A Sampling-Based Approach to Estimate Tree-Ring Width Chronologies

Mads Carlsen

Master's thesis
2025:E59

Lund University

Faculty of Science
Centre for Mathematical Sciences
Applied Computational Science

Centrum Scientiarum Mathematicarum

Master's Theses in Mathematical Sciences 2025:E59
ISSN 1404-6342

LUNFTB-3003-2025

Applied Computational Science
Centre for Mathematical Sciences
Lund University
Box 118, SE-221 00 Lund, Sweden

http://www.maths.lu.se/

A Sampling-Based Approach to Estimate Tree-Ring Width Chronologies

En Samplings-Baserad Metod för att Skatta Kronologier för Årsringsbredd

Mads Egeberg Carlsen

2025

LUND
UNIVERSITY

BERM08

Master Thesis for Applied Computational Science, Physical Geography, 30 credits

Supervisors

Johannes Edvardsson, Department of Geology
Ullrika Sahlin, Centre for Environmental and Climate Research

Mathematics

Lund University

Lund, 2025

Abstract

In the field of dendrochronology, scientists work with the process of ordering different trees in time by using their annual *tree ring width* (TRW). The TRW data are logged in time series and *cross-dated* with each other to find out in which period of time they both were alive. The two TRW series are then conjoined by taking the average of their ring widths in each overlapping year, and this can be repeated adding more and more trees to create a *chronology*. This is done adding one tree after the other, analysing where each individual tree 'fits in', adding the 'best fitting' series, then next best, third best and so forth. This is a time-consuming process, matching TRW series and determining which series fit in and which do not. In this project a workflow was developed with the aim of automatically cross-dating a given number of TRW series in a probabilistic manner, utilising principles from different statistical methods including Markov Chain Monte Carlo (MCMC) and Artificial Bee Colony (ABC). The workflow was developed to sample from the multivariate distribution of the TRW series' datings, and gradually converge towards the distribution showing the TRW series' most probable datings. The developed method proved to be somewhat successful in correctly cross-dating both artificially generated *and* real tree ring data. It furthermore managed to provide insight into the uncertainty of individual trees' ring width data. Although this workflow is an early-stage approach and possible improvements were identified, it could potentially help scientists develop chronologies faster and provide them with uncertainty information, whilst merely acting as a *tool* and not a *replacement* of the scientist.

Keywords: Chronology, cross-dating, dendrochronology, sampling, time series, tree rings, TRW series, uncertainty quantification.

Contents

1 Introduction

Dendrochronology, the analysis of trees' annual growth rings, also known as tree rings, can yield more information than just the approximate age of a tree. While it is true that simply counting a tree's rings can give an estimate of how old the tree is, or was when it died (given that the trunk has all of its rings intact), measuring and analysing the rings more thoroughly can help answer questions about undated pieces of wood as well as environmental changes or the climate in the past (Fritts, 1976; Cook and Kairiūkštis, 1990, Bauer, 2018). The primary concept is relatively simple; if two trees have been living at the same time and in the same approximate area, a section of their tree ring widths (TRWs) should have a similar pattern. For example, if a tree in a forest germinated in 1850 and died 1990, while another tree in the same forest might have started producing rings in the year 1910 and continued living till present day, these trees should in principle have similar TRW patterns in the given period from 1910 to 1990 (Fritts, 1976). That is, if one were to construct time series of the two trees' TRWs and compare them, they would not only be able to tell the approximate age of the two trees. One would also be able to tell the exact year the older tree died by 'matching' these two time series where they 'fit together'. The trees' years of origin, which are derived from the internal *lag* where the series fit together, are from hereon denoted as the TRW series' internal *datings* or simply *dates*.

Matching TRW series in this fashion is possible due to the biological processes in trees and how they grow under different environmental conditions. Trees grow wider rings during more favourable conditions and thinner ones during poorer conditions, mainly depending on moisture availability and temperature (Fritts, 1976; Stoller-Conrad, 2017). Since this relation between TRW and climate conditions is present, TRW series from many trees of different ages can be matched or 'dated' as described above, and used to gain insight in climate variability many centuries or even millennia ago (Fritts, 1976; Cook and Kairiūkštis, 1990; Wilson et al., 2016; Edvardsson et al., 2024). These time series constructed by matching multiple individual TRW series are named *chronologies*.

1.1 The problem

TRW chronologies are constructed by finding the point where the ring patterns of two trees 'match' and computing the average ring width of each year (Cook and Kairiūkštis, 1990). A common approach is to find the 'best matching' pair of TRW series and then successively add the remaining series one by one (Edvardsson et al., 2024). When constructing chronologies, the dendrochronologist performs a lot of the analysis manually, comparing matches between individual series and selecting the best match. This manual inspection is crucial for the quality of chronology construction, and computers are used for the computational and statistical parts of the analysis. Yet, the manual work is still time consuming, and quantifying the *uncertainty* of the data is part of this manual work. Furthermore, since chronologies are constructed by pairwise comparison of TRW series, the result is influenced by the order in which the series are considered. The uncertainty of the points, where the TRWs have been determined to overlap, could have a large impact on the chronology's reliability, as the uncertainty of one TRW series' date could potentially influence the dating of the series added next.

Although there only exists one *true* dating between two TRW series, there might be more than one dating which returns a high correlation. Furthermore, high-correlation datings cannot be expected to occur in a normally distributed manner, but rather in a more sporadic multi-modal distribution. Understanding the nature of a chronology's uncertainty could therefore potentially aid dendrochronologists develop accurate chronologies whilst saving time, and possibly help by automatically identifying weak TRW series showing high uncertainty.

1.2 Aims

The three main goals of this project were (i) helping the dendrochronologist save time, (ii) making a *probabilistic* chronology using all given TRW series *at once* for the dendrochronologist to compare with, and (iii) quantifying a chronology's uncertainty based on its constituent TRW series. Based on these goals, the aims of this project were to 1) develop a probabilistic method to estimate tree-ring chronologies, 2) apply this method on tree ring data to quantify uncertainty in an estimated chronology and 3) compare the results generated with the probabilistic method and the currently applied method using the same data.

A workflow will be developed to sample from the multivariate distribution of dates, and this workflow will be utilised to produce a probabilistic TRW chronology based on the estimated marginal distributions produced by the workflow. In this way it will *aid* the scientist as a tool, not *replace* them or make their current work more redundant. It could potentially make the development of large data TRW chronologies for e.g. climate reconstruction quicker and also contribute with an additional quality check of the data.

2 Theory and Methods

2.1 Current analysis methods

As mentioned previously, it is common to construct chronologies starting with a best matching pair and successively adding TRW series one by one. Here, each individual in a set of TRW series is attempted matched with all other individuals. Afterwards, chronologies are constructed from the best matching pair, where after the remaining individuals are attempted matched against the newly made chronology. This is done by adding the second best matching series, then the third best and so forth, until no more matching TRW series remain (Edvardsson et al., 2024). Furthermore, one TRW series may very well be constructed from two or more individual TRW series from the same tree (measured 90° to 180° apart), having been matched and averaged in order to minimize noise (Edvardsson et al., 2024).

The procedure of matching these TRW series, which is based on both statistical analysis and visual comparisons between TRW series, is called *cross-dating* (Baillie and Pilcher, 1973; Fritts, 1976), and one method used to find the point where they match is called *cross-correlation*. This method returns the correlations and possibly accompanying t-values between two time series on an interval of different *lags* (or dates) between the two series (Baillie and Pilcher, 1973). However, when compiling a chronology a part of the matchmaking still relies on the dendrochronologist's judgment. A visual inspection of the TRW

series is performed by the scientist, as the human judgment is based on a highly evolved skill of pattern recognition (Holmes, 1983; Cook and Kairiūkštis, 1990). Therefore, software like TSAP-Win will present the scientist with multiple 'best matches', i.e., points of overlap where the correlation and significance are highest, for them to assess visually (Rinn and Heidelberg, 2011). When these proposed matches are analysed, a number of different measures are calculated and taken into account. Some of these are the *Gleichläufigkeitswert* (Glk, Eckstein and Bauch, 1969; Buras and Wilmking, 2015) and various versions hereof, a number of different t-values (student's t-value, Student, 1908; TVBP, Baillie and Pilcher, 1973; TVH, Hollstein, 1980), and *Cross Date Index* (CDI) being a combination of the above (Rinn and Heidelberg, 2011; Edvardsson et al., 2024). These are all in place to aid the dendrochronologist determining which of the suggested points of overlap is most likely to be the correct one, some of which might not even be possible. It is possible for a, say 80-year long TRW series to be dated against a chronology to start in e.g. year 1965 based on a program's correlation measures. This would however mean that the TRW series would stretch beyond year 2040, which is not possible as of the current year.

