FORGENIUS

Improving access to FORest GENetic resources Information and services for end-Users

Deliverable D3.5

Fitness (fertility estimated by drones, realised fecundity estimated via seedlings modelling) & selection gradients (relationship fitness and trait proxy) for at least one species

Planned delivery date (as in DoA):M35 30/11/2023

Actual submission date: 24/11/2023, month M35

Workpackage: WP3

Workpackage leader: INIA

Deliverable leader: UMR

Version: 1.0

Project funded by the European Commission within the Horizon 2020 Programme	
Dissemination Level	
PU Public	**PU**
CI Classified, as referred to Commission Decision 2001/844/EC	
CO Confidential, only for members of the consortium (including the Commission Services)	

Research and Innovation action: GA no. 862221

Start date of the project: January 1st, 2021

TABLE OF CONTENTS

1 Summary

In this report for the FORGENIUS project, we provide a summary of the findings from Deliverable 3.5, focusing on tree-level environmental, phenotypic and genomic data for maritime pine (*Pinus pinaster*) genetic conservation units (GCUs). Our approach involved the use of unmanned aerial vehicles (UAVs) to gather comprehensive environmental data from two distinct locations, Tocchi (Italy) and Lacanau (France), as well as individual tree-level cone production data. Furthermore, we provide model-derived realized fecundity estimates (i.e. number of offspring established at the time of sampling) using MEMMseedlings, CERVUS and NMπ software, based on the genotypes and spatial positions of both adult trees and juveniles. Outputs were consistent across methods, with few trees contributed to successful regeneration. The model-based approach (transgenerational measures) is complementary to that using UAV (one year measures).

Selection gradients estimating the relationship between traits (estimated in WP2 and WP3, see D3.1, D3.2 and D3.3) and fitness-proxy (i.e. fertility and realised fecundity estimates), were corrected for spatial autocorrelation and environmental variability (both detected by UAVs), two factors that may bias the output. Results detected significant selection gradients with some traits harbouring a potential to evolve. Overall, our multifaceted approach is providing valuable insights into the resilience and adaptability of maritime pine in the face of environmental changes.

2 Introduction

The far-reaching objective of WP3 is to characterize genetic conservation units (GCUs) with dynamic descriptors that can inform our ability to forecast the adaptability of forest genetic resource (FGR) collections, as well as their resilience to future environmental threats. Building such indices requires the estimation of components of fitness in field conditions (i.e., seed production), the estimation of heritability of adaptive traits (see D3.3), and the estimation of the relationship of adaptive traits to fitness (i.e., selection gradients) in contrasted environments. To estimate fitness in the field we relied on two complementary approaches: 1) fertility estimated by unmanned aerial vehicle (UAVs). This approach provides access to seed counting for a given fruiting season across the entire GCU; 2) realised fecundity estimated via full mating models. This approach uses genomic resources developed in WP4 to model fecundity based on a parentage analysis for adults and juveniles. It further provides estimates of realised fecundity of adult trees across several generations of juveniles. Adaptive traits were measured both directly in the field (e.g., height, DBH) and indirectly via a NIRS approach (see D3.3). The relationship between fitness and adaptive traits was estimated with quantitative genetic models (i.e., selection gradients). In this Deliverable we present outputs for *Pinus pinaster*, one of the four species investigated in WP3 for which we have the most advanced results. The same methodology will be applied for the remaining species. Two ecologically contrasting GCUs were analysed per species, each with 500 georeferenced adult trees and 250 juveniles. Comparing results between GCUs will provide insights on real-time adaptive responses induced by ongoing environmental changes.

3 Results

3.1 Fertility and environmental variables estimated by drones

Data Collection and Processing

Aerial RGB images were obtained through UAV flights at two locations: Tocchi (Tuscany) and Lacanau (French Atlantic coast). To ensure the geospatial accuracy of the data, seven ground control point (GCP) surveys were conducted, which were used for photogrammetric georeferencing. With these high-precision data, a 3D point cloud was generated to derive environmental variables and morphological features, as well as orthoimages with a 5 mm resolution for cone classification. Afterwards, we present briefly the methods used for each variable.

- Pine-cone fertility

To register cone productivity, we rely on a machine learning (ML).

- Morphological traits based in tree crown (ITC)

The delineation of individual tree crowns (ITC) was estimated from dense 3D point clouds. ITC is crucial for recording cone productivity per tree. The ITCs not only record the position of the tree crown but also generate a list of morphological traits based on crown measurements projected onto a 2D canopy.

- Tree Competition Index based on Neighborhood

In the process of calculating indices that describe interactions within forest neighborhoods, the initial essential step involves identifying groups of trees comprising each neighborhood. Subsequently, the calculated neighborhood indices are presented, offering a detailed insight into these interactions.

Tree Raw Densities. For the assessment of tree densities, both overall and specific to tree species, measurements are provided in units of square meters per hectare (m^2/ha^{-1}). These measurements are based on the area that each tree occupies.

Tree Densities Using Angular Occupancy. To quantify tree densities in a different manner, the angular approach is employed. This approach calculates the sum of angles occupied by trees within a given neighborhood.

Tree Densities Using Angular Larger Occupancy. Furthermore, an option is available to refine density computations by considering only neighboring trees that surpass the size of the focal tree, anticipating weaker competition from smaller neighbors within the study system. By including larger neighbors exclusively, a more tailored assessment of competitive interactions can be attained.

- Hydrological variables (SWC and STWI)

To estimate hydrological variables dependent on topography, we employed a kriging method along with topographical data and soil water content (SWC) *in situ* measurements. To compute the Topographic Wetness Index (STWI), we first generated a Digital Terrain Model (DTM) from the dense point cloud captured by a UAV. Next, we calculated the terrain slope and adjusted slope values to prevent mathematical errors. Slope values were converted to radians, and we determined the Contributing Area (AC), representing the region contributing to water flow at each cell of 10 cm. Finally, we applied the SAGA toolkit to obtain the STWI.

