sensors to produce the imagery, which may introduce noise into the analyses that we present herein ~\citep{vogelmann2016}. The only remedy to this concern about time span is therefore to appeal to other data sources, such as aerial or ground-based photos ~\citep{jimenezgarcia21,peterson22}.

This study covered an impressive expanse in western North America, from 9.4°N in Costa Rica north to 54.1°N in southwestern Canada, and from the shores of the Pacific Ocean to the eastern edge of the Rocky Mountains in Colorado. However, this geographic span includes relatively fewer high mountain peaks in Mexico and Central America, at least compared with the northern peaks in the study; a further possible limitation of our work stems from the broad latitudinal gaps in northern Mexico. Finally, our inability to associate specific treelines with specific high peaks north of southernmost Canada meant that the highest-latitude peaks could not be included in the study. The former concerns can be remedied by broadening the area of study and analysis still further, perhaps globally, but the latter concern will remain complicated, as very high latitude peaks tend to be mostly above treeline, such that we do not see a way to create a peak-based analysis of those regions.

Finally, a concern could be that of anthropogenic effects that are not related to climate. That is, although we eliminated from consideration any peaks that had human activities visible at the peak or near treeline (e.g., agricultural activities), we could not control for changing practices of fire control, for example. In this sense, if fire control has been implemented or has become more effective over the past few decades, that—unrelated to climate—could elevate NDVI owing to reduced fire-based removal of vegetation. We hope that the broad variety of peaks included in this study will avoid any confounding effects of this concern.

#### \subsection{Conclusions and Next Steps}

The results of this study point rather dramatically to a major knowledge gap regarding high-elevation vegetation dynamics. That is, the bias of treeline studies away from tropical regions and towards temperate-zone and boreal-zone regions coincides—unfortunately—with

the most dramatic regions of treeline elevational shifts. As we have pointed out in previous contributions ~\citep{jimenezgarcia21}, treelines in the Tropics ~~and their dynamics and their dynamics~~ remain little-documented and poorly characterized.

At the same time, the results of this study and others ~\citep{peterson22,singh12} indicate that remote-sensing data streams are both relevant and informative, ~~and have been incorporated into many treeline studies ~\citep{garbarino2023}, and have been incorporated into many treeline studies ~\citep{garbarino2023}~~. Although the detail available in on-the-ground studies cannot be achieved, significant insight can indeed be gained from satellite-based observations and data streams, ~~particularly when multiple data streams are integrated ~\citep{garbarino2023}, particularly when multiple data streams are integrated ~\citep{garbarino2023}~~. As such, we are in the process of extending this approach globally ~~and using more diverse remote sensing data streams and using more diverse remote sensing data streams~~, in the hope of garnering additional useful insights into patterns of treeline change worldwide, and into processes that drive treeline change phenomena.

~~%% The following commands are for the statements about the availability of data sets and/or software code corresponding to the manuscript.%%~~

~~%% It is strongly recommended to make use of these sections in case data sets and/or software code have been part of your research the article is based on.%%~~

~~%%~~

~~\codedataavailability{All data and code are available on a public Github repository found at the following URL: <https://github.com/jocori/GeographicTreelinePatterns.git>} %% use this section when having data sets and software code available%%~~