Implementing Hidden Markov Model using Tree Ring Data

Overview

Coral reefs, glaziers, ice caps, and tree rings allow paleoclimatologists to model the history of the Earth's natural environment [1]. These natural records provide information about temperature, precipitation, and other aspects of climate. As trees grow, they form rings of light wood in the spring and dark wood in the summer, thus the yearly growth of trees can be studied using the width of tree rings. Trees are also an indicator of climate change; ideal climate conditions such as sunlight, rainfall, and temperature enable significantly faster growth, which can be analyzed via tree rings [2]. Studying tree rings from many trees across a biome allows scientists to understand how the climate may be changing and affecting the forest ecosystem [3]. In dry environments (low latitudes or high elevation), tree rings typically record moisture, and in (high latitudes or high elevation), the ring widths are a proxy for temperature [4].

Project Objective:

The objective of the project is to create a Hidden Markov Model (HMM) using proxy tree ring data to describe how the observable events of tree ring growth occurred depending on unobservable climate factors. Moreover, providing a graph which analyzes and gives clear representation of the data manipulation.

<u>Sensor Model</u>: Tree ring data

<u>Transition Model</u>: Climate data extrapolated from the sensor model

Data and Model

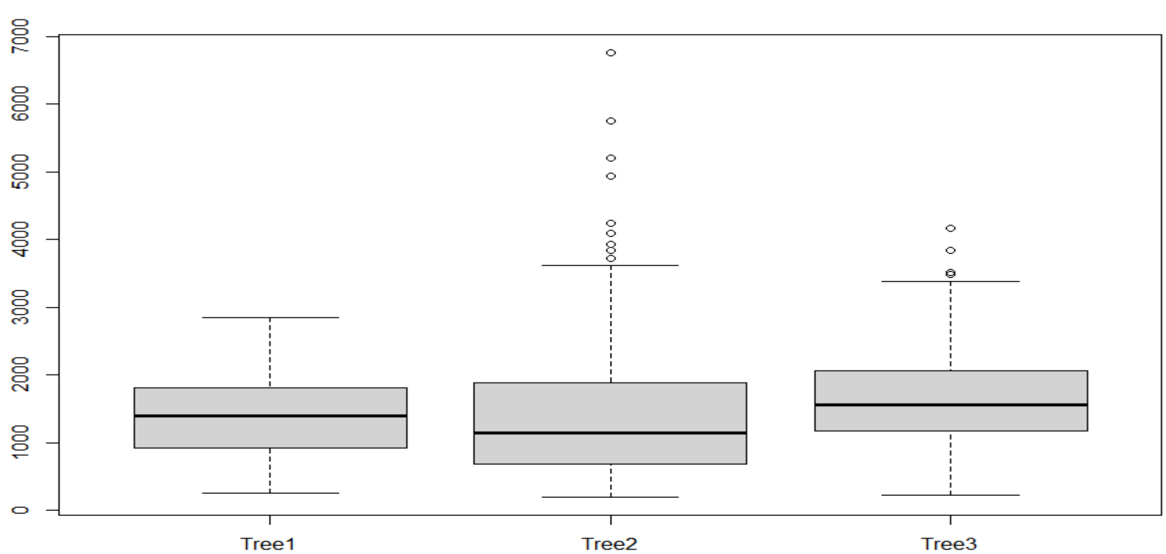
The source data used in this analysis is “Periodicity of western spruce budworm in Southern British Columbia, Canada. Forest Ecology and Management” from the National Centers for Environmental Information. It contains tree ring width data from several trees belonging to the same location i.e., Okanagan Valley, BC

[Paleo Data Search | Study | National Centers for Environmental Information (NCEI) (noaa.gov)]

