Maximum height of mountain forests abruptly decreases above an elevation breakpoint

Aitor AMEZTEGUI[1,2*]; Marcos RODRIGUES[1,2]; Pere Joan GELABERT[1,2]; Bernat LAVAQUIOL[1]; Lluís COLL[1,2]

[1] Department of Agriculture and Forestry Engineering, University of Lleida, Lleida, Catalonia, Spain

[2] Joint Research Unit CTFC-AGROTECNIO, Ctra. Sant Llorenç de Morunys km.2, E-25280 Solsona, Catalonia, Spain

* Correspondence: Aitor Ameztegui, Department of Agriculture and Forestry Engineering, University of Lleida, Av. Rovira Roure 191 25198, Lleida, Catalonia, Spain

aitor.ameztegui@udl.cat

ABSTRACT

Canopy height is an excellent indicator of forest productivity, biodiversity and other ecosystem functions. Yet, we know little about how elevation drives canopy height in mountain areas. Here we take advantage of an ambitious airborne LiDAR flight plan to assess the relationship between elevation and maximum forest canopy height, and discuss its implications for the monitoring of mountain forests' responses to climate change. We characterized vegetation structure using Airborne Laser Scanning (ALS) data provided by the Spanish Geographic Institute. For each ALS return within forested areas, we calculated the maximum canopy height in a 20 x 20 m grid, and then added information on potential drivers of maximum canopy height, including ground elevation, terrain slope and aspect, soil characteristics, and continentality. We observed a strong, negative, piece-wise response of maximum canopy height to increasing elevation, with a well-defined breakpoint (at 1623 ± 5 m) that sets the beginning of the relationship between both variables. Above this point, the maximum canopy height decreased at a rate of 1.7 m per each 100 m gain in elevation. Elevation alone explained 63% of the variance in maximum canopy height, much more than any other tested variable. We observed species- and aspect-specific effects of elevation on maximum canopy height that match previous local studies, suggesting common patterns across mountain ranges. Our study is the first regional analysis of the relationship between elevation and maximum canopy height at such spatial resolution. The tree-height decline breakpoint holds an intrinsic potential to monitor mountain forests, and can thus serve as a robust indicator to appraise the effects of climate change, and address fundamental questions about how tree development varies along elevation gradients at regional or global scales.

Keywords: airborne laser scanning, canopy height, climate change, mountain forests, Pyrenees.

INTRODUCTION

Elevation is a strong handicap for the development of tree vegetation in mountain areas. This phenomenon is particularly evident at the treeline, i.e. the altitudinal limit of upright tree growth (Kullman, 2002; Körner, 2012). The treeline has received much attention in recent decades due to the interest in studying vegetation at the limit of its physiological capacity, and because its relation to temperature makes it an ideal early indicator of the responses of vegetation to climate change (Holtmeier & Broll, 2020). The limitation to tree development at the treeline responds to a common biological cause that applies across latitudes (Körner & Paulsen, 2004; Körner, 2012), and is related to the temperature and length of the growing season. Accordingly, Paulsen & Körner (2014) determined the position of the potential treeline – the natural climatic limit of tree growth without human influence – across the globe. In many mountain systems, however, this potential treeline does not overlap the actual one due to the long history of anthropic modifications (Harsch *et al.*, 2009; Ameztegui *et al.*, 2016).

We know much less about how elevation limits tree growth below the treeline. Does elevation pose a gradual limitation to the development in height of tree vegetation? Does it occur abruptly? In the latter case, from which elevation does it become a limit to the development of trees? These are questions that remain without a clear answer, despite the importance of canopy height as an indicator of forest biomass and carbon storage (Thomas *et al.*, 2008), productivity (Socha *et al.*, 2020), biodiversity and other ecosystem functions (Price *et al.*, 2011; Tao *et al.*, 2016).

Reasons behind this gap in knowledge include the difficulty of measuring tree or canopy height in the field, especially in remote places with complex reliefs (Wang *et al.*, 2019; Holtmeier & Broll, 2020). Traditional studies have addressed this issue through transects or field plots spread over relatively small areas (Payette *et al.*, 1989; Camarero &

Gutiérrez, 2004; Batllori & Gutiérrez, 2008). In recent years remote sensing data has opened the possibility to study forest ecosystems at much larger spatial extents (Coops, 2015; Gómez *et al.*, 2019; Blanco *et al.*, 2020). In particular, light detection and ranging (LiDAR) sensors can provide direct measurements of forest vertical structure over vast areas (Wulder *et al.*, 2012; Wang *et al.*, 2016), and have been employed to map forest canopy height, canopy cover or aboveground biomass (Lefsky *et al.*, 2005; Simard *et al.*, 2011; Wang *et al.*, 2016). To date, such maps have been based on large footprint, spaceborne full waveform LiDAR sensors, which offer global – yet incomplete – coverage at the expense of coarse spatial resolution (Wulder *et al.*, 2012). In this sense, steep slopes are known to broaden the waveform of large footprint LiDAR sensors, making canopy height estimation very problematic (and often unreliable) over mountainous regions (Duncanson *et al.*, 2010; Wulder *et al.*, 2012). In response, initiatives to map global canopy height have deliberately excluded many mountain regions (Wang *et al.*, 2016). Conversely, 'local' approaches have opted for *adhoc* Airborne Laser Scanning (ALS), which offers finer resolution (Wulder *et al.*, 2012; Mao *et al.*, 2019). ALS-based estimations achieve similar or even greater accuracy than field measurements (Duncanson *et al.*, 2010; Wang *et al.*, 2019), though they are more difficult to scale up towards regional or global analyses.

In this study, we aim to quantify the relationship between elevation and maximum canopy height for an entire mountain range (the Pyrenees), taking advantage of an ambitious ALS flight mission that covers the entire Spanish territory (PNOA). We specifically want to answer the following questions: (a) is there a critical elevation threshold from which the relationship begins to occur? (b) are the threshold and the strength of the relationship species-specific? c) is this relationship mediated by other physiographic variables such as aspect? This is the first study to approach these issues at such a broad geographical extent. This will allow us to identify whether the relationships and patterns observed are regionally consistent or dependent on local factors, and discuss the implications for the functioning and service provision of mountain forests, and its potential use to monitor the responses of mountain forests to climate change.

MATERIALS AND METHODS

Study area

Our study area was the Spanish Pyrenees, a range of mountains in southwest Europe that arranges from west to east in the border between France and Spain and covers 50,000 km^2, reaching more than 3,000 m at their highest summits (Fig. 1). The high altitudinal gradient as well as the influence of the Atlantic Ocean in the West and the Mediterranean Sea in the East strongly regulate the climate and therefore the type of vegetation (Fig. 1; Table S1.1). In the west, beech (*Fagus sylvatica* L.) becomes dominant at montane elevations (> 1000 m). In the Central and Eastern part, the climate becomes continental, and the foothills are mostly dominated by evergreen or marcescent oaks, while pines become predominant at higher elevations, and Atlantic species such as beech or fir (*Abies alba* Mill.) are restricted to the most humid valleys. Pines distribute in a clear elevation gradient according to their autoecology: Scots pine (*Pinus sylvestris* L.) is the most common species in the montane range (1300 to 1700 m). From here the main species is the Mountain pine (*Pinus uncinata* Ram. ex DC), which reaches up to 2200-2300 meters, and constitutes the upper limit of the forest (treeline) throughout the massif (Fig. 1c). It should be noted that in the Pyrenees, the treeline is generally well below its potential limit, which some authors place around 2400-2500 meters (Ninot *et al.*, 2008). This is due to the intense history of exploitation and pressure by man, who for millennia has cleared and burned the alpine forests to favor pasture for livestock (Ameztegui *et al.*, 2016).

ALS data source

We characterized vegetation structure using Airborne Laser Scanning (ALS) data provided by the Spanish Geographic Institute (IGN) via the National Plan for Aerial Orthophotography (PNOA). The datasets were captured between 2008 and 2011 (first PNOA flight) using a small-footprint discrete-return airborne sensor (Eastern Pyrenees Leica ALS50 and Central and Western Pyrenees Leica ALS60), operating at near infrared wavelength (1.064 µm) and ±28º scan angle from the nadir. The nominal point density in the study area is 0.5 point/m^2, with a vertical accuracy of ±0.2 m and a horizontal accuracy of ≤0.3 m. Data were delivered in 2 × 2 km tiles of preprocessed data points, in LAS binary file format (v. 1.2), with up to four returns recorded per pulse, and classified following the standards of the American Society for Photogrammetry and Remote Sensing (ASPRS). We selected, downloaded and processed the 3,140 tiles that intersected the limits of the Pyrenees according to the Global Observatory of the Pyrenees (OPCC).