Transforming the TRW data by de-trending and normalising them in various ways is commonly done before analysing and cross-dating them (Baillie and Pilcher, 1973; Hollstein, 1980; Holmes, 1983). This is done as some trees exhibit an underlying trend in ring thickness as the tree ages, and as a result of disturbance and competition (Hollstein, 1980, Cook and Kairiūkštis, 1990).

After compiling the chronology, it is common to check the completeness of the individual TRW series, and a common tool for this is the computer program COFECHA (Holmes, 1983; cited 7,493 times according to Google Scholar on May 24th 2025). A tree ring series could be subject to partially absent or missing rings, as well as double rings (Fritts, 1976; Holmes, 1983) which is what COFECHA helps detect. Upon checking the TRW series for errors, it can be estimated how well the compiled chronology predicts what is known as the *underlying population signal* or *population average*. A measure called *Expressed Population Signal* (EPS; Wigley et al., 1984) is often used for this purpose, albeit sometimes slightly differently than originally intended (Buras, 2017). Based on the limited number of TRW series the chronology is made from, it *approximates* the squared Pearson's correlation coefficient r^2 of the constructed chronology's prediction of the underlying population signal (Buras, 2017).

2.2 A probabilistic sampling-based approach

In statistics, one differentiates between the *sample* and the *population*, where the former is used to estimate properties of the latter. This is also the case for dendrochronology, where the sample consists of a finite set of TRW series which can be used to estimate yearly TRW values in the underlying population signal (Wigley et al., 1984). A chronology constructed from a finite number of TRW series is therefore an estimation of an underlying population chronology in a given time period. The issue at hand is that, unlike in univariate statistics where one might estimate the population average of a given variable, when working with TRW series one does not simply compute an average. The series might be of different lengths and have different datings, but more importantly, some of the internal datings of said series are likely unknown. As already introduced, a TRW series has only one true dating with another TRW series, but finding this manually for a lot of TRW data is time

consuming. It is also not a simple task to do automatically either, as these dates follow an *unknown multivariate* distribution, so finding them directly is a challenge.

Estimating parameters following a specific target distribution is sometimes done by utilising principles from Markov chains and MCMC[1] analysis. Here a chain of samples or *states* is drawn, one state after the other, and every state undergoes an accept/reject test. In this test, every new state is compared to the previous and is accepted with a given *probability* which is determined by comparing the two states in a certain way. This accept/reject test usually considers the target distribution itself as a function of the new and old states, as well as a *conditional* distribution. If this test is set up appropriately, the distribution of the samples in the Markov chain will at some point converge towards the target distribution (Robert and Casella, 2010). However, if the target distribution is unknown, one cannot use this method directly, and an acceptance probability has to be constructed in a way which leads the sampling towards the target distribution. That is, if one can sample states and evaluate them in another way which is somehow characterised by the unknown distribution, an algorithm to estimate the unknown distribution can be established. It could therefore be possible to estimate the distribution of TRW series datings by sampling sets of these and accepting/rejecting them in some appropriate way.

Another method which can be used to solve computational optimisation problems is Artificial Bee Colony (ABC). This general method is based on the behaviour of bees looking for the *best food source*, which in computational optimisation can be translated to the *optimal solution* to a problem (Karaboga, 2009, Akbari et al., 2010). This is suitable when the issue at hand involves finding the global optimum in a solution-space which potentially has multiple local optima. Here bees of different roles are sent out into the field looking for good food sources, some bees being *foragers*, some *onlookers* and some being *scouts*. The role of the foragers and onlookers is looking for sites in a rather strategic way, basing their search area and movements on the *quality* of the current food source. The role of the scouts on the other hand is to look for food sources in a more stochastic manner, not basing their movement and search area on food source quality (Akbari et al., 2010). When talking about *quality* of a food source (possible solution) the amount of *nectar* is used as an analogy for the *performance* of the given solution to the problem (Karaboga, 2009). This could be the solution's mean squared error (MSE), accuracy, or Pearson's r^2 to mention a few. Once good food sources have been found, the onlookers evaluate the best of these, and the neighbourhoods around these best sites are explored for even better sources. If the quality of the previous food source is lower than the new one, the previous site is forgotten and the cycle continues. If a site cannot be improved upon after a set number of trials the site is abandoned (Karaboga, 2009). This method has been found to be successful at finding solutions to optimisation problems, and more so than other optimisation algorithms such as Particle Swarm Optimisation (PSO) and Genetic Algorithms (GA) (Akbari et al., 2010).

2.3 Data and principles

The TRW data in this project consists of TRW series from individual trees being averages of at least 2 samples from each tree. These series have not been normalised nor de-trended,

[1]Markov Chain Monte Carlo

but have been inspected for errors with the program COFECHA. The work in this project is based on data by Edvardsson et al. (2024), who primarily used the previously mentioned Glk and TVBP in the cross-dating procedure.

When building a workflow to estimate a chronology and hence its uncertainty, the fundamental principal in this project was to sample from the multivariate distribution of datings in a chronology. This were to be done by sampling N samples (ideally till the sampled datings converge) in the following steps:

1. Initiate TRW datings and construct the consequent chronology from a K number of TRW series.
2. For every tree[2] k in the K trees:
   (a) Assign multiple changes in dating of tree k.
   (b) Select the 'best' of the proposed datings.
3. Keep these K (potentially) new dates as a *sample* $\bar{d}_n$.
4. Sample $\bar{d}_{n+1}$ based on sample $\bar{d}_n$ according to step 2.
5. The 'goodness' of sample $\bar{d}_{n+1}$ is compared with that of $\bar{d}_n$, and $\bar{d}_{n+1}$ is kept with a probability determined by this comparison.
   - If $\bar{d}_{n+1}$ is accepted, it is kept and now denoted $\bar{d}_n$.
   - If $\bar{d}_{n+1}$ is rejected, it is discarded and *resampled* according to steps 4. to 5.
6. Repeat steps 4. to 5. until $n = N$.

Initiating the TRW datings would be done randomly and the initial chronology would be constructed from a K number of TRW series based on these datings. Utilising principles from the ABC method, these trees would then individually be assigned slight changes in dating, from which the 'best' dating would be selected and kept. After cycling through the K trees, the set of K potentially new positions would then be kept as the *initial sample*, and a new sample would be taken based on this set in a similar fashion as the previous one. As a principle from MCMC methods, each sample after the initial one would undergo a probabilistic test based on its 'goodness' to determine whether it should be kept or not. After a given number of samples, the observed burn-in period (the period before convergence) would be discarded, marginal distributions of the tree-datings could be visualised and a probabilistic tree ring chronology could be constructed based on these. This would provide guidance as to which internal datings might be the *true* datings of each TRW series, and hence what the true chronology might look like.

[2]From hereon, "tree" and "TRW series" are used interchangeably, as one tree is represented by one TRW series.

2.4 Definition of a chronology and statistical model

Let $\bar{c} = (w_1, w_2, \ldots, w_\Lambda)$ be the underlying chronology on an interval consisting of Λ ring width data points (w_λ), being estimated using K TRW series $(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_K)$, each series representing one and only one tree numbered k $(k = 1, 2, \ldots, K)$. Then let $\bar{d} = (d_1, d_2, \ldots, d_K)$ be the internal start datings of the TRW series in a chronology $\hat{c}$, being an estimation of $\bar{c}$. As an example, suppose a chronology $\bar{c}$ were to be estimated on an interval using two TRW series, $\bar{y}_1$ and $\bar{y}_2$, with lengths of e.g. $l_1 = 120$ and $l_2 = 135$ and datings $d_1 = 1965$ and $d_2 = 1995$. The situation would then be given as

$$\bar{d} = (1965, 1995), \quad \bar{y}_1 = (w_1, w_2, \ldots, w_{120}), \quad \bar{y}_2 = (w_1, w_2, \ldots, w_{135}).$$

The approximation of the underlying population signal $\bar{c}$ would then be given as

$$\hat{c} = (w_1, w_2, w_3, \ldots, w_{30}, w_{31}, w_{32}, \ldots, w_{163}, w_{164}, w_{165}).$$

Here w_1 in $\hat{c}$ corresponds to w_1 in $\bar{y}_1$ and w_{165} in $\hat{c}$ corresponds to w_{135} in $\bar{y}_2$. The ring width w_{30} in $\hat{c}$ corresponds to the mean value of w_{30} from $\bar{y}_1$ and w_1 from $\bar{y}_2$, and so forth for all other overlapping values within $\hat{c}$.

