Understanding Climate Change Implications on Tree Growth

Refiz Duro[1,*], Anahid Wachsenegger[1], Anita Zolles[2], Jasmin Lampert[1], Carlos Landivar[2], Larissa Posch[3], Hanna Koloszyc[4], Martin Gritsch[4], Günther Bronner[5], Albert Villalobos Gasca[5], Sean Cody[6], Franz Martin Rohrhofer[6], and David Conti[6]

[1] AIT Austrian Institute of Technology GmbH
[2] Bundesforschungs- und Ausbildungszentrum für Wald, Naturgefahren und Landschaft
[3] E.C.O. Institut für Ökologie
[4] GeoVille Information Systems and Data Processing GmbH
[5] Umweltdata GmbH
[6] Know Center Research GmbH

*Correspondence: refiz.duro@ait.ac.at

Forests are a significant carbon sink, enabling natural carbon dioxide removal from the atmosphere and helping mitigate climate change and its consequences. It is thus critical to preserve and conserve forests, with the good health of trees being a core element in achieving this goal.

Tree growth is a key indicator of tree health, however, the growth dynamics and underlying physiological processes are highly complex. Growth models have been available for several decades, however, the input data to these models have not always been sufficient or satisfactory due to the limitations of spatial or temporal resolutions and measurement accuracies, resulting in limited power of model outputs. Furthermore, climate change has a considerable impact on growth by introducing climatic extremes (i.e. drought, heat, intense and frequent storms) and phenological patterns (i.e., vegetation season length), disrupting the usual growth cycles, thus further increasing the inaccuracies and limiting the prediction power of existing growth models.

More advanced and diversified measurement devices have become available in the meantime, contributing to a more comprehensive and diverse understanding of tree growth and forest dynamics. The data availability has in turn led to the increased popularity of applying AI technologies and advanced modeling approaches. This positive development, however, comes with its own challenges that range from inaccurate tree positions (making it difficult to join datasets), existing data gaps due to, e.g., power supply issues (introducing uncertainties in training data sets), and integrating data for modeling collected by highly distinct detection methods and devices (increasing model complexities and decreasing understanding of model decisions). The aim of addressing these and additional related challenges has been taken on by the AI4Trees project consortium made up of partners from research and industry, which bridges and combines data science, forestry and ecology research.

Data collection and related challenges

Considering the data, in-situ tree growth measurements using *diameter at breast height* (DBH) are usually performed using **dendrometers**, which are highly sensitive measuring instruments that detect minuscule changes in tree diameter (e.g., 1-5 micrometers). The electronic, automatic dendrometers provide high-frequency data hourly or even more frequently. These data are highly valuable to capture the variations due to daily tree growth cycles (e.g., night vs day) and direct environmental impacts (e.g., frost, rain, hot weather). Medium-frequency tree growth data are collected through manual inspections and read-outs of analog dendrometers once every two weeks, while low-frequency data are collected every 5 years with at-hand instruments. The inherent sensitivity of the measurement tools employed to collect the data, along with human-induced measurement errors, inevitably increase the potential of errors, necessitating a thorough data curation procedure.

An additional approach to directly measure the size of a tree is to use **terrestrial laser scanning (TLS)** that now a days provides ultra-high 3D imaging resolution. The measurements are usually performed when trees are free of leaves to reveal highly detailed variables like trunk size, branch volume, tree height, and canopy structure. The tree growth is revealed by comparing data from consecutive measurement runs with several months or more in between. The TLS approach produces high-quality data, however, issues can be introduced by missing trees (due to, e.g., tree cuts between the measurement runs), or data extraction procedures for single trees. Furthermore, merging dendrometer and TLS data is important to truly benefit from the available data, but co-locating the trees in different data sets can be challenging. Another component important to analyze tree growth is **environmental data**, which are more standard and less challenging compared to the high-frequency tree growth data. These data are regularly collected providing valuable insights and data sets on air and soil temperature, precipitation, as well as moisture at different soil depths. Finally, the tree health and forest conditions can be assessed by using **Earth observation** sensor data through vegetation-relevant indices (e.g., NDVI) providing a possibility for regular and scaled-up forest health monitoring. Additionally, satellite imagery with different spatial resolutions can be used to explore the potential for individual tree detection and monitoring of forest disturbances.

From data to probabilistic to machine learning models

All these datasets contribute to understanding the complexities of climate change impacts on tree growth and forest dynamics. Considering the availability and measurement frequency of different data sets, various modeling approaches were implemented. High frequency dendrometer data were used as the core data of the development artificial intelligence models. The medium and lower frequency tree growth data were likewise used to develop probabilistic and Gaussian process models, as well as to support the AI models by filling (temporal) data gaps and for accuracy corrections. Additionally, sub-meter Very High Resolution (VHR) satellite imagery was used to develop a CNN-based U-Net model to evaluate the potential for detecting tree crowns and tracking changes in forest structure.

The inherent data sparsity in low-frequency measurements forces developing simple probabilistic models that can reliably predict long-term tree growth. Single tree growth over time exhibits largely linear behavior in these managed forests. This simplicity in growth dynamics allows us to implement straightforward linear growth models, enriched with probabilistic frameworks to provide explicit certainty estimates. By incorporating probabilistic modeling, we leverage low-frequency time series data and link them to more costly TLS data through techniques that preserve accuracy and propagate certainty. Furthermore, Gaussian Processes model periodic tree growth using biweekly DBH data and linear trends, capturing high-frequency cycles with uncertainty-aware predictions. Both methods advance climate-sensitive forest management by integrating diverse, cost-efficient data sources with enhanced predictive reliability. In parallel, high-frequency tree growth data were leveraged to develop a machine learning tree growth prediction model using a data-centric paradigm to achieve a high performance AI-model. To achieve this, an analysis of existing open-source frameworks for data analysis highlighted a gap in current tools and the need for tailored solutions to address the specific requirements of forestry-related data. Upon finding the suitable approach, a strong correlation between soil moisture, temperature, and individual tree growth was identified, and the importance of including these environmental factors as well as the tree competition parameters in predictive models was underscored, as they play a vital role in accurately modeling tree dynamics and growth patterns. Lastly, initial machine learning model forecasting results demonstrated high accuracy, providing a robust foundation and serving as a baseline for the development of more sophisticated models.

These insights collectively advance the understanding of forest dynamics and offer a pathway toward more effective data-driven decision-making in forestry.