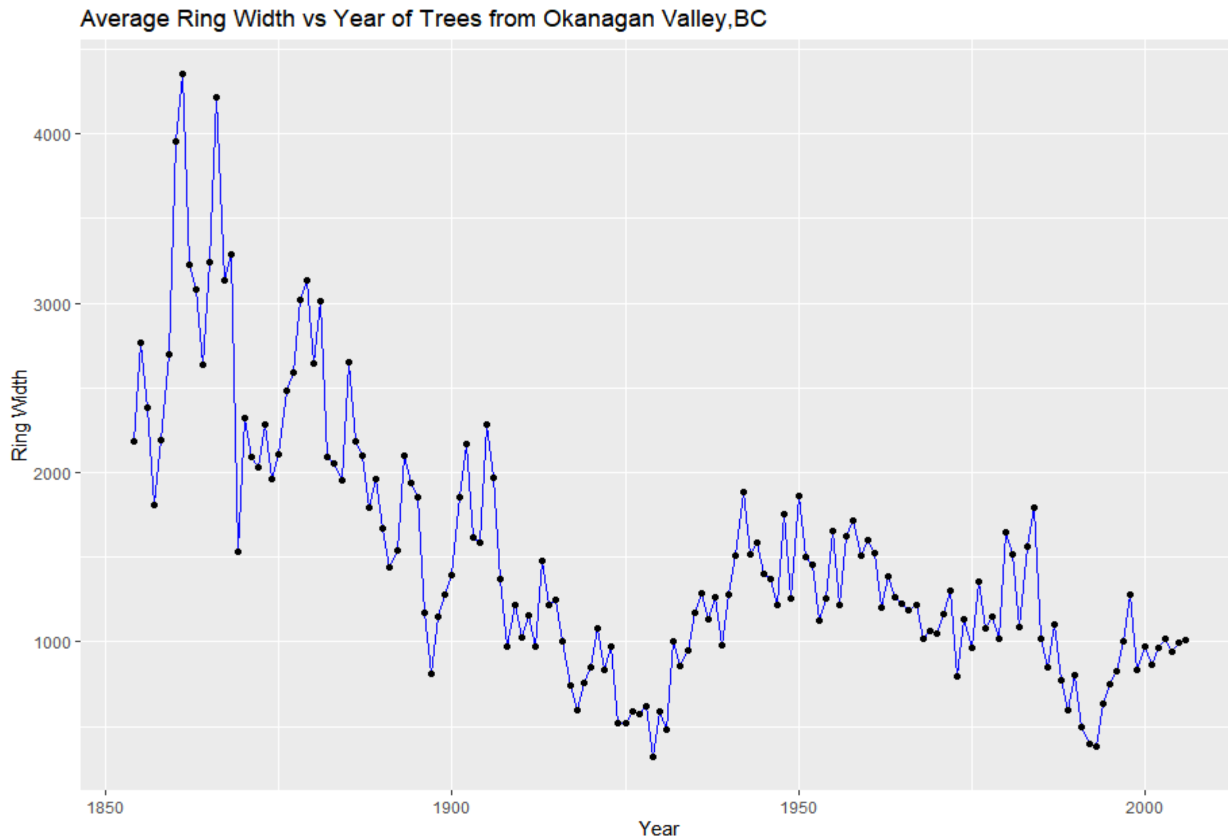
Data of three trees from a similar time period and having a similar age (+- 1 year) was used to calculate average ring width from 1854 to 2006 .

Initial Investigation

Presence of outliers in Tree 2 and Tree3 data, this might be caused by exceptionally favorable climate for tree growth which suggest periods of warm and rainy climate



Average Ring width vs Year plot



After analyzing the plot above we can conclude the following:

- A period of high growth from 1850 to 1875
- Then a period of declining growth rate thus narrower tree rings bottoming at 1930's
- A period of better growing conditions, hence wider rings until ~ 1960's
- Unfavorable conditions again from 1960's until ~1980's

Parameters used to fit HMM

The data parameters were fitted to the HMM tool as follows:

<u>Number of stations</u>

3 (3 trees in the dataset)

<u>Number of sequences</u>

1 (1 dataset per tree per year)

<u>Length of sequences</u>

153 (153 different years worth of data)

<u>Number of hidden states</u>

2 (Significance of this value is uncertain, but has not been observed to affect the simulated data)

<u>Number of components in rainfall distribution per station</u>

3 (tree number, year, tree ring width)

Justification of parameters, sequences and number of sequences

- Parameters- Tree rings width alongside how they've been changing over years were the two main parameters used in the model to analyze the data. These parameters were really useful in finding out the hidden and known states for the HMM model.
- Sequences- The statistics provided by the sampling of data over 153 different years clearly represented how the tree rings width has been changing over the years.
- Number of sequences- Only 1 dataset was used for the observations and was enough for the data manipulation due to its wide range of values and information recorded over hundreds of years.

Justification for model

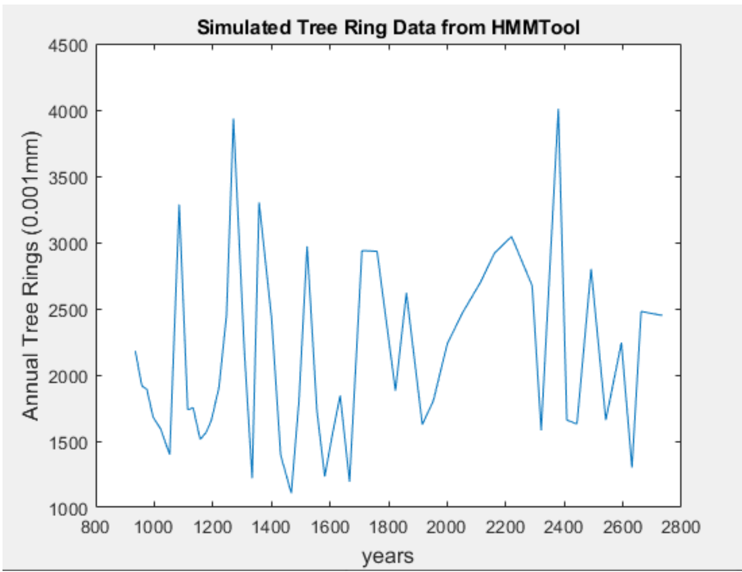
Since there was not enough time to create a predictor file for the Non Homogeneous Hidden Markov Model, the model used is the Hidden Markov Model.