To determine the relationships between the two measurement methods, Pearson and Spearman correlations were calculated, aiming to identify linear and monotonic relationships, respectively. Additionally, linear regression analyses were conducted to ascertain the extent to which STWI values could predict SWC measurements. Subsequently, visual analyses were employed to estimate the level of agreement between the two methods, highlighting any systematic bias and the consistency, or lack thereof, of the observed differences.

Results

Pine-cone productivity

We visually counted a total of 1148 young cones of *Pinus Pinaster* in Lacanau and 1147 in Tocchi from 5 mm resolution orthoimages (Figure 1). This visual-sampling-validated cone classification showed a comparable number of young cones at both sites. Tocchi had nearly three times more mature cones than Lacanau, and Tocchi had three times more mature cones than young cones, while Lacanau had twice as many young cones as mature cones.

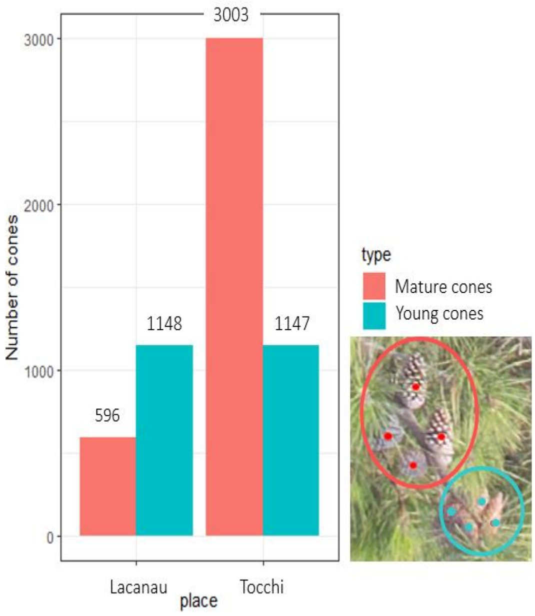


Figure 1 – Visual estimation of young and mature cones on orthoimage of 5 mm resolution for the Lacanau and Tocchi GCUs.

The Weka FastRandomForest classifier's sequential training results for automated young cone classification demonstrated minimal bias. This indicates that images taken at 15 meters above the canopy surface had sufficient resolution for identifying young cones. However, variations in flight altitude, canopy heterogeneity, overlapping leaves, shadows, and similar leaf-cone textures affected automatic recognition at specific GCU sites.

Morphological traits based in tree crown (ITC)

The analysis of PCA (Figure 2) for the Tocchi and Lacanau plots reveals consistent patterns in the relationships between key morphological traits. As seen in the plots, robust positive correlations between zmax and zmean underscore a direct relationship between the maximum height and the mean height of the trees. The selected variables—zmax, zmean, zkurt, zskew, pzabovezmean, n, area, and zsd—provide a non-redundant and accurate representation of tree morphology.

PCA reinforces the importance of these variables by revealing that they are significant factors in the morphological variability of trees. The first principal components in Tocchi and Lacanau explain a substantial portion of the variance, with 'area' standing out as a distinctive factor. Although the standard deviation of heights (zsd) is less prominent in Tocchi, it notably contributes to the observed variability in Lacanau.