Now leaving this example; since the datings in $\bar{d}$ are uncertain quantities they were in this project defined as a vector of dependent random variables. Because they are dependent, their uncertainty can be described by a joint probability distribution. As this distribution is unknown, a simple multivariate normal distribution is chosen as the *prior*. The expected value of the datings would be a K-dimensional point given as a vector $\bar{\mu} = (\mu_1, \mu_2, \ldots, \mu_K)$, and the variance would be described with a covariance matrix $\mathbf{\Sigma}$. This prior is given as

$$\bar{d} \sim \mathcal{N}_K(\bar{\mu}, \mathbf{\Sigma})$$

where K is the dimensionality of the normal distribution, and the covariance matrix $\mathbf{\Sigma}$ is two dimensional with a $K \times K$ shape.

2.5 Sampling

In this section, the statistical method of 'sampling' used in this project is described. This refers to the process of essentially making qualified 'guesses', i.e., drawing sets of numbers which are not quite random yet not completely deterministic. A 'sample' is such a set of numbers i.e., a set of potential datings. The upcoming figures which have the x-axis labelled 'Sample' or 'Sample number' thereby refers to how many of such samples had been taken up until that point, and is somewhat analogical to time.

2.5.1 Initialisation and correlation

Initialising the vector $\bar{d}$ was done by drawing integer elements of the mean vector $\bar{\mu}$ from a uniform distribution given as $\mu_k \sim \mathscr{U}(0, L)$. Here μ_k is the k'th element of $\bar{\mu}$, and L denotes the maximum length of a *continuous* chronology made from the K trees. This is simply defined as $L = \sum_k^K len(\bar{y}_k)$, where $len(\bar{y}_k)$ denotes the length of the TRW series of the k'th tree. This method was taken from the ABC principles, as all bees are scattered randomly

across the sampling space from the start. The covariance matrix $\mathbf{\Sigma}$ was initially set to represent independence between all datings in $\bar{d}$ since the dependence between the dates was unknown to begin with, meaning it was set to be the $K \times K$ identity matrix $\boldsymbol{I}^{K\times K}$. Finally K elements of $\bar{d}$ were sampled from the subsequent multivariate normal distribution using the initialised $\bar{\mu}$ and $\mathbf{\Sigma}$. Upon having sampled $\bar{d}$, the chronology based on these dates was compiled *excluding* the k'th TRW series $\bar{y}_k$, and the Pearson correlation coefficient between $\bar{y}_k$ and the chronology of the remaining trees was computed. This coefficient was calculated on the interval between the dating of $\bar{y}_k$ (denoted d_k^*) and the end of $\bar{y}_k$ being equal to $d_k^* + len(\bar{y}_k)$. This was done for all K TRW series, and this correlation coefficient is denoted *leave one out* (LOO) correlation which was defined as

$$r_k|d_k = \frac{COV(\hat{c}_{-k}|\bar{d}_{-k}, \bar{y}_k|d_k)}{\sqrt{V(\hat{c}_{-k}|\bar{d}_{-k})} \cdot \sqrt{V(\bar{y}_k|d_k)}}. \tag{1}$$

Here r_k is the leave one out correlation of tree k given a specific date d_k. The 'sub-chronology' *without* tree k is denoted $\hat{c}_{-k}$ given the dates $\bar{d}_{-k}$ being the date-vector *without* d_k^*. A similar coefficient is used in the program COFECHA. Note that only values on the interval $[d_k^*, d_k^* + len(\bar{y}_k)]$ where $\bar{y}_k$ and $\hat{c}_{-k}$ o are used.

2.5.2 Date selection

For each TRW series $\bar{y}_k$, several *new* datings were proposed, which were sampled at the 1st, 10th, 30th, 70th, 90th and 99th percentiles of a normal distribution centred around the tree's *original* dating (d_k^*). In the analogy of the bee colony, these were six new potential food sources centred around the original location, which the forager bees would explore. The standard deviation of this normal distribution would depend on the tree's LOO correlation given the original dating.

$$d_{k,\xi} \sim \mathcal{N}(d_k^*, s_k), \quad \xi = 1, \ldots, 6 \tag{2}$$

$$s_k = \frac{L}{20} \cdot \frac{1}{(\|r_k|d_k^*\| + 1)^5}. \tag{3}$$

The subscript ξ denotes the *proposed* or *suggested* dating. The standard deviation's dependency on the correlation of the previous proposed solution is again borrowed from the ABC method, where the movement of a forager bee is dependent on the quality of the food source. The definition of s_k is arbitrarily chosen as being relatively low when the absolute value of $r_k|d_k^*$ is close to 1, and relatively high when it is close to zero. This is to widen the reach of the new date selection if the current position d_k^* returns a correlation close to zero, while making it more narrow when the correlation approaches ± 1. The reason for making expression (3) symmetric around zero is, that when the correlation is close to -1 at a given dating, a relatively high correlation could be expected at another dating close to the current one. This is furthermore the reason why s_k is set to never approach zero on the interval where $r_k|d_k^* \in [-1, 1]$, as the neighbourhood around d_k^* should be explored for candidates returning a high LOO correlation if $r_k|d_k^*$ is close to -1.

Additionally, two 'wildcard positions' were considered as well as the six drawn from the normal distribution (2) in order to better prevent getting stuck in a *local optimum*. That is,

the situation where $r_k|d_k^*$ is relatively high and no better d_k exists in the local neighbourhood[3] of d_k^*, whilst the optimal d_k exists outside this neighbourhood. These suggested dates had the same function as the scout bees in the ABC method, and as these scouts are supposed to be relatively few compared to the foragers the number 2 was chosen. These wildcard positions were drawn from uniform distributions given as $d_{k,7} \sim \mathscr{U}(d_k^* - L, d_k^*)$ and $d_{k,8} \sim \mathscr{U}(d_k^*, d_k^* + L)$. That is, one wildcard position on each 'side' of d_k^*. This would provide a much wider range than the one provided by the distribution described by expression (2) and would serve as a potential way out of a local optimum.

2.5.3 Sample acceptance

Upon having sampled eight potential datings for the given tree, the LOO correlations for that tree given these datings were calculated. The dating which resulted in the highest LOO correlation (also considering the original dating d_k^*) was the one to be kept, which is unlike the ABC method. In ABC each position is kept with a certain probability depending on their quality. After finding the optimal date for the given tree, the LOO correlations of all TRW series in the chronology were recalculated. The next tree was then given eight new date proposals and the cycle continued till the last tree had been considered. The final set of datings was the *initial sample* $\bar{d}_0$ and the average LOO correlation of this sample $\langle r \rangle_0$ was calculated. The next sample was taken based on $\bar{d}_0$, cycling through all the trees, suggesting and updating dates like described above.

As sample one ($\bar{d}_1$) had been taken, it was *accepted* with a certain probability p_A. This was when the accept/reject stage from MCMC took place. If accepted, sample $\bar{d}_1$ would lay the foundation to the sampling of $\bar{d}_2$, just as $\bar{d}_0$ had to the sampling of $\bar{d}_1$, and if rejected $\bar{d}_1$ would be resampled. This acceptance probability should lead the sample chain towards the desired distribution of the dates, which in theory should correspond to the locations resulting in the highest correlation between the TRW series. It was therefore set to depend on the current and previous samples' average LOO correlation, and p_A was hence defined as

$$p_{A,n} = \frac{\langle r \rangle_n + 1}{\langle r \rangle_{n-1} + 1} \wedge 1, \quad n > 0. \tag{4}$$

That is, if the average LOO correlation of a sample n was better than the previous sample's LOO correlation, sample n would be accepted. If not, the sample n would be accepted with a probability determined by how much higher the previous sample's average correlation was compared to the current sample's. 1 is added to both the numerator and the denominator to make the fraction positive in any given situation. In the unlikely case that $\langle r \rangle_{n-1} = -1$ sample n would be accepted regardless of $\langle r \rangle_n$.

2.5.4 Additional measures

To further help prevent getting stuck in a local optimum, the date which was kept was selected *randomly* from the nine dates *disregarding* their LOO correlations entirely if the sampling seemed to have converged. This was determined to be the case if the average LOO

[3]The range between the 1st and 99th percentile of the distribution (2).

correlation had remained the same for 500 samples in a row. This should in theory help the workflow explore new arrangements of the TRW series, and consequently lead it towards a new and possibly better optimum. If no new optimum would be found, it should tend back towards the previous one. It would also risk taking the sampling out of the *global* optimum and into a local one, and so it was assumed to populate the marginal distribution of each tree's datings, and find the true dating more often than any local optimum.