Processing of ALS data, maximum canopy height and environmental variables

After filtering for those points classified as ground or vegetation (ASPRS classes 2, 3, 4 and 5), we normalized the point cloud by subtracting the elevation of a 5x5 meter digital terrain model (produced from the same ALS data) using the function *lasnormalize* as implemented in the *lidR* R package (Roussel *et al.*, 2020). Point cloud data were then aggregated to a 20-m grid cell using the *grid_metrics* function in *lidR*. To reduce the influence of sampling bias from possible errors in ALS surveys, and since we were interested in the maximum canopy height in each point of the territory, we retrieved for each cell in the grid the median of vegetation height returns above the 95th percentile in height (*top_height*), following Mao et al. (2019). We used the Spanish Forest Map 1:50,000 to restrict the analyses only to forested sites, and to assign each cell in the grid to a particular dominant forest species. Direct comparison of the ALS-derived height values with ground truth values derived from the Spanish National Forest Inventory (IFN; Direccion General para la Biodiversidad, 2007) is not possible due to methodological differences between both data sources. Instead, we compared the overall height distribution between the two data sources for each main tree species in the study area (Fig. S1.1). This allowed us to verify that our filters correctly excluded errors in the ALS surveys and assigned ALS data to the main species, producing reasonable *top_height* values for each of the species (Mao *et al.*, 2019).

We then added information on potential drivers of maximum canopy height – including physiographic, climatic and soil-related variables – to each cell in the grid. Ground elevation, terrain slope angle and aspect were obtained from the ALS-derived 5 m DTM. Aspect values were then reclassified into north (values between 315 and 45º) and south (between 135 and 225º); we also derived quantitative indicators of northness and eastness as the cosinus and the sinus of terrain aspect, respectively. We calculated the distance to the sea as a proxy for climatic continentality. Soil characteristics were obtained from the SoilGrids database (Hengl *et al.*, 2017), and included depth to bedrock and soil texture (proportion of clay, silt and sand). Finally, we derived climatic variables – mean annual temperature and annual precipitation – from the WorldClim database (Fick & Hijmans, 2017). All variables were resampled to the 20 x 20 m working resolution (see Fig. S1.2 to S1.13).

Statistical analyses

Since we were interested in modeling the response of the potential maximum development of tree vegetation, we aimed to remove from the dataset those cells in which, for many possible reasons, the tree vegetation has not reached its full potential height (poor soil, early stages, management and other disturbances, etc.). To do so, we grouped all the observations located above 1200 m into 500 equal interval elevation classes and selected, for each elevation class (2.6 m width each), only those cells with *top_height* values above the 95th percentile for that class (Coll *et al.*, 2011). The resulting variable was further referred to as the maximum canopy height (*max_height*). It represents the maximum height that vegetation can reach for a given elevation interval, and was termed as the dependent variable in our models. Since the choice of filtering percentile is somewhat arbitrary, and to assess the influence of this choice on our conclusions, we also built models in which the maximum canopy height was determined by selecting observations above the 90th percentile for each elevation class, and the results are shown in Supplementary Materials.

After visual exploration of the data, we assessed the relationship between elevation and *max_height* by fitting log-linear segmented regression models, an analysis in which the independent variable is partitioned into intervals and a separate regression is fitted into each interval. A segmented (or broken-line) relationship is defined by the slope beta coefficients (β1 and β2) and the breakpoints (ψ) where the slope of the relation changes (Equation 1).

$$\log(\max\_height) = \begin{cases} \alpha_1 + \beta_1 \cdot Elevation \; \forall \; Elevation \leq \psi \\ \alpha_2 + \beta_2 \cdot Elevation \; \forall \; Elevation > \psi \end{cases} \qquad (Eq. 1)$$

where α_1 and α_2 are the intercepts, and β_1 and β_2 are the slopes of the relationship below and above the breakpoint, respectively, whereas ψ is the breakpoint, i.e. the value of the independent variable where the slope of the relationship changes.

The model simultaneously yields point estimates and standard errors of all the model parameters, including the breakpoints. This allowed us to obtain not only the slope of the relationship between both variables (β_2), but also to determine the threshold at which this relationship commences (ψ, i.e. the breakpoint). We obtained the model parameters (β_1, β_2, and ψ) by bootstrapping, to avoid the effects of the huge sample size on the significance of the parameter estimators (White *et al.*, 2014), and to avoid the potential misspecification of the model due to spatial autocorrelation. Thus, we fitted 1000 models with a subsample of ≈10,000 randomly chosen data points (5,000 for calibration and 5,000 for validation) for each realization. We retrieved the mean and the standard deviation of the breakpoint position (ψ) and the slope before and after the breakpoint (β_1 and β_2) as parameter estimates, and the R-squared (R^2) and root mean standard error (RMSE) – calculated using the validation sample – as indicators of model performance. We assessed the support for the segmented regression model by comparing its performance to that of a non-segmented log-linear model via the differences in Akaike's Information Criterion (AIC) and R^2.

We also evaluated if additional variables could further explain the variation of maximum canopy height. To do so, we only kept observations above the elevation breakpoint as determined by the segmented model. Then we fitted univariate lineal models including as predictors elevation, soil characteristics – soil depth and texture –, climatic variables – mean annual temperature and annual rainfall – and physiographic variables (continentality, northness, eastness, and terrain slope). We also fitted a 'full model' that considered all the predictors. We investigated the change in model performance (R^2) of each univariate model, focusing on the comparison with the 'full model' and the univariate elevation model.

To assess the effect of aspect on the relationship between elevation and *max_height*, we repeated the analysis after segregating the sample into aspect classes. That is, we determined *max_height* per elevation class separately for north-facing and south-facing slopes, and then we fitted a segmented regression – as specified above – for each aspect class. We repeated the same procedure for each main tree species, splitting the sample according to the four main species in our dataset: *Pinus sylvestris* (49.6% of the laser returns above 1200 m), *Pinus uncinata* (27.5%), *Fagus sylvatica* (8.0 %), and *Abies alba* (3.9%). We did not include *Quercus* species in this analysis because – although abundant in the original sample – they were only present at low elevations (below 1500 m; Fig. 1). All the statistical analyses were conducted using R version 3.6.1 (R Core Team, 2018) and the package *segmented* (Muggeo, 2020), and the variables were log-transformed when needed to meet the assumption of normality.

RESULTS

Response of maximum canopy height to elevation

Estimations of tree canopy height varied between 11 and ca. 35 m (Table S1.1), and in general were higher at both ends of the Pyrenees, where the oceanic influence allows the presence of species from temperate forests such as beech or fir (Fig.2). There was a clear breakpoint in the response of maximum canopy height to elevation, which occurred at an elevation of 1623.3 ± 4.7 m (Fig. 3). Above this threshold, maximum canopy height decreased at a rate of 1.7 meters per each 100 m gain in elevation, whereas below this point, maximum canopy height was independent to elevation (Figure 3; Table 1). The results obtained across 1,000 bootstrap models were very consistent and showed a high robustness in the estimation of all the regression parameters (Fig. S1.14, see Methods for details on bootstrapping). The existence of the breakpoint is confirmed by the better fit of the stepwise model with respect to alternative linear and non-linear models (Table 1).

Elevation explained around 65% of the variability in maximum canopy height (Table 2). Climatic variables, particularly mean annual temperature, were also good predictors of maximum canopy height, but with less predictive ability than elevation (R^2 = 0.36 and RMSE = 2.9 m for temperature; 0.18 and 3.3 for annual precipitation). The effect of the other potential predictors was negligible, with the exception of soil depth (R^2 = 0.14; RMSE = 3.4; see the relationship of maximum canopy height with all explanatory variables in Fig. S1.15 – S1.18). However, when combining elevation with climatic variables or soil depth into a single model, the predictive ability remained similar to that of the univariate elevation model (Fig. 4). This suggests that the explanatory effect of climatic and soil-related variables is mainly due to their covariation with elevation (Pearson's r for mean annual temperature = -0.89, for precipitation = 0.74, for soil depth = -0.70).

The results using percentile 90 were very similar to those obtained for percentile 95. There was also a strong support for the existence of a breakpoint, which the models located at 1648 ± 6.4 m in elevation, i.e. only 25 m above the breakpoint detected for p95 (Fig. S1.19). Above this threshold, maximum canopy height decreased at a rate of 1.7 meters per each 100 m gain in elevation, identical to the rate detected for percentile 95. The goodness of fit of the percentile 90 models was in turn slightly poorer, with a mean R^2 = 0.47.