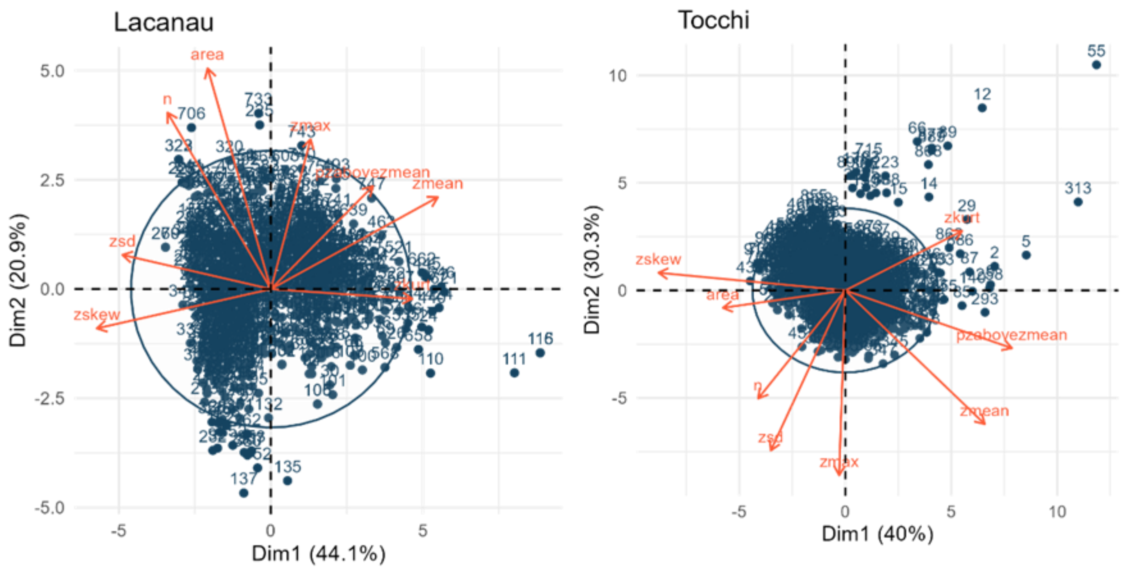


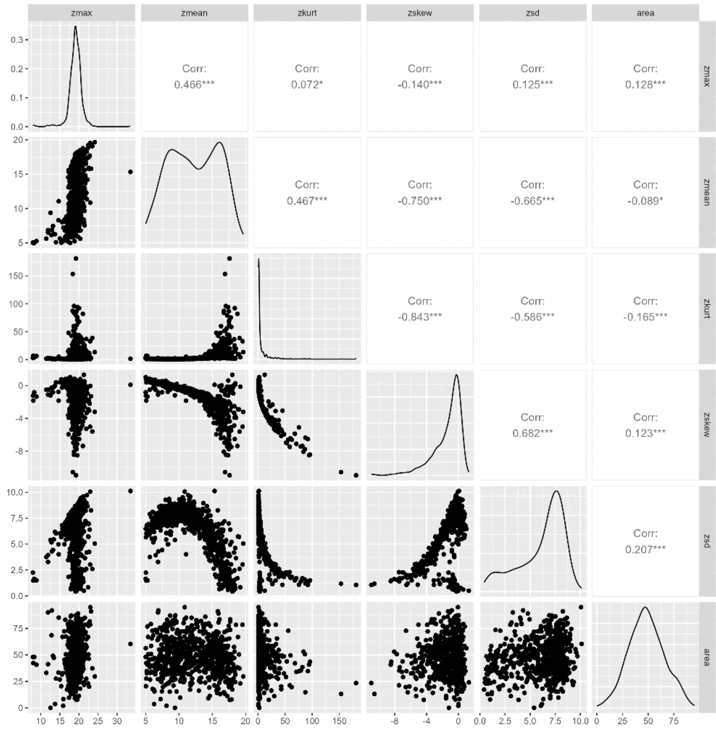
Figure 2 – Principal Components Analysis between the selected morphological traits for both GCU.

The comparison of the Tocchi and Lacanau plots (Figure 3) highlights patterns and differences that may indicate environmental or genetic influence on tree morphological traits. In Tocchi, the strong positive correlation between maximum height (zmax) and mean height (zmean) of trees underscores a direct relationship between these dimensions, while the negative correlation between kurtosis (zkurt) and skewness (zskew) suggests that trees with sharper heights are less asymmetric. The moderate correlation between the standard deviation of heights (zsd) and area implies a connection between height variability and the extent of the tree crown, although maximum height alone does not reliably predict crown area.

In Lacanau, the dynamics are similar but with distinct nuances: the correlation between zmax and zmean, although positive, is slightly lower, and the negative relationship between zkurt and zskew is more pronounced, reflecting less asymmetry in sharper height distributions compared to Tocchi. Furthermore, the positive association between zsd and area is stronger, indicating a more defined relationship between height variability and crown area. Similar to Tocchi, maximum height does not emerge as a strong predictor of crown area.

The careful selection of variables such as zmax, zmean, zkurt, zskew, zsd, and area for this study is crucial, as each provides valuable and complementary information about tree morphology. By integrating these morphological traits with fertility data at the individual tree level, one can begin to draw connections between the physical shape of trees and their reproductive success. Ultimately, this multidimensional approach clarifies the interactions between morphology and underlying genetics.

Lacanau



Tocchi

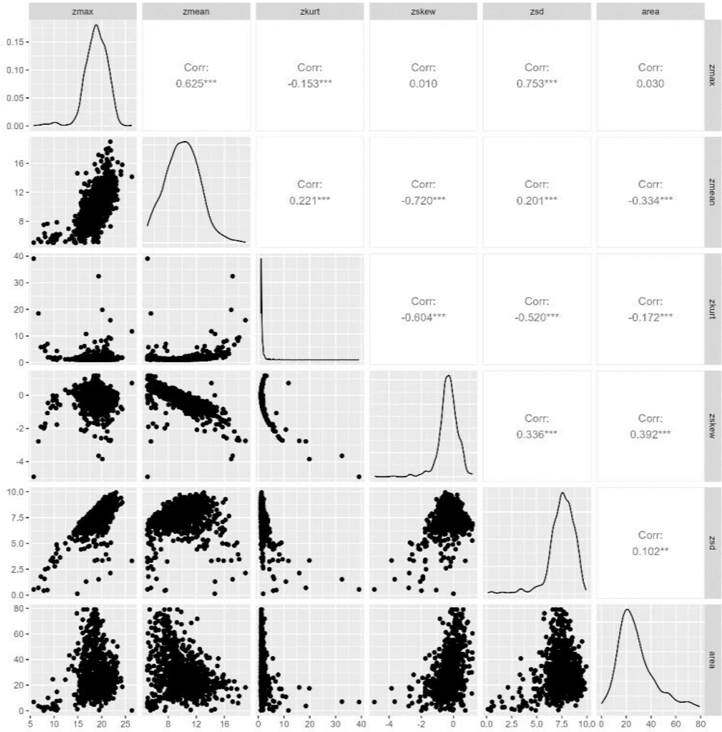


Figure 3 – Relationships between the selected morphological traits fot both GCUs.

Tree Competition Index based on Neighborhood

The structure of *Pinus Pinaster* forests in Tocchi and Lacanau, analyzed through tree raw density and angular occupancy, reveals distinctive patterns of tree competition and interaction at each site. With an analysis radius of 30 meters, suitable for capturing tree competition dynamics, notable differences in terms of tree distribution and density are revealed.

In Tocchi, tree raw density varies widely from 4.52 to 40.01 m²/ha, indicating pronounced spatial heterogeneity. This diversity is visually reflected on the density map (Figure 4) with a range of colors denoting areas of different tree occupancy. Angular occupancy values in Tocchi are lower, suggesting more scattered competition and greater space availability for each tree, resulting in a more open forest structure with a variety of microhabitats.

In contrast, Lacanau exhibits high and consistent tree density, with less variability, with tree raw density varies widely from 14.78 to 42.14 m²/ha indicating a more uniform forest structure. The map of this site shows a more homogeneous color pattern, indicating intense competition for resources such as light and nutrients. In Lacanau, both angular density and the density of larger trees are higher, highlighting a forest with regeneration and competition dynamics influenced by dominant trees.

In addition to raw density, angular occupancy reflects the proportion of space each tree occupies within its neighborhood. In Tocchi, angular density ranges from 0.19 to 2.25, while in Lacanau, it ranges from 0.81 to 3.33. These results indicate greater competition for space in Lacanau, where trees tend to occupy larger angles within their established neighborhood.

The angular occupancy of larger trees ("Angular Large") also differs between the two sites. In Tocchi, this occupancy is lower, with values ranging from 0.01 to 2.23, implying a lesser influence of large trees on the forest structure. In Lacanau, although the range is narrower (0.01 to 1.15), it suggests a significant presence of large trees that could exert a greater impact on resource competition.