Additionally, gaps in the chronology were expected to appear and remain as the sampling would go on, due to there being no restrictions on minimum number of overlapping years. These gaps were therefore removed every time a sample was accepted but before the sample was saved. This was done by assigning the date of the TRW series following a gap to the first year of the gap, meaning it would not change the average LOO correlation (see fig. 1). This was done whilst ensuring that the relative positions of any following TRW series would be preserved and no *additional* overlap was introduced.

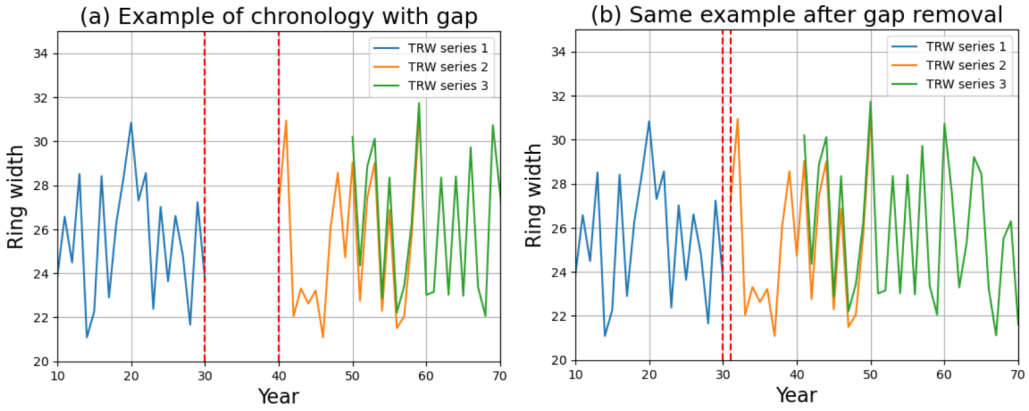

Figure 1: Showing the principle idea behind the process of gap removal. The initial gap in this arbitrary example is 10 years wide (a) and is reduced to just one year (b). It is not reduced to zero years in order to not introduce any additional overlap which would otherwise occur between TRW series 1 and 2. Note that this is purely an illustration of the gap removal process, and that the series are not real TRW data.

In order to keep the datings' reference point constant the longest TRW series was selected as the *anchor* series, keeping its dating fixed and moving the other series relative to this anchor. Furthermore, in order to keep all dates positive the anchor date was *boosted* so that it would have the constant value of $L \cdot 3$, which would ensure that the estimated chronology would never reach dates below zero as long as gaps were removed regularly.

The fundamental assumption behind this methodology was, that all the TRW series would more often than not fall into their respective right places as sampling would go on. If this assumption were to be correct, the chronology given by the emerging marginal distributions would be the *optimal* one. The marginal sampled distributions should thereby change less and less with each sample.

2.6 Post sampling

The estimation of the chronology and its subsequent uncertainty was done upon completing the sampling, by presenting the trees' marginal distributions as violin style ridge plots. Before constructing graphics, the *burn-in* period was determined by visually inspecting the evolution of the average LOO correlation and selected individual trees' dates. This period being the interval before possible convergence was removed from the sampling to potentially make the emerging distributions more noise-free. The chronology $\hat{c}$ was constructed based on the datings found at the most prominent mode in each tree's marginal distribution. In principle, if a tree's dating-distribution would show one clear and narrow peak, the tree's true dating would have a *low* uncertainty. Conversely, if a tree's distribution would be wide and show no or multiple peaks its true dating would have a relatively *high* uncertainty. The anchor series would in principle be the only tree which uncertainty would not be apparent from its resulting distribution, as this tree would always take on the same dating.

The output of this workflow was thereby intended to be a visualisation of the trees' probabilistic datings and the variation hereof. This could aid the user in the process of chronology construction by visually presenting the trees' most probable datings and their uncertainty in the form of their marginal distributions.

2.7 Data generation and workflow application

In order to test the workflow, a number of test series was generated artificially where the true datings were known. The true arrangement of the test series was generated such that every series would have an overlap of at least 75 years with at least one other test series, meaning this arrangement would have no gaps. Three datasets of such test series were generated, where the number of test series (K) were 5, 10 and 15, which would help gaining insight into how well the workflow would perform on gradually larger and larger datasets. These three datasets were made by first creating an underlying test chronology $\bar{c}_t$, which is analogue to the previously introduced $\bar{c}$. These underlying test chronologies were generated by randomly selecting starting points between 0 and 200, whereafter every other point was sampled from a normal distribution centred around the previous one and with a standard deviation of 5. A 1000 years long chronology $\bar{c}_t$ was generated, and the value of $min(\bar{c}_t) + 10$ was finally *subtracted* from every value in order to ensure all values were at least 10 units above 0 (see example in fig. 2).

From this chronology each test series in a dataset was taken as a subset of data points in a given range, and a normally distributed error ($\mu = 0, \sigma = 5$) was added. The test series were distributed such that $K - 2$ series were evenly spread out with an overlap of 75 years with the next series. The start dates of the remaining two series were then selected randomly between 0 and the last test series' start date - that is, the two last series would be *contained* by the $K-2$ other series (see fig. 3). This element of randomness was introduced to examine whether series which overlap with potentially more than two other series would be correctly dated as 'well' as the others. All series were 120 years long, except for one which was 150 years long. This longer series would act as the anchor series previously described, and it was chosen to be a series with a true dating *well inside* the master chronology. This was done in order to examine its effect on the series 'surrounding' it, that is to see if the series with true

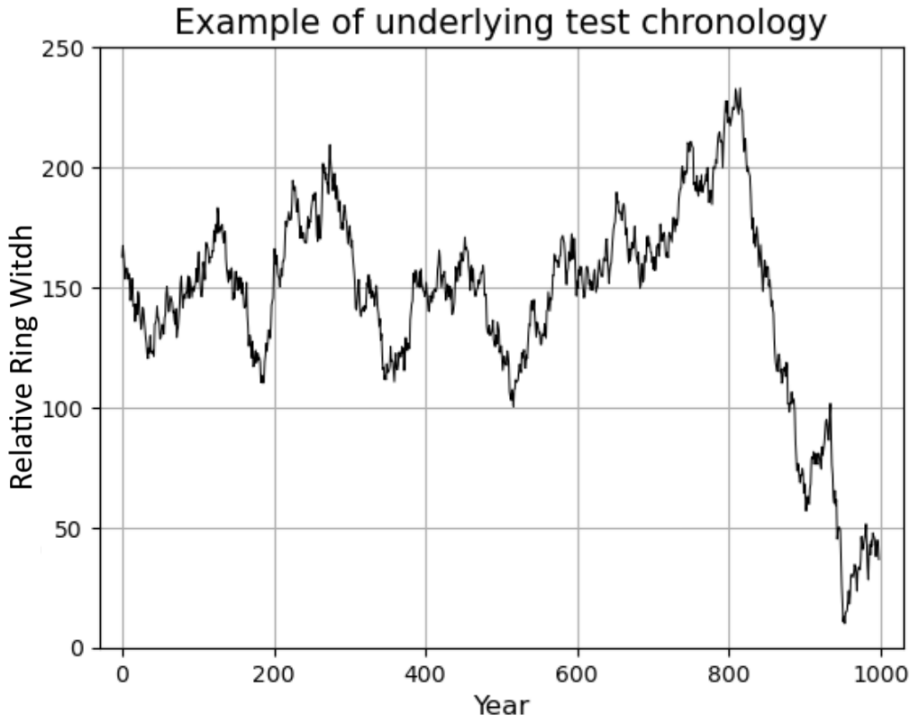

Figure 2: Showing one of the underlying chronologies in its full length. It is this time series that one of the four test datasets was based on. Note that the y-axis lacks a unit, due to this being simulated data. The scale is somewhat representative of actual data, where the unit typically is in 1/100 mm.

datings close to that of the anchor series seemed more stable than others. Four such datasets were generated, where one test chronology would lay the foundation to three test datasets consisting of 5, 10 and 15 TRW series each - a total of 12 sets of test series. Four datasets, rather than a single one, were generated to better capture the workflow's performance.