Aspect and species-specific effects of elevation on maximum canopy height

The drop in maximum canopy height with elevation was much more pronounced (-2.4 m/100 m vs. -1.3 m/100 m) for the northern slopes, where it also started at a slightly lower elevation (1657 vs. 1674 m; Fig. 5), although without significant differences in the breakpoint position (Table 1). The maximum height of the vegetation below the breakpoint was up to four meters taller on the northern aspects (27 vs. 23 m), but due to the faster decline in maximum height, trees become taller in southern orientations from elevation 2100 onwards (Fig. 5). Models adjusted for north-facing aspect trees showed a better fit than those for south slopes ($R^2 = 0.66 \pm 0.008$ vs. 0.40 ± 0.011), as well as more robust parameter estimation (Fig. S1.20-S1.22).

Fitting separate models for each species revealed an unequivocal breakpoint only for the two species growing in the subalpine belt: *Pinus uncinata* and *Abies alba*. For these two species, the model captured 60 and 87% of the variation in maximum canopy height, respectively, 20 points more than alternative linear models (Table 1). The relationship profile was quite similar to the one observed in the general analysis, with a slight decrease in height until a certain elevation threshold, above which the effect of elevation was much sharper, and twice as strong in *Abies* than in *Pinus* (Fig. 6).

In the two other species (*Pinus sylvestris* and *Fagus sylvatica*) the goodness of fit of the models indicates a much poorer ability to predict maximum canopy height with elevation (R^2 = 0.17 and 0.24), and stepwise models showed similar explanatory ability than log-linear models (Table 1). The breakpoint for these two species was detected at elevations at which their presence becomes testimonial (Fig. 1 and Fig 6). For *Pinus sylvestris*, the rate of decrease in maximum height before the threshold was the highest of all species, and the breakpoint did not occur until 1915 meters, which is close to the upper elevation limit of the species in the Pyrenees. Moreover, the log-linear model explained a similar amount of the variation in canopy height, which indicates low support for the existence of a breakpoint in the "maximum height-elevation" relation. In the case of *Fagus sylvatica*, parameter estimations show a bimodal distribution that indicates little support for the piecewise response (Fig. S1.23 – S1.25).

DISCUSSION

Maximum canopy height decreases with elevation only above a threshold

We observed a clear, negative, and piecewise response of maximum canopy height to increasing elevation. The piecewise and negative response was observed regardless of other factors such as slope, orientation or the dominant tree species. Interestingly, the relation between maximum canopy height and elevation is not gradual, but starts at a certain point, evidencing that elevation begins to restrain the height of trees further below the treeline, but above the trailing edge of species' range. Furthermore, the models fitted with tree heights above the 90th percentile yielded the same patterns as those above the 95th percentile, demonstrating that the relationship between canopy height and elevation holds irrespective of the height indicator chosen.

It is clear that ecological processes in mountains are not driven by elevation itself, but by the various factors that are correlated with it (e.g. temperature or rainfall) (Rumpf *et al.*, 2018; Körner & Spehn, 2019). Previous studies conducted on tropical and temperate biomes present strong evidence on the prominent role of water availability in canopy height (Klein *et al.*, 2015; Tao *et al.*, 2016; Zhang *et al.*, 2016), supporting the hydraulic limitation hypothesis that has also been verified at the individual tree level (Koch *et al.*, 2004; Moles *et al.*, 2009). In contrast, energy limitation was more important in boreal forests, where temperature is more limiting to trees (Zhang *et al.*, 2016). In our case, the decrease in maximum canopy height with elevation seems to be primary related to the adiabatic gradient, i.e. the decrease in temperature with elevation, rather than to changes in soil properties or water availability. These results suggest that energy limitation is also the most decisive factor in mountain environments, but the generality of this finding has yet to be confirmed in other mountain ranges. Notwithstanding, the observed humpshaped relationship seems to indicate that more than one variable may be involved, as already reported for boreal forests in Alberta (Mao *et al.*, 2019). Elevation, in any case, seems to integrate very clearly the various causes that govern the maximum height that tree vegetation can reach.

The height-elevation threshold as a tool to monitor climate change effects

The existence of a clear elevation threshold above which canopy height begins to diminish unveils the potential of this threshold as a monitoring tool to assess the effects of climate change on mountain forests at regional or global scales. Despite the attention devoted to the treeline as an indicator of vegetation responses to climate (Paulsen & Körner, 2014), many treelines have been historically modified by human activity, hampering the detection of climatic responses (Harsch *et al.*, 2009; Ameztegui *et al.*, 2016). In contrast, our threshold presents a series of advantages. By considering the maximum height of the vegetation along elevation gradients, the position of our limit is not sensitive to anthropic factors, and may thus be used as an alternative indicator to study the responses of species related to the changes in climate. Moreover, our indicator, based on tree growth, is likely to respond more readily to environmental changes, although this remains to be verified. In order for the treeline to move upwards, a series of processes must take place successively – seed production and dispersal, germination and establishment, survival, growth… – each depending on the climate in different ways. Many treelines are therefore very inert to change, and it is common to detect the effects of climate change as density changes below the treeline rather than as actual displacements of the limit itself (Camarero & Gutiérrez, 2004; Batllori & Gutiérrez, 2008). Future research may elucidate to what extent the indicator we present here responds to environmental changes more or less rapidly and accurately.

Several arguments support the use of elevation instead of climate variables as a monitoring tool. First, elevation seems to integrate well a variety of environmental variables – temperature, precipitation, soil properties – which often are correlated both among them and with elevation. Second, and more importantly, it is difficult to find climatic data with the required spatial detail, particularly in mountain areas. Although global datasets such as WorldClim (Fick & Hijmans, 2017) have made worldwide climatic data readily available, their quality is spatially unequal, and the density of climate stations commonly gets scarce precisely in mountain regions (Paulsen & Körner, 2014). For instance, only around 2% of the weather stations in Spain are located above 1,500 m (Gonzalez-Hidalgo *et al.*, 2020). This issue may not be so severe for global analyses, but becomes critical if mountain areas are to be targeted. Moreover, the rapid change of precipitation over short horizontal distances is often not well captured by climate databases, leading to potential biases in the estimation of its role as driver of ecological processes. Finally, most of these databases provide static information, which prevents their use to monitor the response of species to climate change.

Vegetation height decreases faster at northern-slopes and for subalpine species

Beyond 1600 meters, the maximum canopy height decreased at a rate of 1.7 meters for every 100 meters of increase in elevation, identical to the rate reported for a pine-dominated treeline in the Swiss Alps (Coops *et al.*, 2013). However, both the position of the breakpoint and the magnitude of the response were not general, but sensitive to factors such as species or slope orientation. The faster response of canopy height in northern aspects corresponds with their higher productivity at low elevations, and is also consistent with previous studies that locate the Pyrenean treeline at higher elevations on the southern slopes due to differences in thermal balance and dynamics in snow cover (Ninot *et al.*, 2008). Very similar patterns have also been observed in the Swiss Alps, where responses of vegetation height were also 70% faster on northern slopes, as observed here (Coops *et* *al.*, 2013). The similarity in patterns in both massifs suggests a common response that deserves further study.

Interestingly, the accuracy of the regression model was much higher for species typical of higher elevations (*Pinus uncinata* and *Abies alba*; Table 1). These species, which rarely grow below 1300-1500 m, mostly thrive in the Pyrenean subalpine belt, which is characterized by relatively wet but cold and windy climate. In such conditions, its growth potential in height is likely to be more limited by temperature changes associated to elevation than by soil- or precipitation-related variables (i.e. soil depth, water and nutrient availability), which can be more limiting at lower elevations. Accordingly, we only found a limited effect of soil characteristics on maximum canopy height, which can be explained by the covariation of the former with elevation. These results support previous studies at finer scales with seedlings of these species planted along elevation gradients (Ameztegui & Coll, 2013; Coll & Ameztegui, 2019). The relationship between elevation and maximum canopy height was much less clear for montane species, which suggests that the elevation constraint begins above the upper limit of these species, where only a few individuals can grow under favorable microclimatic conditions (only 3.5% of the observations for montane species were located above the breakpoint, as compared to 75% for *Pinus uncinata,* see Fig. 1). It remains to be determined whether climate change can alter this behaviour, favouring the upwards migration of these species and a greater dependence on elevation.