In summary, the inclusion of angular density and angular occupancy of larger trees complements the understanding of forest structure and ecological interactions in both sites. Tocchi exhibits a more heterogeneous structure with spatially dispersed competition, while Lacanau displays a more uniform and dense forest, indicative of more direct competition among trees. These findings underscore the importance of spatially explicit assessment in understanding the ecology of mono-specific *Pinus pinaster* forests.

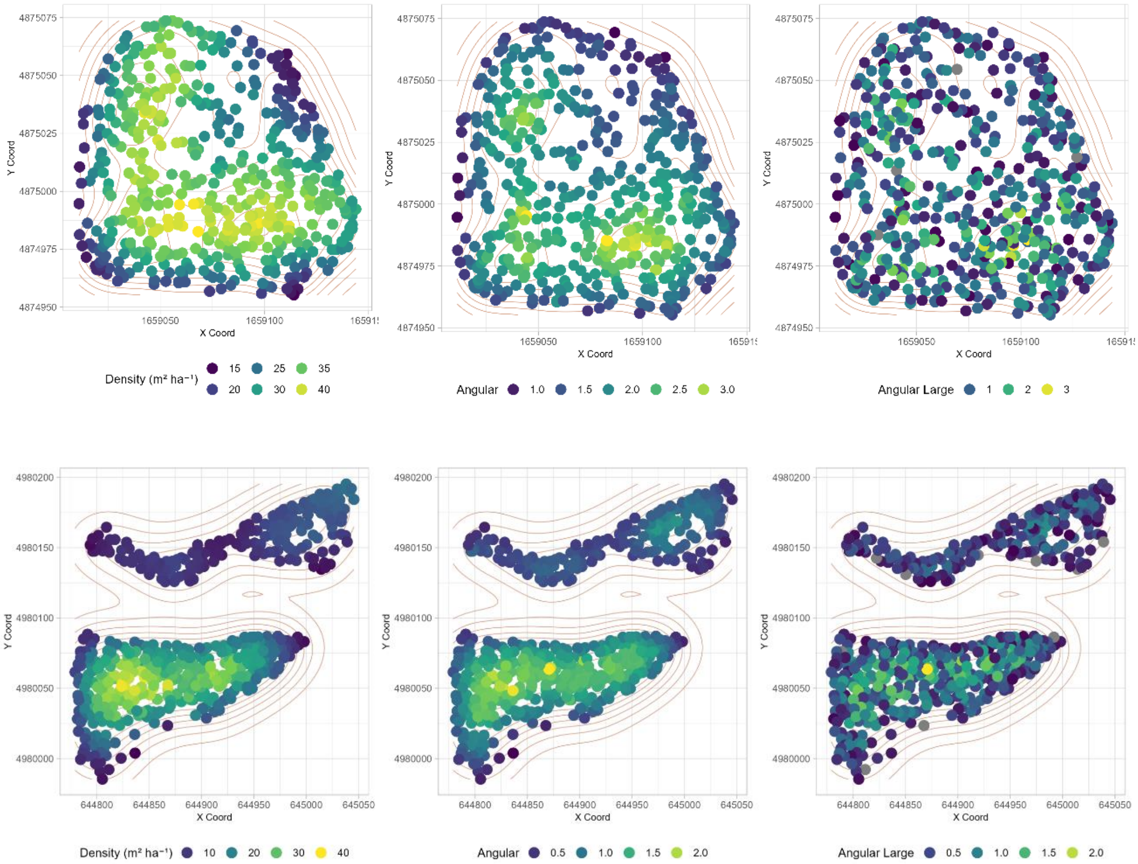


Figure 4 - Descriptive mapping of the statistics of the Density Raw, Angular and Angular Large indices for both GCU.

In the study of hydrological variables for the Tocchi and Lacanau plots, mapping (Figure 5) and statistical analyses were applied to explore the relationship between the Soil Topographic Wetness Index (STWI) and *in-situ* Soil Water Content (SWC). Calculations of Pearson and Spearman correlations revealed that, in both sites, the relationships between SWC and STWI measurements are weak.

These results were reflected in the conducted linear regression analyses, which showed a low R^2 value in both locations, indicating that STWI does not provide a reliable prediction of *in-situ* soil moisture for these areas. Interestingly, the regression model for Lacanau indicated a negative relationship between SWC and STWI, with a significantly negative slope. This suggests that contrary to the assumption that areas with higher topographic moisture would have higher soil moisture, the opposite occurs in Lacanau: an increase in STWI is associated with a decrease in SWC.