Additionally, a copy of every test dataset was made and an *odd* series was introduced, which would have a maximum cross-correlation of strictly less than 0.7 with the underlying chronology, with a minimum overlap of 5 years. This addition was done in order to test the workflow's ability to detect a series which does not 'fit in' with the rest. These odd series were generated in a similar fashion as the test chronologies, with the only difference being the length of 120 years instead of 1000. Once these test datasets had been generated, the workflow was run using every dataset - with and without the odd series. The results were then plotted and the workflow's performance was assessed according to the resulting date-distribution of each series and their respective true dates.

Finally, the workflow was applied to datasets from actual trees. It was initially applied to data on tree rings from trees which all had true dates relatively *close* to each other. A dataset of 10 TRW series taken from *living* trees, hence with known dates, was sampled and the resulting uncertainty was interpreted in the same fashion as with the test series. The workflow was then used to estimate the uncertainty of another dataset of real TRW data, only now from trees which were not necessarily living. This meant that the true dates were potentially more 'spread out' than those of the living trees. The results were

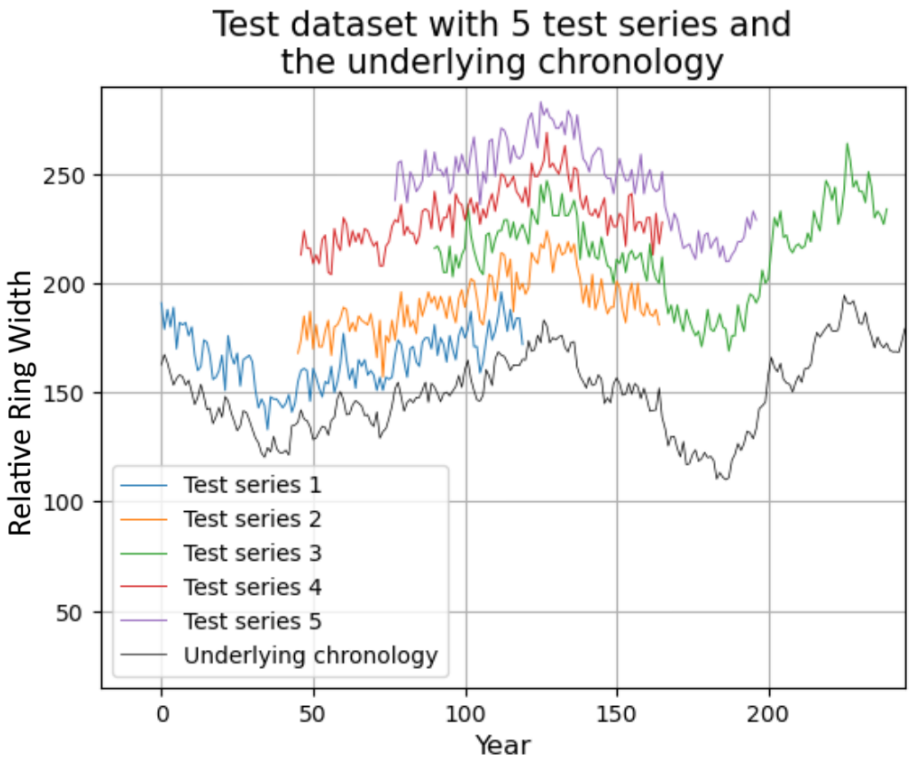

Figure 3: Showing one of the the datasets with 5 generated test series ($K = 5$) as well as the test chronology as an example. The 'odd' series is not visualised here as it has no true dating. Note that the series have been separated with 20 units (corresponding to 0.2 mm) along the y-axis in order to better display them and their relative positions. Also note how *Test series 4* and *5* are fully 'contained' by *Test series 1* to *3* as described above.

interpreted as in the two previous scenarios, except for the date comparison as these were unknown. The resulting probabilistic chronology was however compared to the one compiled by a dendrochronologist using the same data. The workflow application on different kinds of real data was done to test its capabilities on series which are 'stacked' (in is the case of the dataset with living trees) and series which are potentially more scattered.

3 Results

3.1 Generated data

Before plotting the data the convergence and burn-in periods of the data were tested by running the workflow with one generated dataset without odd series twice, both with 5, 10 and 15 series.

The number of samples for each number of trees were set to 3500, 7000 and 11000 for 5, 10 and 15 trees respectively. As can be seen in figures 4, 5 and 6 below, it was not clear at which point the burn-in periods had ended, as neither the average LOO correlation nor most of the selected tree datings ever seemed to properly converge. In some cases however, some datings were apparently more favoured than others, which is apparent in figure 4d and figure 6d. In the case of the 5-series datasets a 1000 sample burn-in period was deemed appropriate, as it would exclude some of the seemingly erratic behaviour in the beginning of the sampling (see fig. 4b and c).

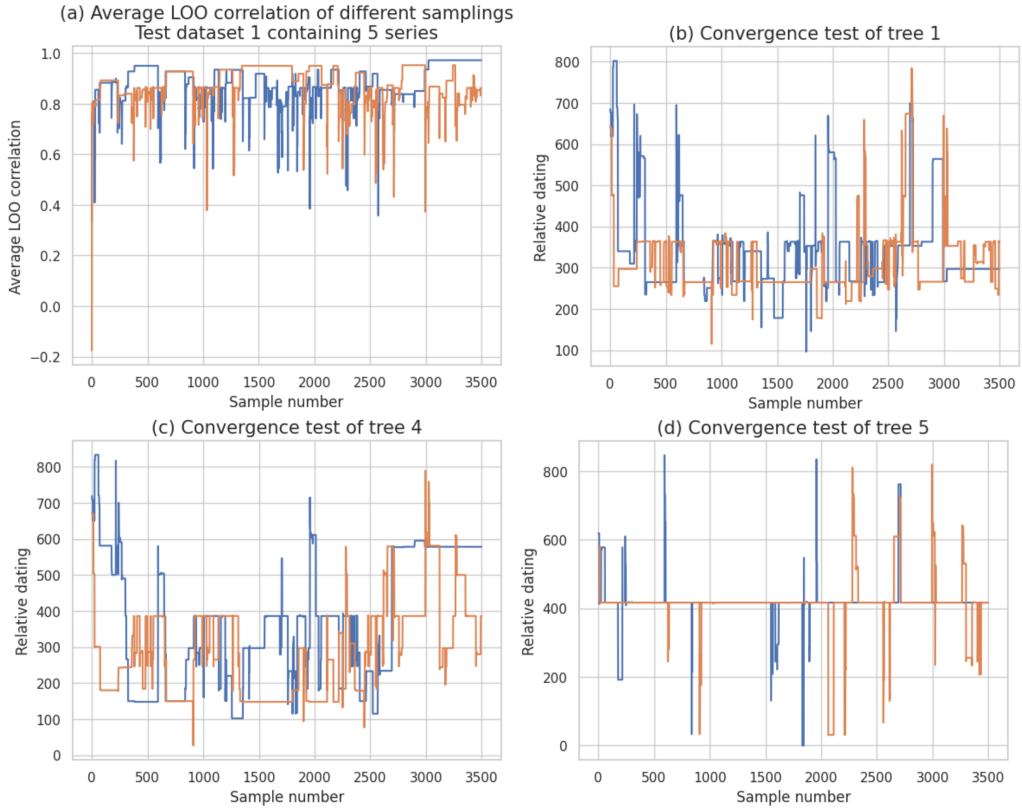

Figure 4: Showing the evolution of the average LOO correlation (a) and the datings of three selected test series (b-d) from two workflow runs both using the same 5-series dataset.

As for the 10- and 15-series datasets, this distinction between burn-in and noise was even more challenging. Therefore, the burn-in periods of these datasets were set to be around half of the datasets, more specifically 4000 samples for the 10-series dataset and 5500 samples for the 15-series dataset.