CONCLUSIONS

Our study is the first regional analysis of the relationship between elevation and maximum canopy height at detailed spatial resolution. By combining thousands of ALS observations, we were able to address fundamental questions about how tree development varies along elevation gradients, and evidence the existence of a solid piece-wise response. The breakpoint in the maximum canopy height – elevation relationship has the prospect of becoming a fundamental tool in the study of responses of mountain trees to environmental changes. Regular monitoring of its position, for example, can be used to assess the effects of climate change on mountain forests, isolating them from the effects – often misleading – of land use changes. The approach is also applicable in any mountain range, and may allow to test the generality of our findings. Finally, recent global monitoring initiatives such as GEDI (Global Ecosystem Dynamics Investigation), specifically designed for the study of vegetation, provide the first comprehensive global LIDAR dataset (Dubayah *et al.*, 2020; Valbuena *et al.*, 2020), and open a promising future for evaluating the relationship between canopy height and environmental and physiographical variables at the global scale.

DATA AND CODES AVAILABILITY STATEMENT

The data that supports the findings of this study is available in FigShare at the private link https://figshare.com/s/f5847dda38b702986a9c

ACKNOWLEDGEMENTS

This work was supported by the Spanish Ministry of Economy and Finance under "Juan de la Cierva" contracts to AA (IJCI-2016-30049) and MR (FJCI-2016-31090), and by the research grant program Ajuts UdL, Jade Plus i Fundació Bancària La Caixa to PJG and BL (Agreements 79/2018 and 87/2020 of the Governing Council of the University of Lleida).

DECLARATION OF INTEREST STATEMENT

We declare no potential competing interests.

REFERENCES

Ameztegui, A. & Coll, L. (2013) Unraveling the role of light and biotic interactions on seedling performance of four Pyrenean species under environmental gradients. *Forest Ecology and Management*, **303**, 25–34.

Ameztegui, A., Coll, L., Brotons, L. & Ninot, J.M. (2016) Land-use legacies rather than climate change are driving the recent upward shift of the mountain tree line in the Pyrenees. *Global Ecology and Biogeography*, **25**, 263–273.

Batllori, E. & Gutiérrez, E. (2008) Regional tree line dynamics in response to global change in the Pyrenees. *Journal of Ecology*, **96**, 1275–1288.

Blanco, J.A., Ameztegui, A. & Rodríguez, F. (2020) Modelling Forest Ecosystems: a crossroad between scales, techniques and applications. *Ecological Modelling*, **425**, 109030.

Camarero, J.J. & Gutiérrez, E. (2004) Pace and pattern of recent treeline dynamics: Response of ecotones to climatic variability in the Spanish Pyrenees. *Climatic Change*, **63**, 181–200.

Coll, L. & Ameztegui, A. (2019) Elevation modulates the phenotypic responses to light of four co-occurring Pyrenean forest tree species. *Annals of Forest Science*, **76**, 41.

Coll, L., González-Olabarría, J.R., Mola-Yudego, B., Pukkala, T. & Messier, C. (2011) Predicting understory maximum shrubs cover using altitude and overstory basal area in different Mediterranean forests. *European Journal of Forest Research*, **11**, 55.

Coops, N.C. (2015) Characterizing Forest Growth and Productivity Using Remotely Sensed Data. *Current Forestry Reports*, **1**, 195–205.

Coops, N.C., Morsdorf, F., Schaepman, M.E. & Zimmermann, N.E. (2013) Characterization of an alpine tree line using airborne LiDAR data and physiological modeling. *Global Change Biology*, **19**, 3808–3821.

Direccion General para la Biodiversidad (2007) *Tercer Inventario Forestal Nacional (1997-2007)*, Ministerio de Medio Ambiente, Madrid.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W. & Silva, C. (2020) The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing*, **1**, 100002.

Duncanson, L.I., Niemann, K.O. & Wulder, M.A. (2010) Estimating forest canopy height and terrain relief from GLAS waveform metrics. *Remote Sensing of Environment*, **114**, 138–154.

Fick, S.E. & Hijmans, R.J. (2017) WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, **37**, 4302–4315.

Gómez, C., Alejandro, P., Hermosilla, T., Montes, F., Pascual, C., Ruiz, L.A., Álvarez-Taboada, F., Tanase, M. & Valbuena, R. (2019) Remote sensing for the Spanish forests in the 21st century: a review of advances, needs, and opportunities. *Forest Systems*, **28**, 001.

Gonzalez-Hidalgo, J.C., Peña-Angulo, D., Beguería, S. & Brunetti, M. (2020) MOTEDAS century: A new high-resolution secular monthly maximum and minimum temperature grid for the Spanish mainland (1916–2015). *International Journal of Climatology*, **40**, 5308–5328.

Harsch, M.A., Hulme, P.E., McGlone, M.S. & Duncan, R.P. (2009) Are treelines advancing? A global meta-analysis of treeline response to climate warming. *Ecology Letters*, **12**, 1040–9.

Hengl, T., Jesus, J.M. de, Heuvelink, G.B.M., Gonzalez, M.R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S. & Kempen, B. (2017) SoilGrids250m: Global gridded soil information based on machine learning. *PLOS ONE*, **12**, e0169748.

Holtmeier, F.-K. & Broll, G. (2020) Treeline Research—From the Roots of the Past to Present Time. A Review. *Forests*, **11**, 38.

Klein, T., Randin, C. & Körner, C. (2015) Water availability predicts forest canopy height at the global scale. *Ecology Letters*, **18**, 1311–1320.

Koch, G.W., Sillett, S.C., Jennings, G.M. & Davis, S.D. (2004) The limits to tree height. *Nature*, **428**, 851–854.

Körner, C. (2012) *Alpine Treelines: Functional Ecology of the Global High Elevation Tree Limits*, Springer Basel.

Körner, C. & Paulsen, J. (2004) A world-wide study of high altitude treeline temperatures. *Journal of Biogeography*, **31**, 713–732.

Körner, C. & Spehn, E. (2019) A Humboldtian view of mountains. *Science*, **365**, 1061–1061.

Kullman, L. (2002) Rapid recent range-margin rise of tree and shrub species in the Swedish Scandes. *Journal of Ecology*, **90**, 68–77.

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Espirito-Santo, F.D.B., Hunter, M.O. & Oliveira, R. de (2005) Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, **32**.

Mao, L., Bater, C.W., Stadt, J.J., White, B., Tompalski, P., Coops, N.C. & Nielsen, S.E. (2019) Environmental landscape determinants of maximum forest canopy height of boreal forests. *Journal of Plant Ecology*, **12**, 96–102.

Moles, A.T., Warton, D.I., Warman, L., Swenson, N.G., Laffan, S.W., Zanne, A.E., Pitman, A., Hemmings, F.A. & Leishman, M.R. (2009) Global patterns in plant height. *Journal of Ecology*, **97**, 923–932.

Muggeo, V.M.R. (2020) *segmented: Regression Models with Break-Points / Change-Points Estimation*,.

Ninot, J.M., Batllori, E., Carrillo, E., Carreras, J., Ferré, A. & Gutiérrez, E. (2008) Timberline structure and limited tree recruitment in the Catalan Pyrenees. *Plant Ecology & Diversity*, **1**, 47–57.

Paulsen, J. & Körner, C. (2014) A climate-based model to predict potential treeline position around the globe. *Alpine Botany*, **124**, 1–12.

Payette, S., Filion, L., Delwaide, A. & Begin, C. (1989) Reconstruction of tree-line vegetation response to long-term climate change. *Nature*, **341**, 429–431.

Price, M.F., Gratzer, G., Duguma, L.A., Kohler, T., Maselli, D. & Romeo, R. (2011) *Mountain Forests in a Changing World - Realizing Values, addressing challenges.*, Published by FAO/MPS and SDC, Rome.

R Core Team (2018) *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Austria.

Roussel, J.-R., Auty, D., De Boissieu, F., Sánchez Meador, A. & Bourdon, J.-F. (2020) *lidR: Airborne LiDAR Data Manipulation and Visualization for Forestry Applications*,.

Rumpf, S.B., Hülber, K., Klonner, G., Moser, D., Schütz, M., Wessely, J., Willner, W., Zimmermann, N.E. & Dullinger, S. (2018) Range dynamics of mountain plants decrease with elevation. *Proceedings of the National Academy of Sciences*, **115**, 1848–1853.

Simard, M., Pinto, N., Fisher, J.B. & Baccini, A. (2011) Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research: Biogeosciences*, **116**.

Socha, J., Hawryło, P., Stereńczak, K., Miścicki, S., Tymińska-Czabańska, L., Młocek, W. & Gruba, P. (2020) Assessing the sensitivity of site index models developed using bi-temporal airborne laser scanning data to different top height estimates and grid cell sizes. *International Journal of Applied Earth Observation and Geoinformation*, **91**, 102129.