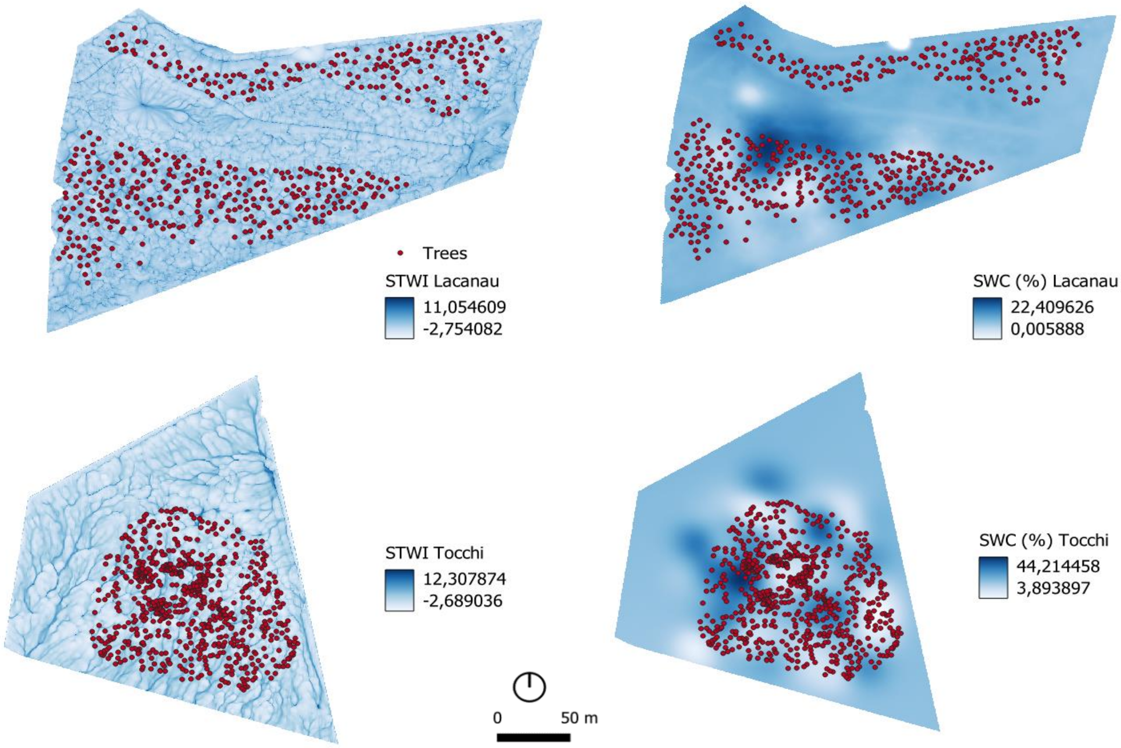


Figure 5. Mapping of Saga Topographic Wetness Index (STWI) and Soil Water Content (SWC) at 10 cm resolution for both GCUs.

Output underscore that STWI and SWC approach soil moisture from different perspectives. While STWI, based on topographic data, reflects a more general and long-term view, SWC provides a momentary snapshot influenced by current conditions such as recent precipitation. The inclusion of elevation data in SWC analysis provides valuable topographic context but does not alter its nature as a momentary measurement.

SWC remains a critical component for understanding soil moisture at a specific time and place and serves as an essential benchmark for relative moisture comparisons within the study area. By combining these measurements with temporal monitoring and STWI analyses, a more comprehensive picture of soil moisture dynamics could be achieved. This integrative approach is crucial for accurately interpreting how soil moisture conditions can influence ecological processes.

3.2 Realised fecundity estimated via seedlings modelling

We estimated the realized effective fecundity from the genotypes and spatial positions of both adults and juveniles with two different methods. We used first MEMMseedlings software (Oddou-Muratorio et al., 2018) which allows to infer jointly individual male and female effective fecundities with pollen and seed dispersal kernels, thus correcting for sampling issues.

In the Lacanau and Tocchi GCUs, the pollen dispersal kernel indicated that the majority of pollen comes from outside the stand (Lacanau: average pollen immigration rate = 94%; Tocchi: average pollen immigration rate = 78%), while more than half of seeds come from inside the stand (Lacanau: average seed immigration rate = 55%; Tocchi: average seed immigration rate = 16%).

Realized effective fecundity, estimated as the number of offspring established at the time of sampling using MEMMseedlings software, showed an asymmetrical distribution in each stand, with higher densities for lower values (Figure 6). This means that most trees have very few offspring while only a few have many, a typical result for continuous forest tree stands. Maximum effective female fecundity was twice as high at Lacanau (effective fecundity = 8.5) as at Tocchi (effective fecundity = 4.8). Overall, effective female fecundity was greater than one for 21 trees (4.2%) at Lacanau, and for 15 trees (3.1%) at Tocchi.

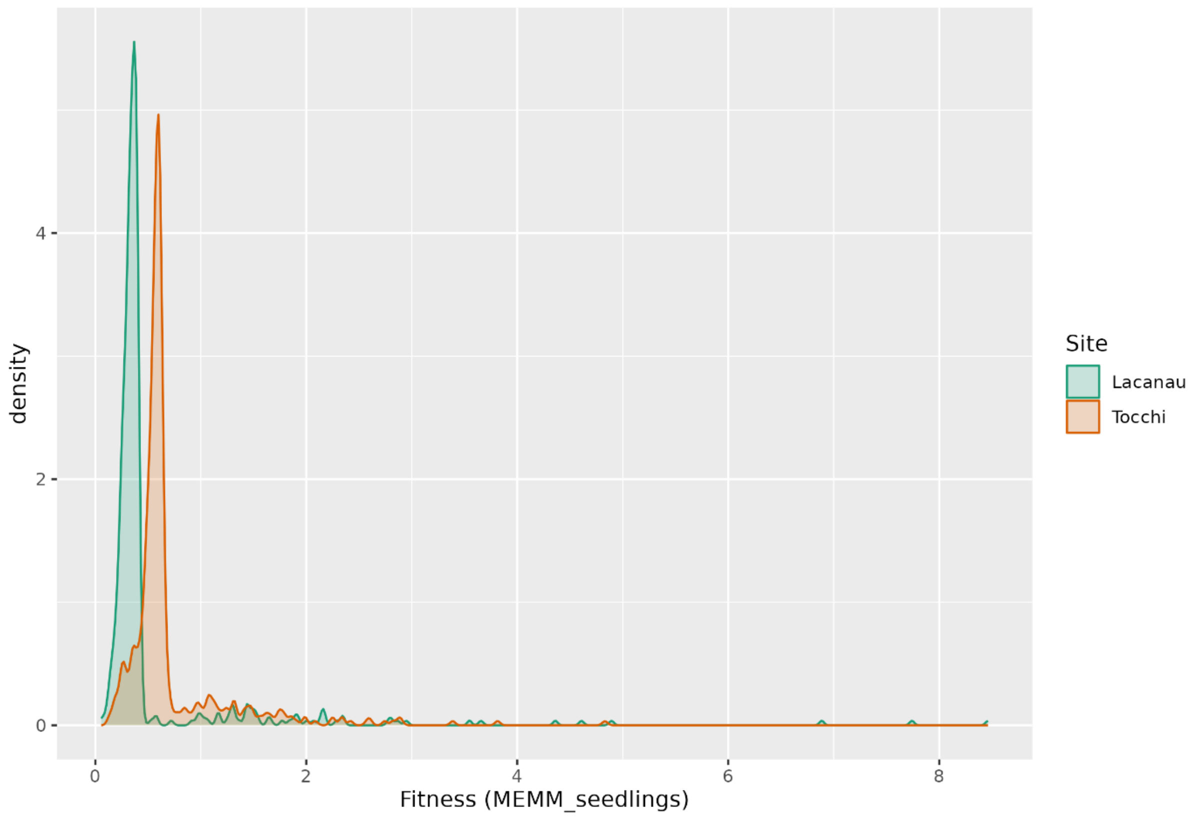


Figure 6: Density curves of the fitness proxies estimated by MEMM software in both GCUs.