The Hidden Markov Model does a reasonably good job at drawing correlations and parallels between multiple factors and linking them to a result. This means that we can employ it to argue whether or not Climate has had an effect on the size of tree ring width.

As the ring width is not a step-by-step phenomenon but rather a function of a cluster of processes working together to contribute to it, the Hidden Markov Model is a suitable tool to analyze it.

Conclusion

The following graph was produced by plotting the data produced by HMMTool in MATLAB.



The results from using the simulation method of HMMTool follows similarly to the observed values. The spike in width around the year 1860 on the "Average Width vs Year of Trees from Okanagan Valley, BC" graph also appears in the simulated tree ring data as well as the sharp dip after it. The slower increase in the width of tree rings from 1920 to 2200 (predicted by Markov model), as evidenced by lower slope of the graphed line indicates that the change in tree width stabilized over this period, suggesting a consistent climate pattern. Even if only the period between 1800 to 2000 had been considered, no drastic changes in the three ring widths are present. Tree ring widths contain valuable information about weather events, but are not a good indicator of climate change because they are dependent on several factors. These factors include growth spurts owing to forest fires reintroducing nutrients back into the soil and reducing competition for nutrients among trees, and water availability [8].

Problems & their Resolutions

Observing Different Types of Data

Since researchers concentrate on Earth's environment by utilizing satellite and instrumental records, we initially had decided to include this among other wide arrays of ecological records. The fact that life expectancies when utilizing satellite and instrumental information have been somewhat short when contrasted with Earth's life is one test that stands in the way of applying this information. Both the satellite records and the instrumental records are excessively short to concentrate on specific environment processes that happen more than hundreds to millennia.

Data Selection Manipulation

Finding a relatable dataset of a nearby local region and converting the source data into a usable structure. Reducing the effect of confounding variable age by selecting sample tree of similar age belonging to a similar time period.

Choosing parameters for the GUI:

Choosing the correct parameters such as the sequence length proved to be a challenge.

HMM tools

The task of downloading and using the HMM tools presented several challenges due to the lack of updates, documentation and portability. Several tools had been audited and tested with many failures and setbacks.

Sources

[1] Natural Resources Canada. 5 Things You Can Learn From Tree Rings.

https://cfs.nrcan.gc.ca/pubwarehouse/pdfs/38898.pdf

[2] Bauer B. How Tree Rings Tell Time and Climate History. Beyond The Data. November 29, 2018.

https://www.climate.gov/news-features/blogs/beyond-data/how-tree-rings-tell-time-and-climate-history

[3] American Museum of Natural History. Studying Tree Rings to Learn About Global Climate. https://www.amnh.org/learn-teach/curriculum-collections/earth-inside-and-out/studying-tree-rings-to-learn-about-global-climate

- "Studying Tree Rings to Learn About Global Climate"
- American Museum of Natural History

[4] National Centers for Environmental Information. How Do Scientists Study Ancient Climates? March 3, 2016.
https://www.ncei.noaa.gov/news/how-do-scientists-study-ancient-climates

[5] Quinn J. Fitting Hidden Markov Models Part I: Background and Methods. July 3, 2018.
https://waterprogramming.wordpress.com/2018/07/03/fitting-hidden-markov-models-part-i-background-and-methods/

[7] Alfaro, R.; Alexso, J.; Lewis, K.; Bhandari, S.; Speer, Periodicity of western spruce budworm in Southern British Columbia, Canada. January 15, 2020
https://www.ncei.noaa.gov/access/paleo-search/study/27432

[8] Raymond C. et al. Understanding and managing connected extreme events. Nature.com. July

2020. https://www.nature.com/articles/s41558-020-0790-4.pdf

Individual Contributions:

Mikhail Sinitcyn: Project research, dataset research, graphing, report writeup, conclusion

Justin Nakatsu: Inputting data into HMMTools, graphing, reporting results

Simarprit Gandhoke: Dataset research, graphing, data cleanup, HMMTools, data analysis

Janit Kumar: Justification for parameter, sequences and number of sequences, project objective, helped with hidden and known states for HMM.

Nicholas Zhang: Report formatting, report writeup, result analysis

Chetan Khosla: Project research, observing the problems when developing the HMM