Tao, S., Guo, Q., Li, C., Wang, Z. & Fang, J. (2016) Global patterns and determinants of forest canopy height. *Ecology*, **97**, 3265–3270.

Thomas, R.Q., Hurtt, G.C., Dubayah, R. & Schilz, M.H. (2008) Using lidar data and a height-structured ecosystem model to estimate forest carbon stocks and fluxes over mountainous terrain. *Canadian Journal of Remote Sensing*, **34**, S351–S363.

Valbuena, R., O'Connor, B., Zellweger, F., Simonson, W., Vihervaara, P., Maltamo, M., Silva, C.A., Almeida, D.R.A., Danks, F., Morsdorf, F., Chirici, G., Lucas, R., Coomes, D.A. & Coops, N.C. (2020) Standardizing Ecosystem Morphological Traits from 3D Information Sources. *Trends in Ecology & Evolution*.

Wang, Y., Lehtomäki, M., Liang, X., Pyörälä, J., Kukko, A., Jaakkola, A., Liu, J., Feng, Z., Chen, R. & Hyyppä, J. (2019) Is field-measured tree height as reliable as believed – A comparison study of tree height estimates from field measurement, airborne laser scanning and terrestrial laser scanning in a boreal forest. *ISPRS Journal of Photogrammetry and Remote Sensing*, **147**, 132–145.

Wang, Y., Li, G., Ding, J., Guo, Z., Tang, S., Wang, C., Huang, Q., Liu, R. & Chen, J.M. (2016) A combined GLAS and MODIS estimation of the global distribution of mean forest canopy height. *Remote Sensing of Environment*, **174**, 24–43.

White, J.W., Rassweiler, A., Samhouri, J.F., Stier, A.C. & White, C. (2014) Ecologists should not use statistical significance tests to interpret simulation model results. *Oikos*, **123**, 385–388.

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W. & Gobakken, T. (2012) Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment*, **121**, 196–209.

Zhang, J., Nielsen, S.E., Mao, L., Chen, S. & Svenning, J.-C. (2016) Regional and historical factors supplement current climate in shaping global forest canopy height. *Journal of Ecology*, **104**, 469–478.

TABLES

Table 1. Summary of the results for the fitted models of maximum canopy height as a function of elevation.

	Breakpoint (m)		β_1		β_2		R^2		ΔR^2	
	Mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
General model	1623.3	4.7	$7.9\cdot10^{-06}$	$7.4\cdot10^{-06}$	$-7.8\cdot10^{-4}$	$6.5\cdot10^{-06}$	0.63	0.004	0.18	0.003
Per aspect classes										
North-facing	1657.1	9.1	$4.2\cdot10^{-5}$	$1.4\cdot10^{-5}$	$-1.0\cdot10^{-3}$	$1.9\cdot10^{-5}$	0.66	0.008	0.21	0.007
South-facing	1674.0	92.9	$-5.2\cdot10^{-5}$	$7.9\cdot10^{-5}$	$-6.4\cdot10^{-4}$	$7.5\cdot10^{-5}$	0.40	0.011	0.07	0.005
Per species										
Pinus uncinata	1782.9	9.1	$-1.3\cdot10^{-4}$	$2.0\cdot10^{-5}$	$-7.8\cdot10^{-4}$	$1.2\cdot10^{-5}$	0.59	0.008	0.14	0.006
Abies alba	1722.3	20.5	$-1.8\cdot10^{-4}$	$2.6\cdot10^{-5}$	$-1.4\cdot10^{-3}$	$9.9\cdot10^{-5}$	0.87	0.012	0.25	0.023
Pinus sylvetris	1915.2	32.4	$-1.8\cdot10^{-4}$	$5.9\cdot10^{-6}$	$-1.2\cdot10^{-3}$	$1.9\cdot10^{-4}$	0.17	0.007	0.01	0.002
Fagus sylvatica	1696.9	135.35	$-1.5\cdot10^{-4}$	$4.3\cdot10^{-5}$	$-1.1\cdot10^{-3}$	$5.8\cdot10^{-4}$	0.24	0.033	0.04	0.015

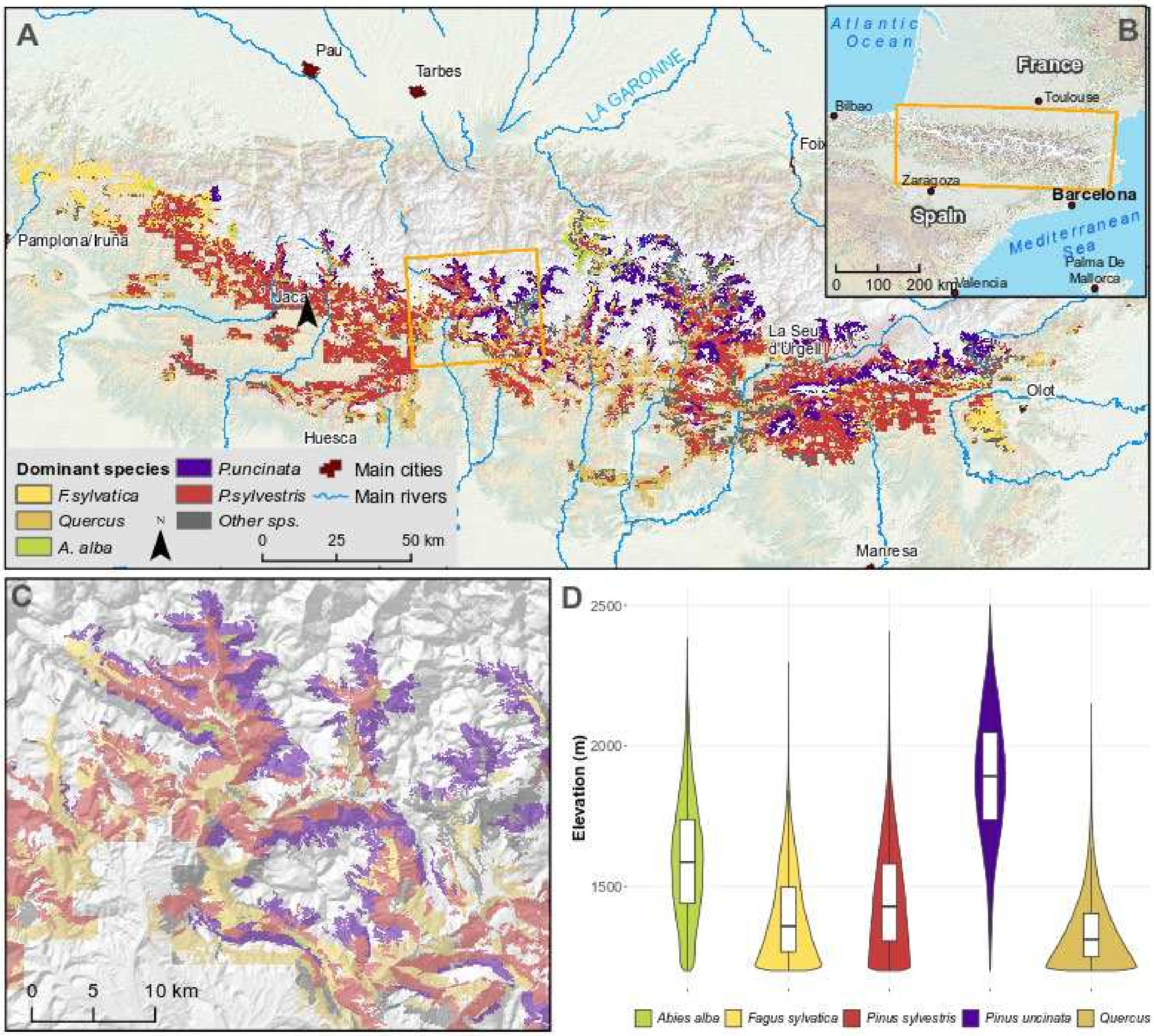
The parameter estimates correspond to a segmented log-linear model in the form: $log(max\_height) = \alpha_1 + \beta 1 \cdot Elevation$ *for elevation < breakpoint; and* $log(max\_height) = \alpha_2 + \beta 2 \cdot Elevation$ *for elevation > breakpoint*. The results are presented for the general model, for a model fitted for each species separately, and for a model fitted for each aspect class separately. Values are average predictions of parameters estimates for 1,000 models fitted to random subsets of the dataset (5,000 points for training and 5,000 for validation). R^2 for each model is calculated as the coefficient of determination of the relationship between the observed data and the predicted data using the validation dataset. ΔR^2 refers to the average increase in R^2 of the segmented model as compared to a log-linear model.