In addition, we assessed the relative fitness of each tree based on fecundity estimates obtained from a categorical parentage analysis implemented in CERVUS 3.0.7 software (Kalinowski et al., 2007, Marshall et al., 1998). A similar asymmetrical fitness distribution as that obtained from MEMMseedlings software, was observed for categorical fecundity (estimated as the number of established offspring) calculated using CERVUS parentage analysis (Figure 7). Maximum effective female fecundity was higher at Tocchi (effective fecundity = 11) than at Lacanau (effective fecundity = 9). Effective female fecundity was greater than one for 49 trees (10.1%) at Tocchi, and for 29 trees (5.8%) at Lacanau.

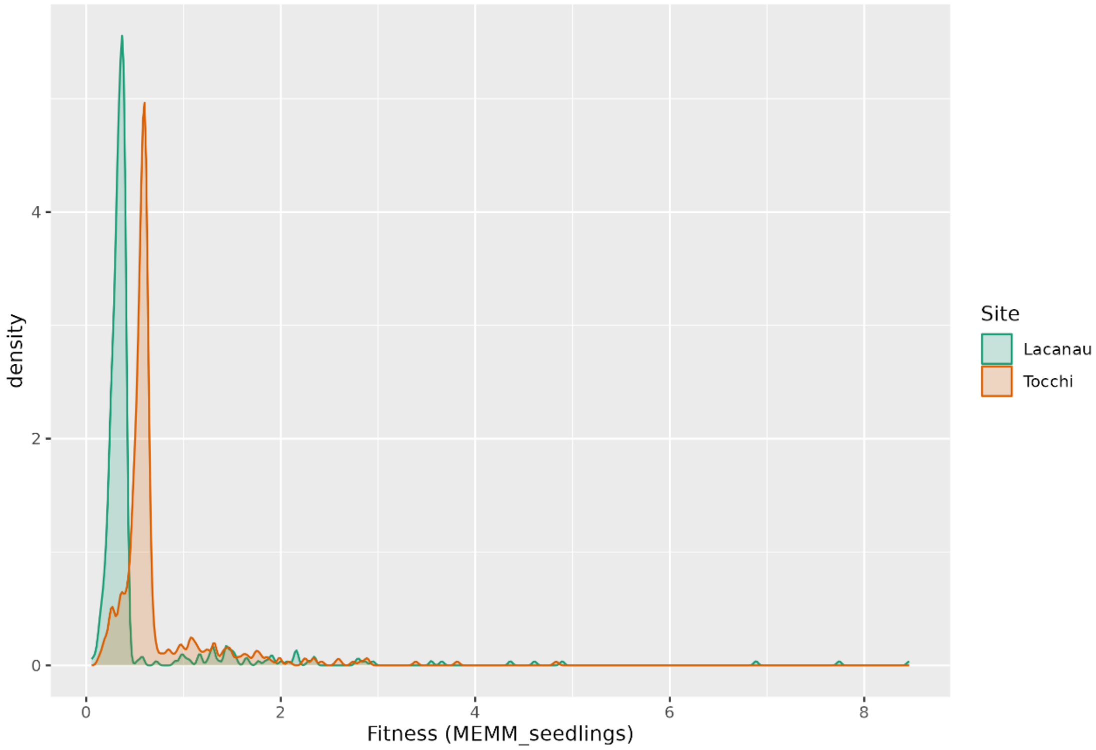


Figure 7: Density curves of the fitness proxies estimated by CERVUS software in both GCUs.

To complement the two approaches described above, we estimated realized fecundity using a third method, implemented in the software NMπ (Chybicki, 2018; Chybicki, 2023). The modelling approach implemented in NMπ differs from previous methods as it is based on an efficient single-step estimation of gamete immigration and dispersal kernels, together with the effects of phenotypic values on fitness (i.e., selection gradients), while minimizing the risk of false positives among phenotypic and ecological determinants of reproductive success.

Output revealed that in both stands, the majority of pollen came from outside the stand (Lacanau: average pollen immigration rate = 95%; Tocchi: average pollen immigration rate = 74%), while more than half of seeds came from inside the stand (Lacanau: average seed immigration rate = 52%; Tocchi: average seed immigration rate = 14%).

The estimates of pollen and seed immigration rates were very similar to those obtained with the MEMMseedlings software.

In Tocchi, female log fecundities ranged from -1.129 to 2.209, with only 3.31% of values significantly larger than zero, while male log fecundities ranged from -1.888 to 2.466, with only 5.58% of values significantly larger than zero (Figure 8). The reliability of our estimates was ensured by acceptance rates for proposed parameter values across MCMC iterations between 29 and 38% (they should be between 25 and 45% according to NMπ user manual).