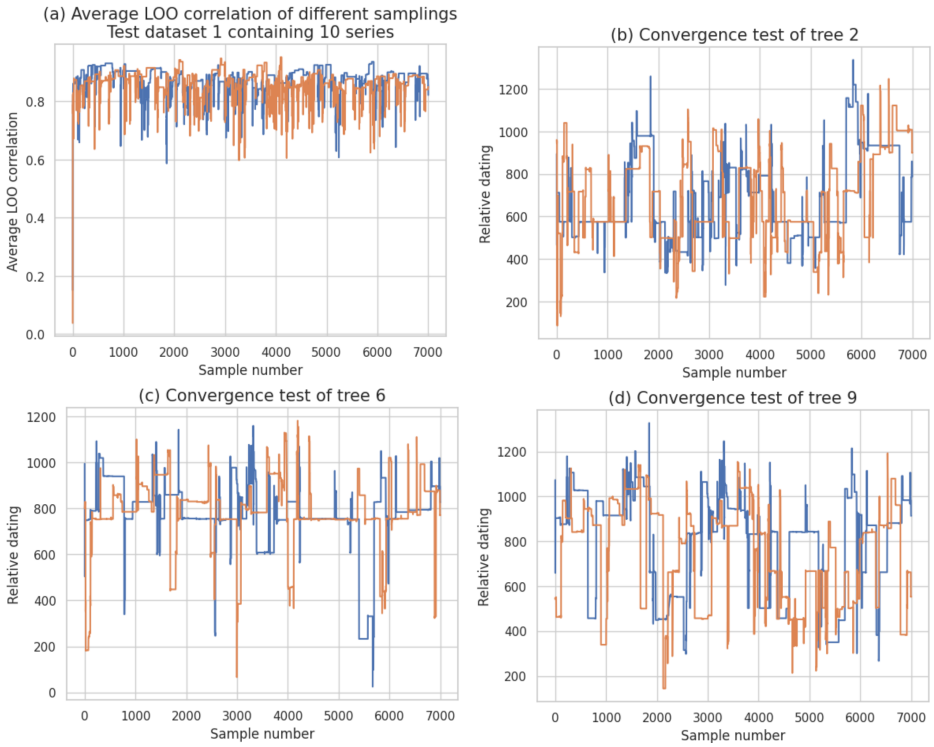

Figure 5: Showing plots of the same concept as figure 4 but using a 10-series dataset.

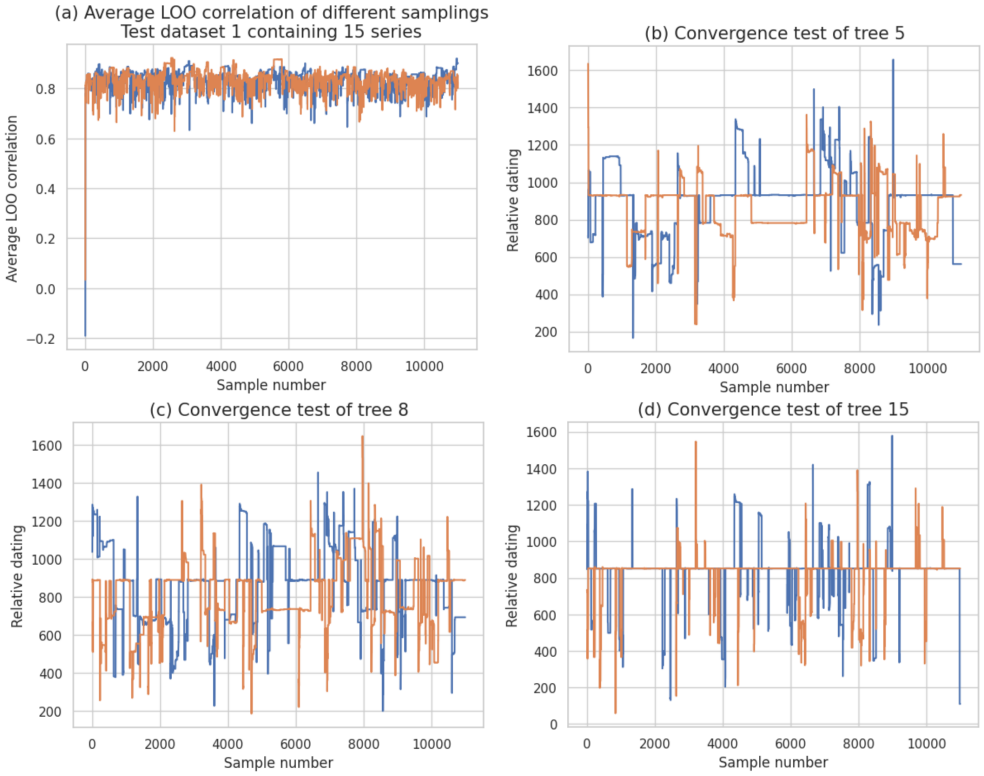

Figure 6: Showing plots of the same concept as figure 4 and 5 but using a 15-series dataset.

Plotting the remaining data after removing the proposed burn-in samples yielded the results shown in figures 7, 9 and 10. Beginning with the smallest test datasets, consisting of five test series each, it was apparent that no series (ignoring the static anchor series) showed consistent variability. All four non-static series showed low variability throughout at least one sampling, and all true datings were in three out of the four samplings located within apparent modes of the marginal distributions (see fig. 7). That is, most trees seemed to have 'found' their true relative dating throughout these three samplings. Tree 1 displayed the 'highest' degree of uncertainty throughout the four runs, and this tree happens to be the one which true dating lied farthest from that of the anchor series. Conversely, the series closest to the anchor did seem to be comparatively stable (trees 2 and 5 in fig. 7a, trees 2 and 4 in 7b, and trees 2 and 5 in 7d), except for in dataset number 3 (fig. 7c) where all non-static series displayed a comparable uncertainty. Trees 4 and 5, which were assigned true starting dates randomly between those of Tree 1 and 3, did not seem any less stable than the ones with evenly spaced true dates.

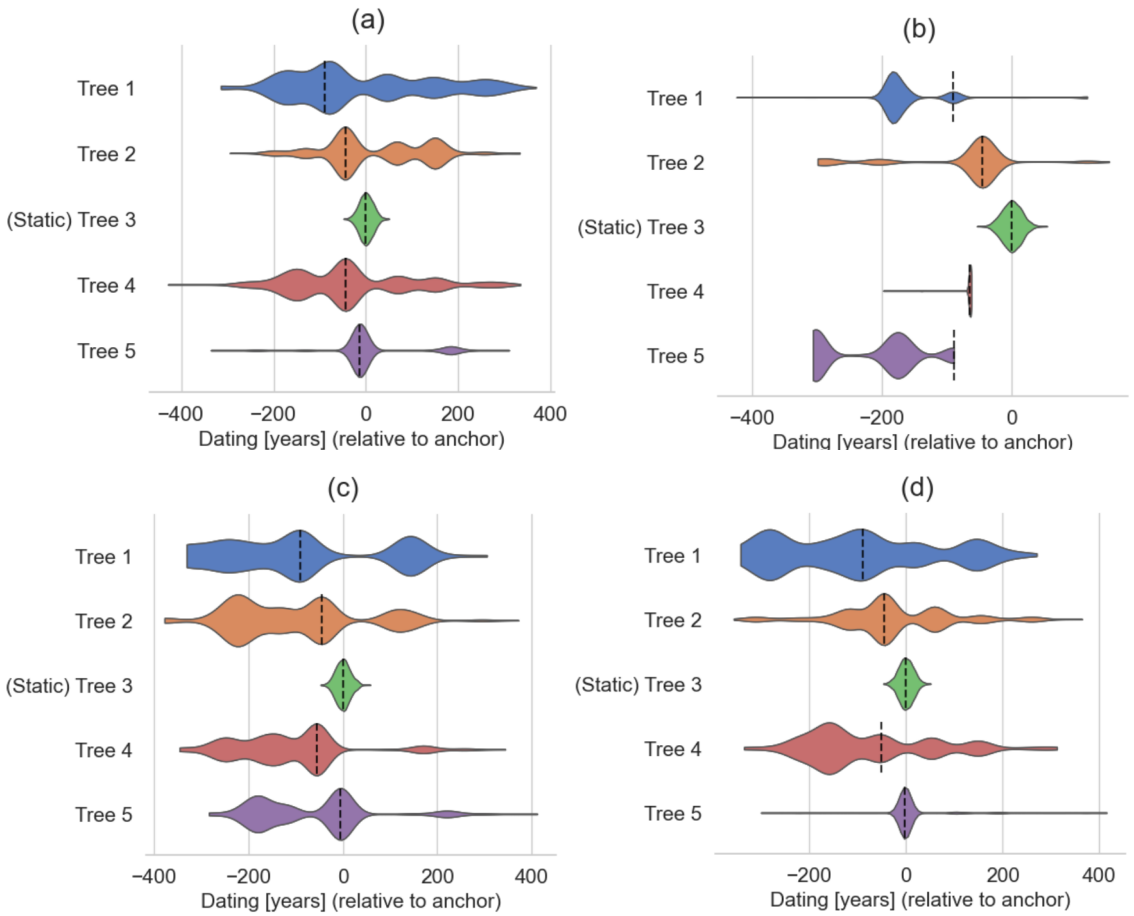

Figure 7: Showing the marginal distributions of test series datings across test datasets containing 5 series. Every tree's true dating is shown as a vertical dashed line. Note that the static tree has normally distributed dates. This is artificially added error which allows this tree to be visible in the plot.