Table 2. Mean and sd of r-squared and RMSE of the 1000 tested models for each realization. Model name refers to the variable included as predictor of maximum canopy height, whereas Full Model refers to a multivariate model including all the possible predictors

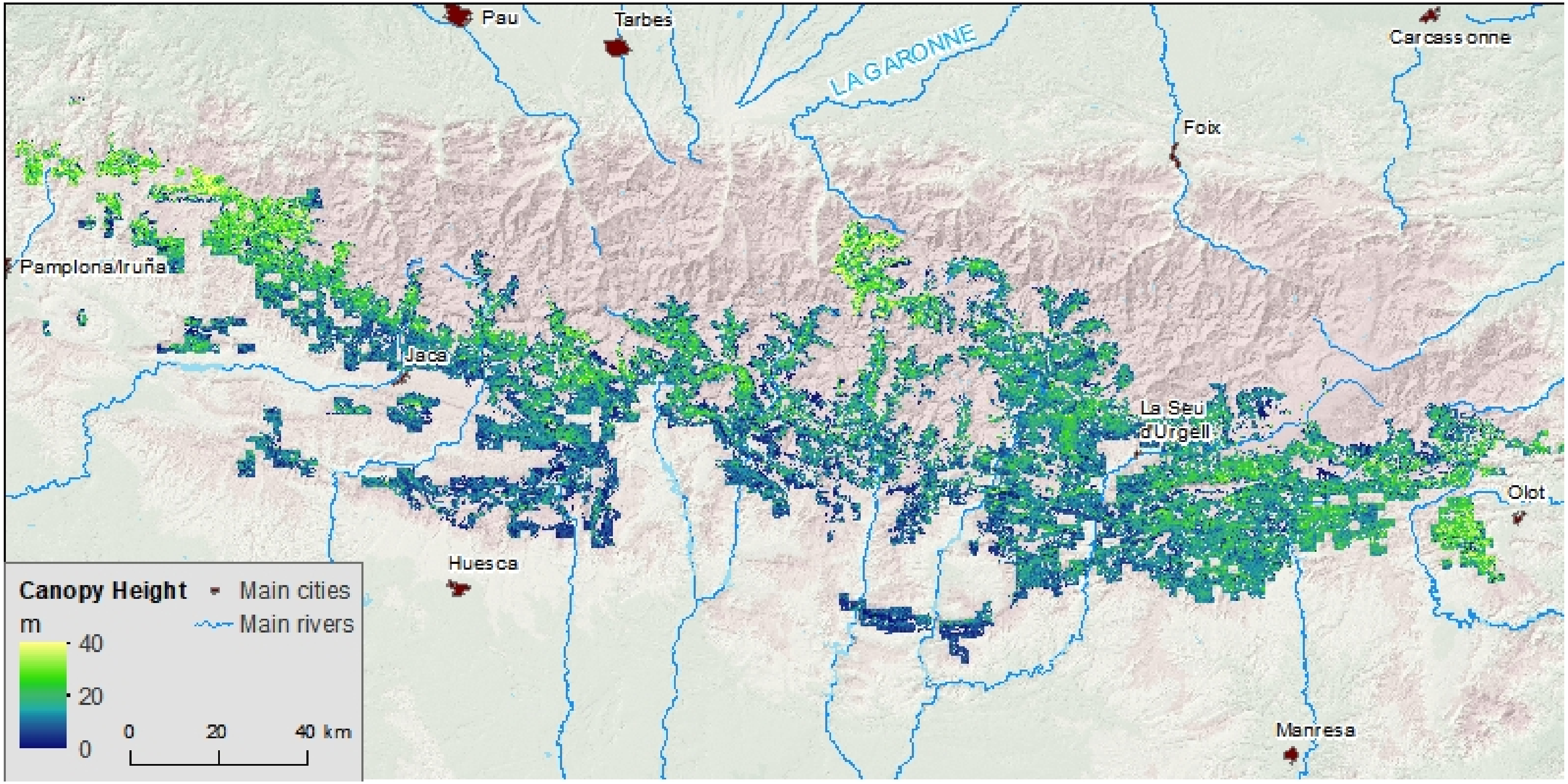
Model	Mean R^2	SD R^2	Mean RMSE	SD RMSE
Elevation	0.63	0.00613	2.25	0.0220
Mean anual temperature	0.358	0.00873	2.94	0.0254
Annual rainfall	0.182	0.00817	3.33	0.0256
Soil depth	0.139	0.00718	3.41	0.0243
Northness	0.030	0.00447	3.62	0.0247
Distance to sea	0.027	0.00379	3.63	0.0258
Sand %	0.023	0.00435	3.63	0.0268
Clay %	0.016	0.00416	3.65	0.0260
Silt %	$-1.2 \cdot 10^{-4}$	0.00305	3.68	0.0246
Slope	-0.0012	0.00277	3.68	0.0257
Eastness	$-7.7 \cdot 10^{-3}$	0.00248	3.69	0.0248
Full model	**0.653**	**0.00570**	**2.17**	**0.0208**

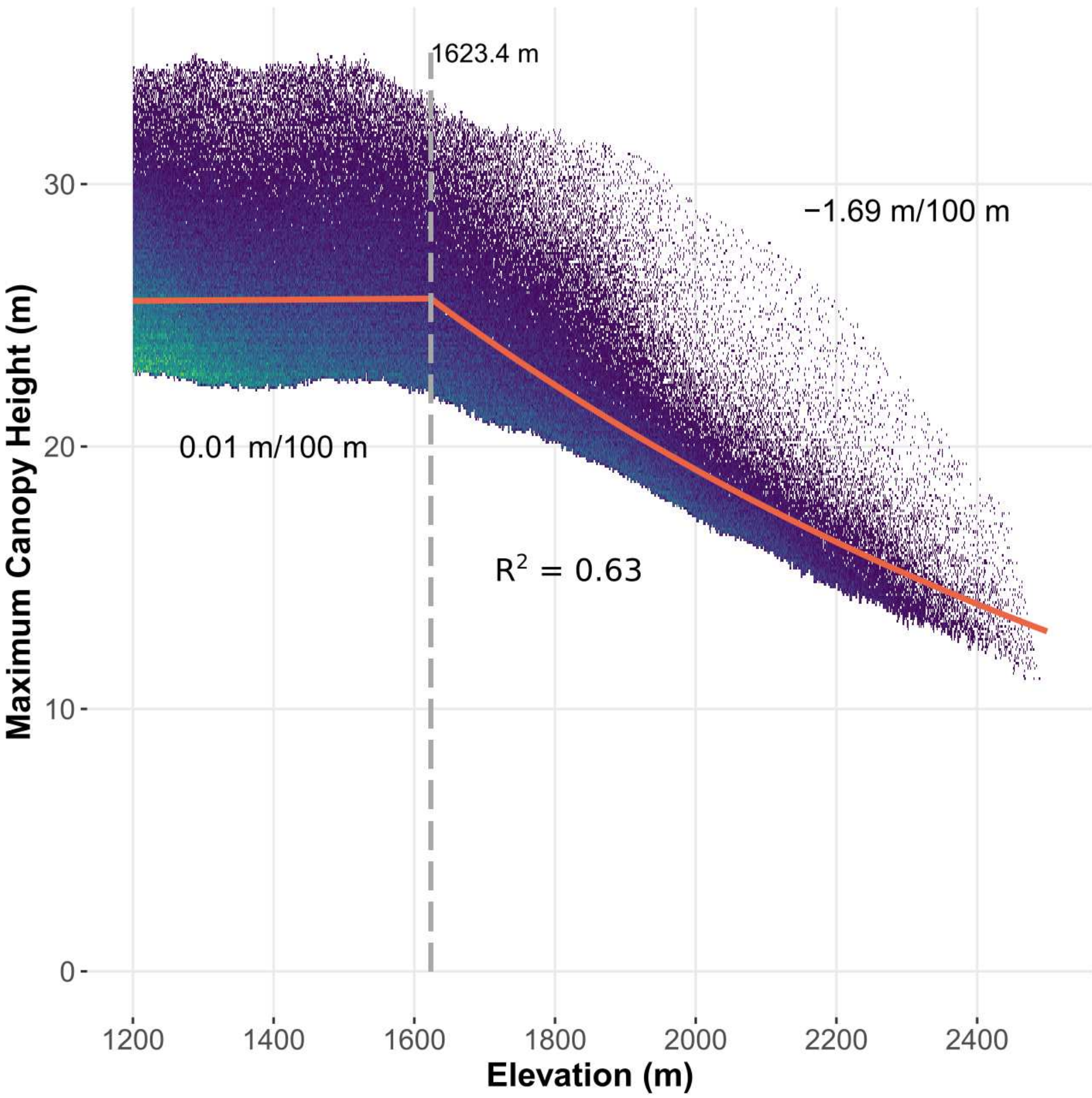
R^2 for each model is calculated as the coefficient of determination of the relationship between the observed and predicted data, using randomly chosen independent datasets for training (5,000 points) and validation (5,000).