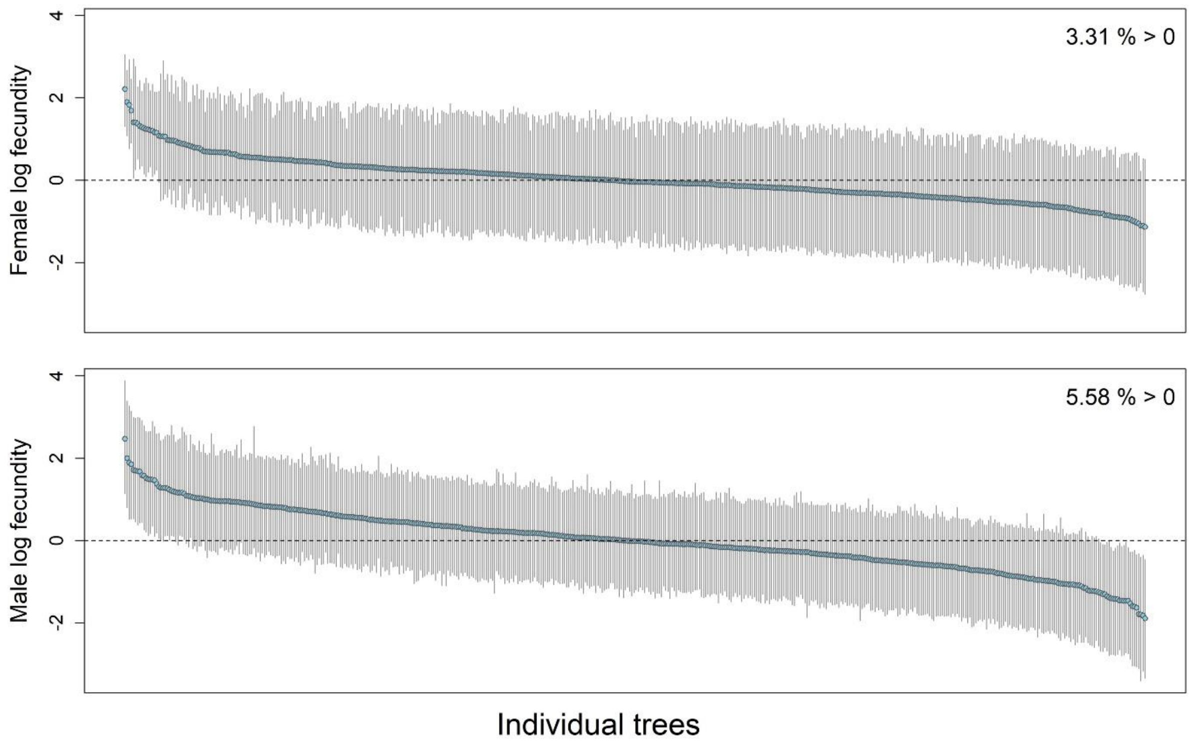


Figure 8 – Individual estimates of female and male realized fecundity as obtained by NMπ analysis for the site of Tocchi.

In Lacanau, female log fecundities ranged from -2.101 to 3.737, with only 6% of values significantly larger than zero, while male log fecundities ranged from -0.084 to 2.228, with no values significantly larger than zero (Figure 9). Unfortunately, acceptance rates for proposed parameter values across MCMC iterations were not between 25 and 45%, indicating that there might be some mixing issues in the MCMC chain.

Overall, the estimates of realized fecundity obtained with NMπ showed a skewed distribution both in Tocchi and Lacanau, consistently with those obtained with both MEMMseedlings and CERVUS.

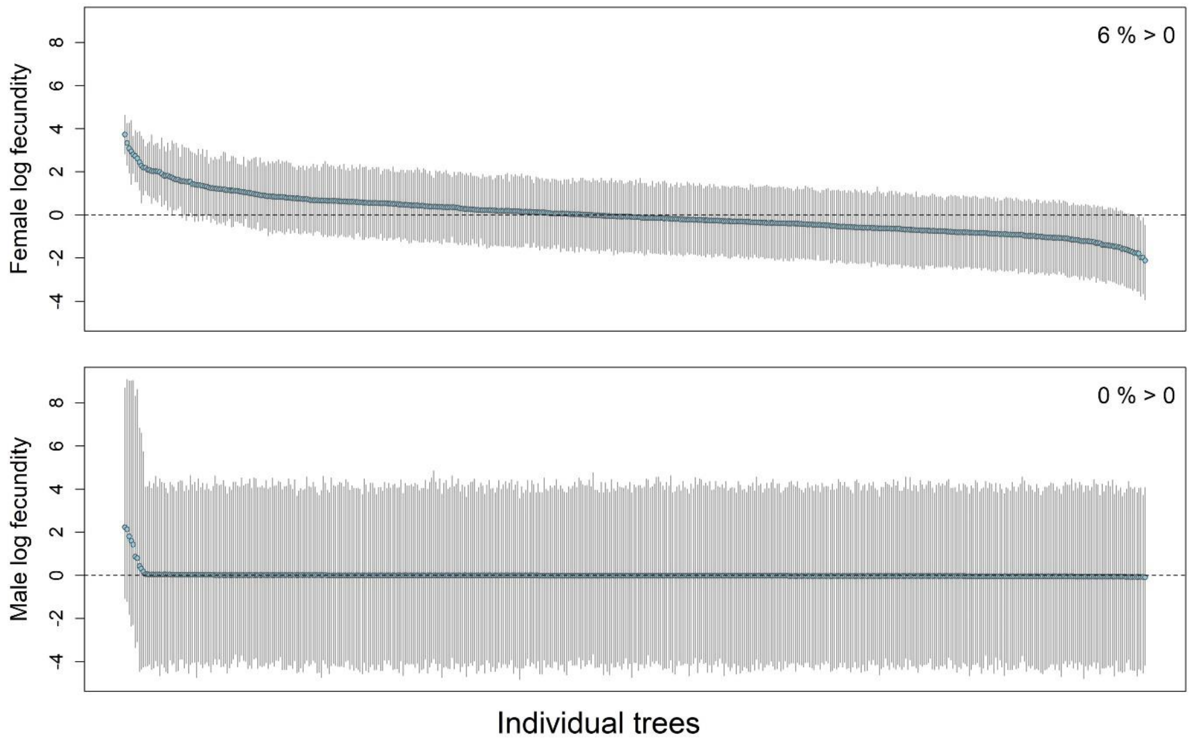


Figure 9 – Individual estimates of female and male realized fecundity as obtained by NMπ analysis for the site of Lacanau.