As figure 7b shows, the distribution of tree 5 barely covered its true dating, and the distribution of tree 4 was very narrow. However, when plotting the full dataset *including* the proposed burn-in period tree 5 showed a distribution covering the true date (see fig. 8). Furthermore, the distribution of tree 4 is more clear and tree 1 has a larger peak at its true dating.

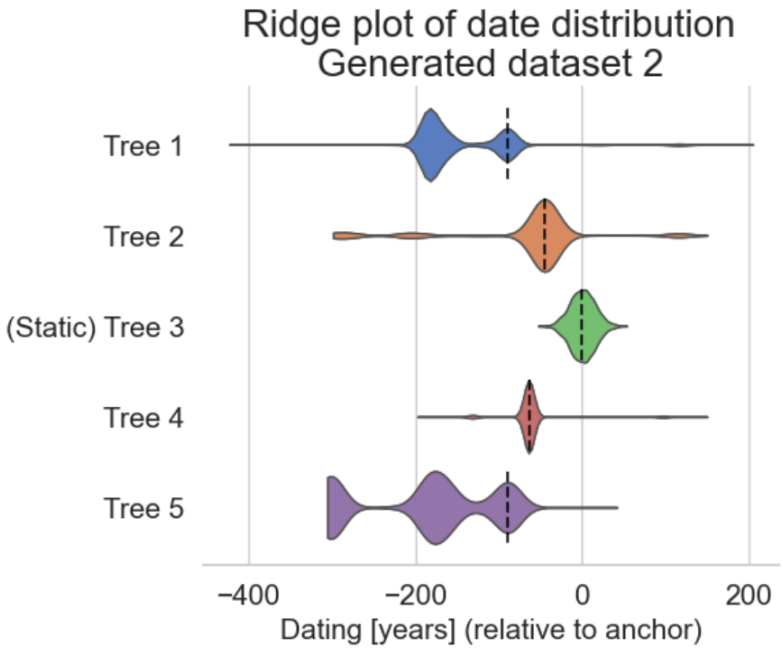

Figure 8: Showing the marginal distributions of the 5 trees in the 2nd dataset *including* the burn-in period.

Applying the workflow to the 10-series datasets yielded the results shown in figure 9. Here it can be seen how, as in the case with 5 series, some trees display high uncertainty in some runs, whilst being very certain in others. Tree 5 is an example hereof, which showed relatively low uncertainty around its true value in run 1 and 4 (fig. 9a and d) but high uncertainty in run 2 and 3 (fig. 9b and c). Running 10 series did furthermore not seem to consistently display a relation between series stability and proximity to the anchor.

Finally displaying the distributions of the four workflow runs using the 15-series datasets (fig. 10) showed some series having constant low uncertainty, but most series having a varying degree hereof. An example is tree 5, which in figure 10d has one clear peak, whilst having two in figure 10a. In 10b the uncertainty is also quite low but the peak appears earlier than the true dating, while in 10c the tree's uncertainty is relatively high. In other instances, such as trees 10, 11 and 12 in figure 10a, the trees seem to have found the true dating relative to *each other*. This is shown as large peaks occurring in what appears to be the right order (and with approximately the right number of years in between peaks) only shifted back in time. In three out of four runs (fig. 10a, b and d) the most stable series seem to be the ones with true dating in the proximity of the anchor series. In figure 10c however, almost all the non-static series seemed comparatively unstable. Across all datasets, the two test series with randomly chosen true dates were not observed to be any less stable in general than the evenly distributed ones.

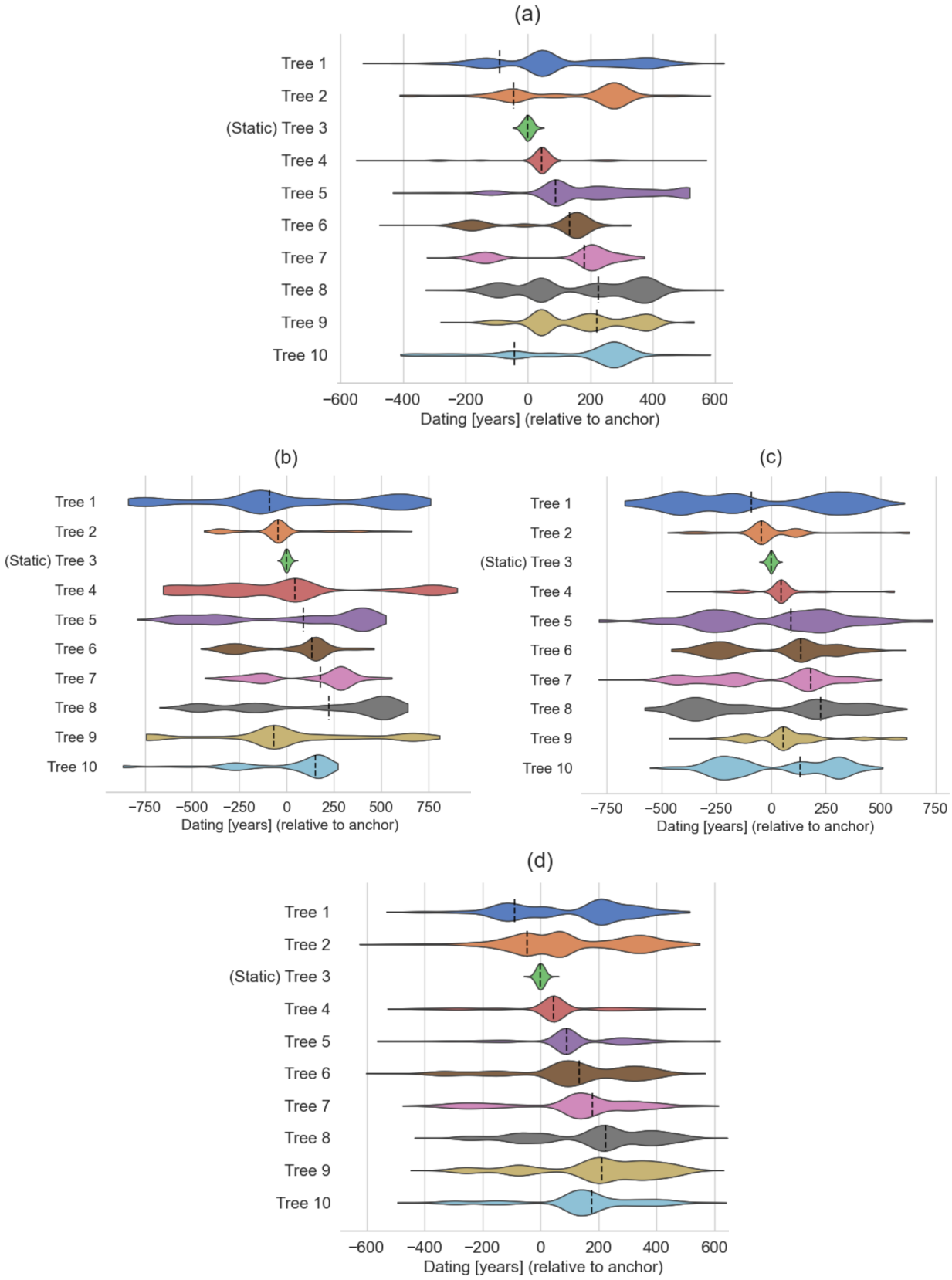

Figure 9: Showing similar results as figure 7 using the 10 series datasets.