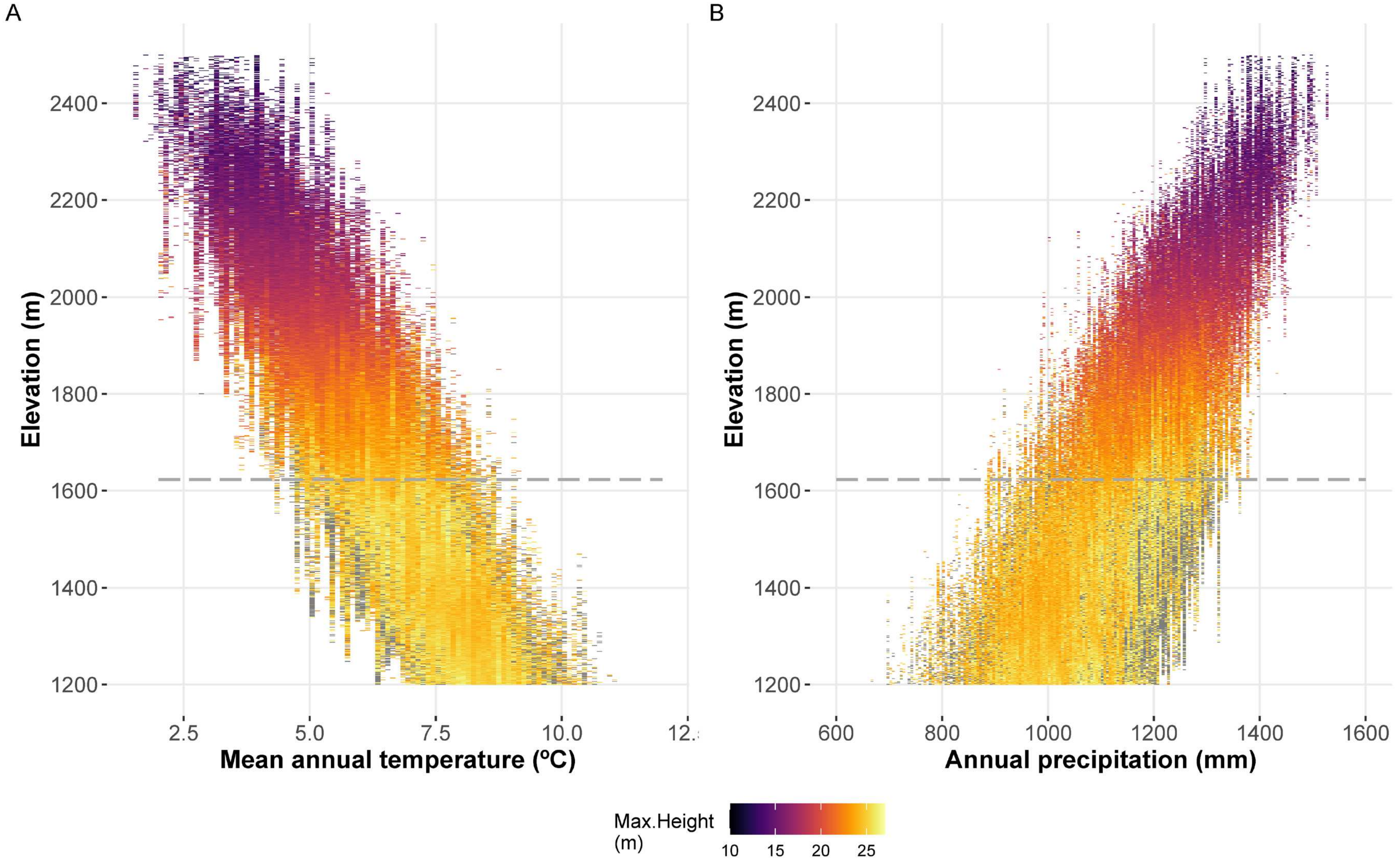
FIGURES



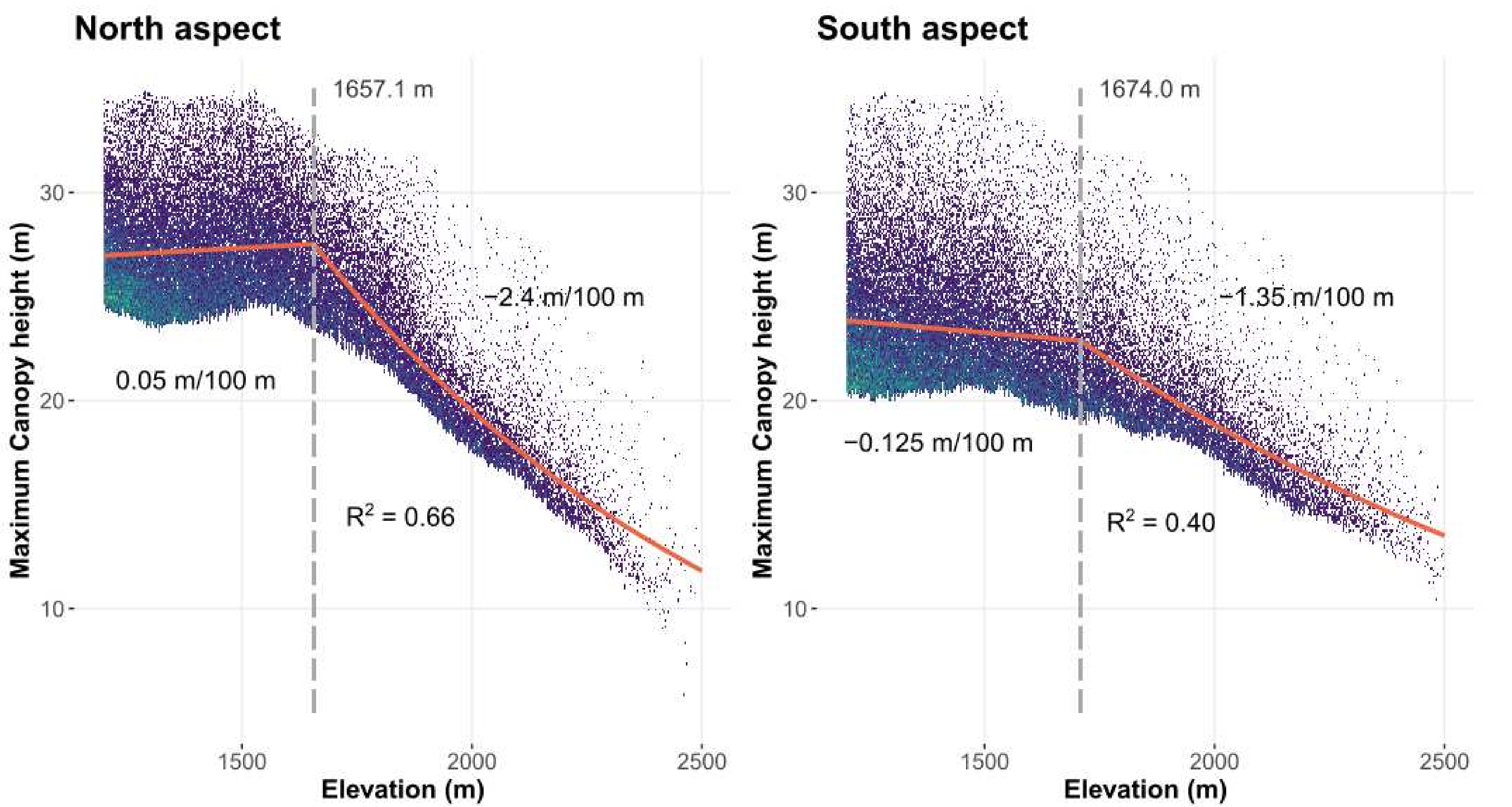
A
B
C
D
Pau
Tarbes
LA GARONNE
Foix
Pamplona/Iruña
Jaca
La Seu d'Urgell
Olot
Huesca
Manresa
Atlantic Ocean
France
Bilbao
Toulouse
Zaragoza
Barcelona
Spain
Mediterranean Sea
Palma De Mallorca
Valencia
0 100 200 km
Dominant species
F.sylvatica
Quercus
A. alba
P.uncinata
P.sylvestris
Other sps.
Main cities
Main rivers
N
0 25 50 km
0 5 10 km
Elevation (m)
2500
2000
1500
Abies alba
Fagus sylvatica
Pinus sylvestris
Pinus uncinata
Quercus