3.3 Selection gradients (relationship fitness and trait proxy)

We used two types of regression-based approaches to estimate univariate phenotypic selection gradients on the four phenotypic traits currently available (DBH, height, cone count and crown area; the two latter originated from UAV data, see above).
First, following Alexandre et al. (2020), we used common linear and quadratic models to which we added covariates linked to tree competition, soil moisture and spatial autocorrelation of fitness (see description in the previous sections).
Second, we used zero-inflated generalized linear mixed models (zero-inflated GLMMs) to compute phenotypic selection gradients. Zero-inflated models allowed us to distinguish sampling zeros from structural zeros for count data. For this type of model, we used the effective female fecundity calculated from CERVUS software and we added as well the three covariates potentially related to fitness components mentioned above.
Both models gave similar results in the two GCUs, with the exception of the zero-inflated model for height in Tocchi (which was marginally significant whereas the linear model was not). While strong selection gradients were found for DBH (see Figure 10), no significant selection gradient was found for the other three phenotypic traits (except for the zero-inflated model for height in Tocchi). The significant selection gradient for DBH in Lacanau had also a significant quadratic component. This can be interpreted as not only big trees having bigger fitness but also small ones. About the covariates,

both the competition index and the spatial autocorrelation of fitness were significant in most models, indicating some microenvironmental variation in the stands affecting fitness, involving both interaction with other maritime pine individuals (i.e., competition) and abiotic environments.

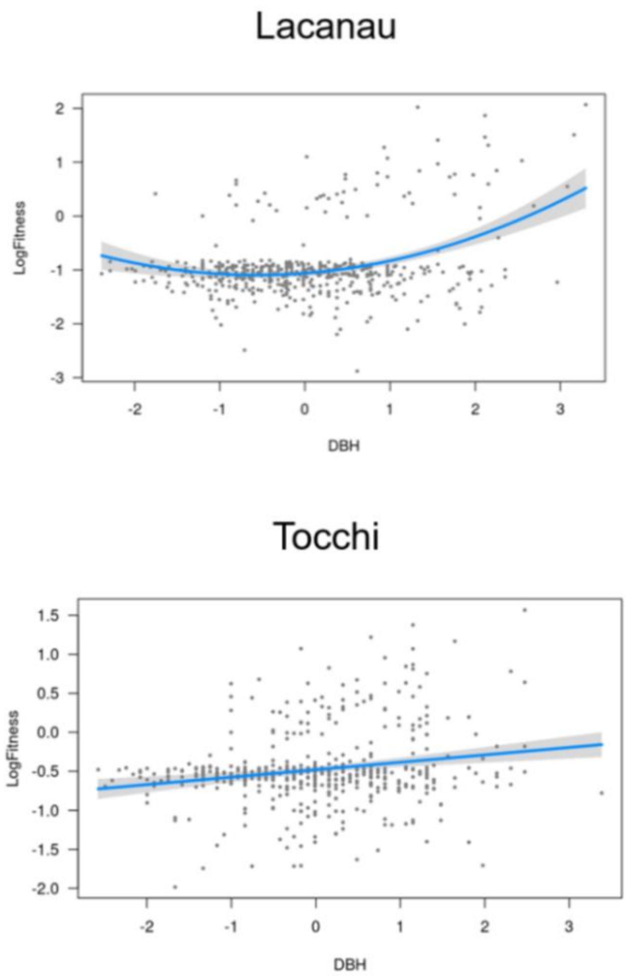


Figure 10: Significant selection gradients for DBH in two *Pinus pinaster* GCUs.

4 Conclusions

Altogether, the results presented here highlight the highly complex approaches used to estimate environmental characteristics and fertility by drones, as well as realised fecundity via seedlings modelling. All these data involved the coordination of many partners and an interdependence of the tasks in order to reach the ambitious goal of WP3, which is to forecast the adaptability of forest genetic resources collections. Up-to-know, output show significant selection gradients in the two GCUs for at least DBH. As more traits are being integrated, we expect to reveal more traits under selection, and we plan to explore multivariate adaptive capacity (i.e., considering all measured traits and their relations) to provide more precise information on the evolutionary potential of these populations.

5 Partners involved in the work

UMR, INRAE, GIS, INIA-CSIC, CNR

6 Annexes

Adams, W. T., & Birkes, D. S. (1991). Estimating mating patterns in forest tree populations. In S. Fineschi, M. E. Malvolti, F. Cannata, & H. H. Hattemer (Eds.), Biochemical markers in the population genetics of forest trees. The Hague: SPB Academic Publishing, pp. 157–172.

Alexandre, H., Truffaut, L., Klein, E., Ducousso, A., Chancerel, E., Lesur, I., Dencausse, B., Louvet, J., Nepveu, G., Torres-Ruiz, J. M., Lagane, F., Musch, B., Delzon, S., & Kremer, A. (2020). How does contemporary selection shape oak phenotypes? Evolutionary Applications 13: 2772-2790. https://doi.org/10.1111/eva.13082

Chybicki, I. J. (2018). NMπ—improved re-implementation of NM+, a software for estimating gene dispersal and mating patterns. Molecular Ecology Resources 18: 159-168. https://doi.org/10.1111/1755-0998.12710

Chybicki, I. J. (2023) NMπ 2.0: Software update to minimize the risk of false positives among determinants of reproductive success. Molecular Ecology Resources 23: 1168-1181. https://doi.org/10.1111/1755-0998.13767.Kalinowski, S.T., Taper, M.L. & Marshall, T.C. (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. Molecular Ecology 16: 1099-1106. http://dx.doi.org/10.1111/j.1365-294x.2007.03089.x

Marshall, T. C., Slate, J., Kruuk, L. E. B., & Pemberton, J. M. (1998). Statistical confidence for likelihood-based paternity inference in natural populations. Molecular Ecology 7: 639-655. https://doi.org/10.1046/j.1365-294x.1998.00374.x

Oddou-Muratorio, S., Gauzere, J., Bontemps, A., Rey, J.-F., & Klein, E. K. (2018). Tree, sex and size: Ecological determinants of male vs. female fecundity in three Fagus sylvatica stands. Molecular Ecology 27: 3131-3145. https://doi.org/10.1111/mec.14770

Dalponte, M. and Coomes, D. A. (2016), Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. Methods Ecol Evol, 7: 1236–1245. doi:10.1111/2041-210X.12575