Pau
Tarbes
LA GARONNE
Carcassonne
Foix
Pamplona/Iruña
Jaca
La Seu
d'Urgell
Olot
Huesca
Manresa
Canopy Height
m
40
20
0
Main cities
Main rivers
0
20
40 km

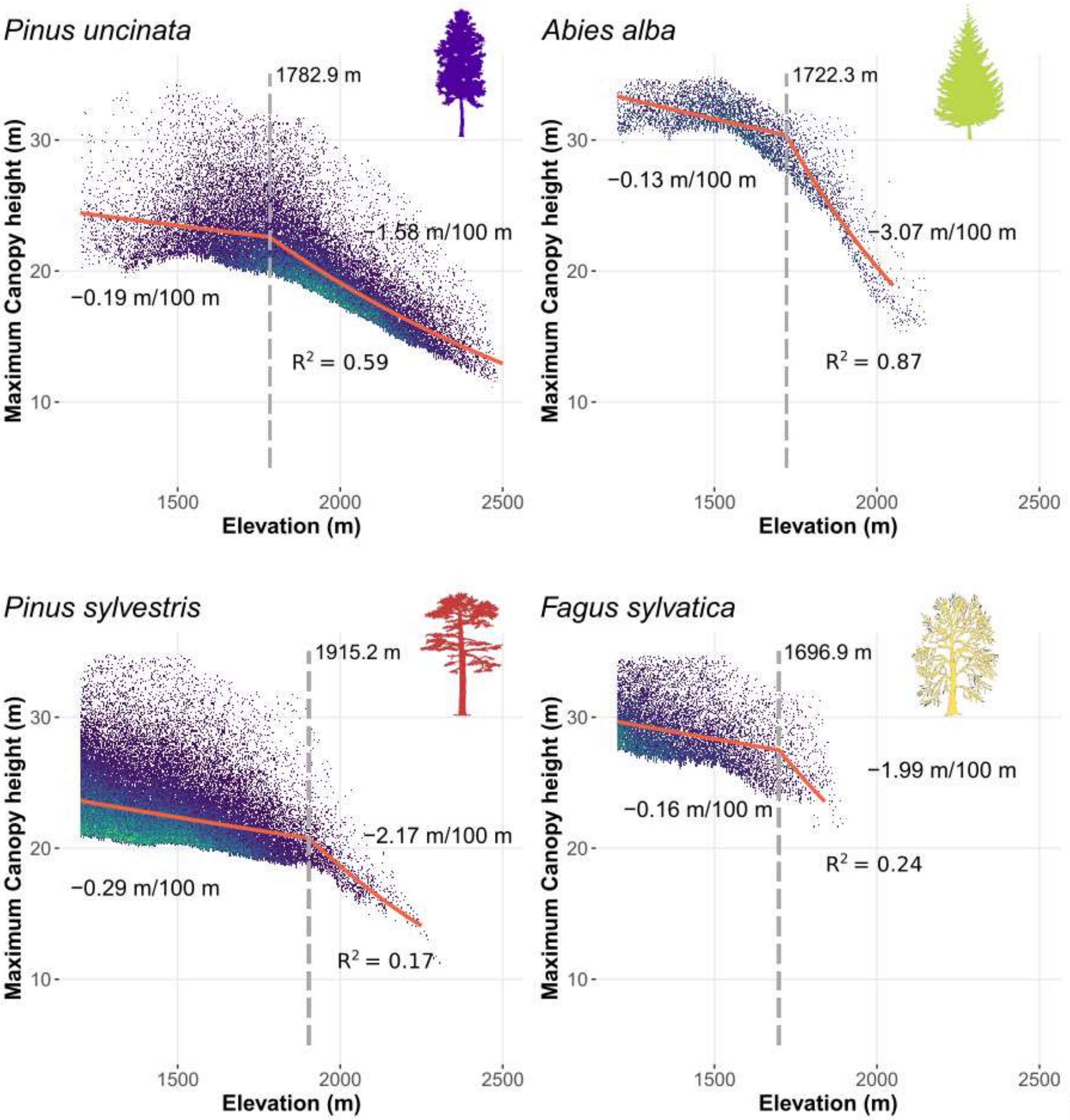

1623.4 m
−1.69 m/100 m
0.01 m/100 m
R^2 = 0.63
Maximum Canopy Height (m)
30
20
10
0
1200
1400
1600
1800
2000
2200
2400
Elevation (m)


A
B
2400
2200
2000
1800
1600
1400
1200
Elevation (m)
2.5
5.0
7.5
10.0
Mean annual temperature (ºC)
600
800
1000
1200
1400
1600
Annual precipitation (mm)
Max.Height (m)
10
15
20
25

567


North aspect
1657.1 m
0.05 m/100 m
−2.4 m/100 m
R² = 0.66
South aspect
1674.0 m
−0.125 m/100 m
−1.35 m/100 m
R² = 0.40
Maximum Canopy height (m)
Elevation (m)
10
20
30
1500
2000
2500

568



Pinus uncinata
1782.9 m
−0.19 m/100 m
−1.58 m/100 m
R² = 0.59
Abies alba
1722.3 m
−0.13 m/100 m
−3.07 m/100 m
R² = 0.87
Pinus sylvestris
1915.2 m
−0.29 m/100 m
−2.17 m/100 m
R² = 0.17
Fagus sylvatica
1696.9 m
−0.16 m/100 m
−1.99 m/100 m
R² = 0.24
Maximum Canopy height (m)
Elevation (m)

569

FIGURE LEGENDS

Figure 1. Location of the study area and distribution of the main species along elevation gradients (A) Distribution of the main forest species across the Spanish Pyrenees; (B) Location of the study area within Southern Europe; (C) Detail of the distribution of the main species along elevation gradients in a valley in the Central Pyrenees; (D) Violin plots showing the overall distribution of the main species across the elevation gradient in the Pyrenees, as observed using PNOA LiDAR data and the Spanish Forest Map.

Figure 2. High-resolution (20 m) canopy height grid of the Spanish Pyrenees as derived from the Spanish Airborn LiDAR plan (PNOA). Canopy height was higher at both ends of the Pyrenees, where the sea influence softens the climate and allows the presence of tree species such as fir or beech.

Figure 3. Relationship between terrain elevation and maximum canopy height across the Spanish Pyrenees, as determined from airborne LiDAR data. Orange lines represent the predictions according to a segmented log-linear regression model, and dashed line represents the breakpoint identified by the same model. Values indicate the approximate rate of change in maximum canopy height for a 100 m change in elevation below and above the breakpoint. The segmented log-linear model is the average prediction of 1,000 models fitted to random subsets of the original dataset. R^2 is calculated as the coefficient of determination of the relationship between the observed data and the predicted data using the validation dataset.

Figure 4. Variation of maximum canopy height with elevation and climatic variables. Maximum canopy height increases with increasing temperature (A) and decreasing precipitation (B) but this relationship is explained by the covariation between elevation and climate variables (see Table 2). Elevation breakpoint is indicated by the dashed gray line.

Figure 5. Relationship between terrain elevation and maximum canopy height in the Spanish Pyrenees, split for north-facing and south-facing slopes. Orange lines represent the predictions according to a segmented log-linear regression model, and dashed line represents the breakpoint identified by the same model. Values indicate the

approximate rate of change in maximum canopy height for a 100 m change in elevation below and above the breakpoint. The segmented log-linear model is the average prediction of 1,000 models fitted to random subsets of the original dataset. R2 for each model is calculated as the coefficient of determination of the relationship between the observed data and the predicted data using the validation dataset.

Figure 6. Relationship between terrain elevation and maximum canopy height in the Spanish Pyrenees, split across the main dominant species. Orange lines represent the predictions according to a segmented log-linear regression model, and dashed line represents the breakpoint identified by the same model. Values indicate the approximate rate of change in maximum canopy height for a 100 m change in elevation below and above the breakpoint. The segmented log-linear model is the average prediction of 1,000 models fitted to random subsets of the original dataset. R2 for each model is calculated as the coefficient of determination of the relationship between the observed data and the predicted data using the validation dataset.