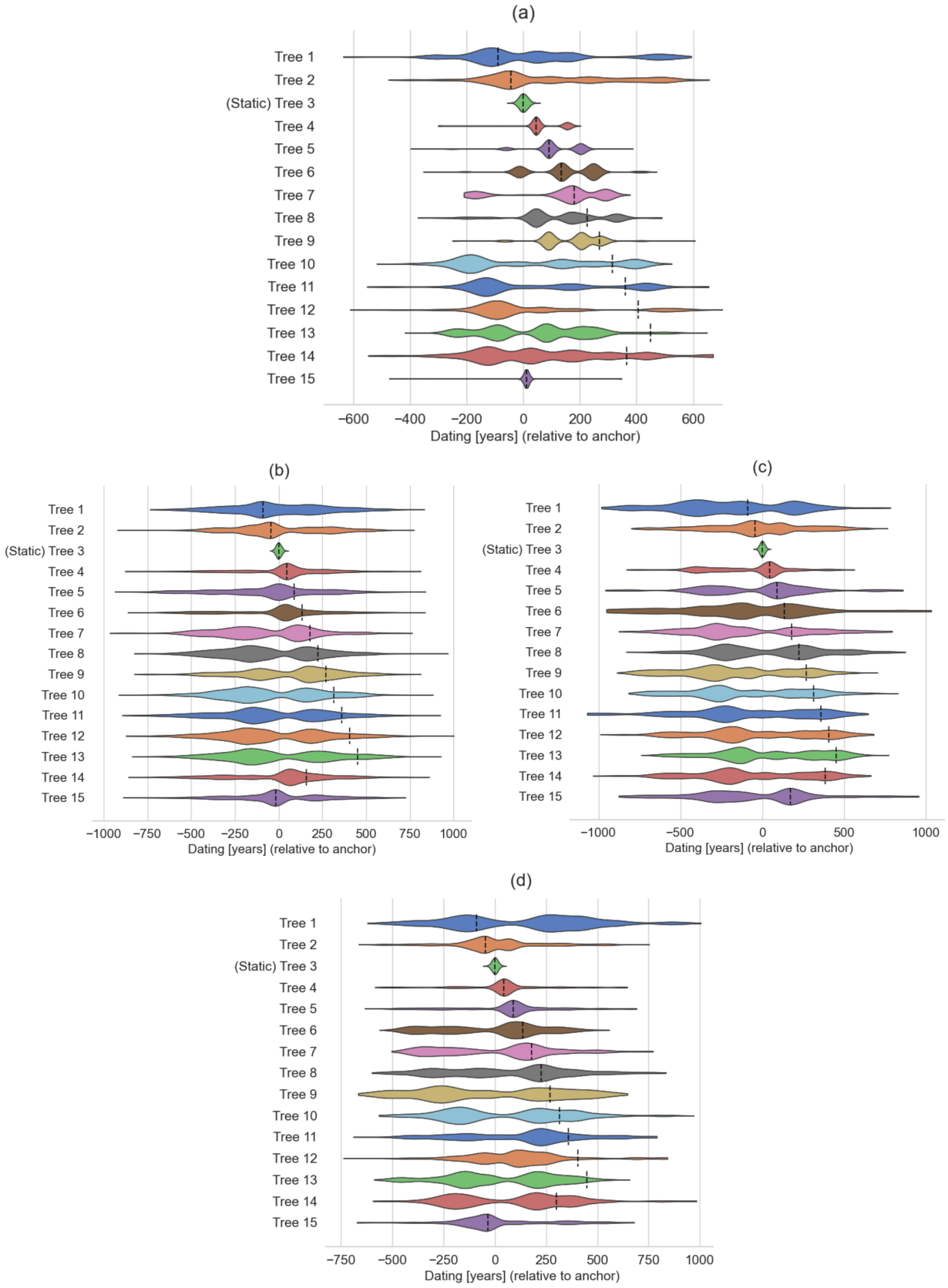

Figure 10: Showing similar results as figure 7 and 9 using the 15-series datasets.

3.1.1 Introducing the odd series

When introducing an odd test series to the datasets, simulating a situation where a TRW series does not fit in with the others, the same numbers of samples and burn-in periods were applied.

In the case of the 5-series datasets, the 6th tree, being the odd one out, showed varying degrees of uncertainty (see fig. 11). When running the workflow using the second dataset the odd tree seemed to display the highest uncertainty, but otherwise had a somewhat comparable or even lower uncertainty than the other non-static test series. It was only in two out of four datasets (datasets 1 and 4; fig. 11a and d) that all trees' dating distributions managed to capture their respective true dating within reasonably apparent peaks.

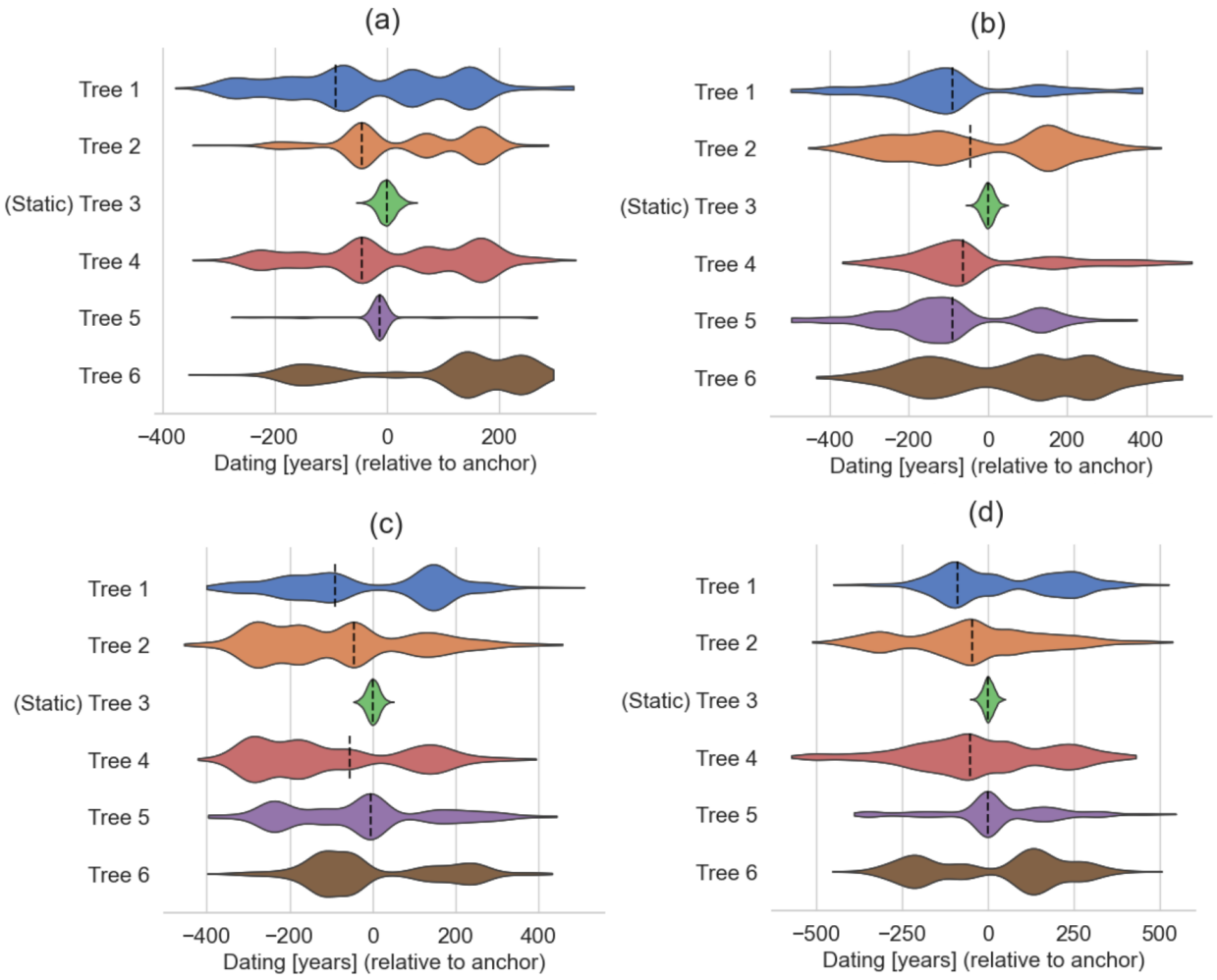

Figure 11: Showing the marginal distributions of test series datings using the 5-series datasets including the 'odd' tree. This last tree does *not* have a dashed black line as it was designed not to 'fit in' with the others.

When considering the results from the 10-series workflow runs, the uncertainty of the odd series seemed rather constant, except in the 4th dataset (fig. 12d) where it seemed comparatively stable. Interestingly, in one dataset the stability of the remaining non-static

series seemed to have been improved from the run before the odd series was introduced. This is apparent when comparing figure 9b with 12b, where the remaining series display much more narrow peaks than without the odd series. This effect was not clearly visible across any other datasets besides this one, and a more common result was no apparent change in uncertainty or a worsening effect. In the remaining three datasets (shown in fig. 12a, c and d) the odd series never seemed to display a low uncertainty, but did show uncertainty comparable with some of the remaining series.

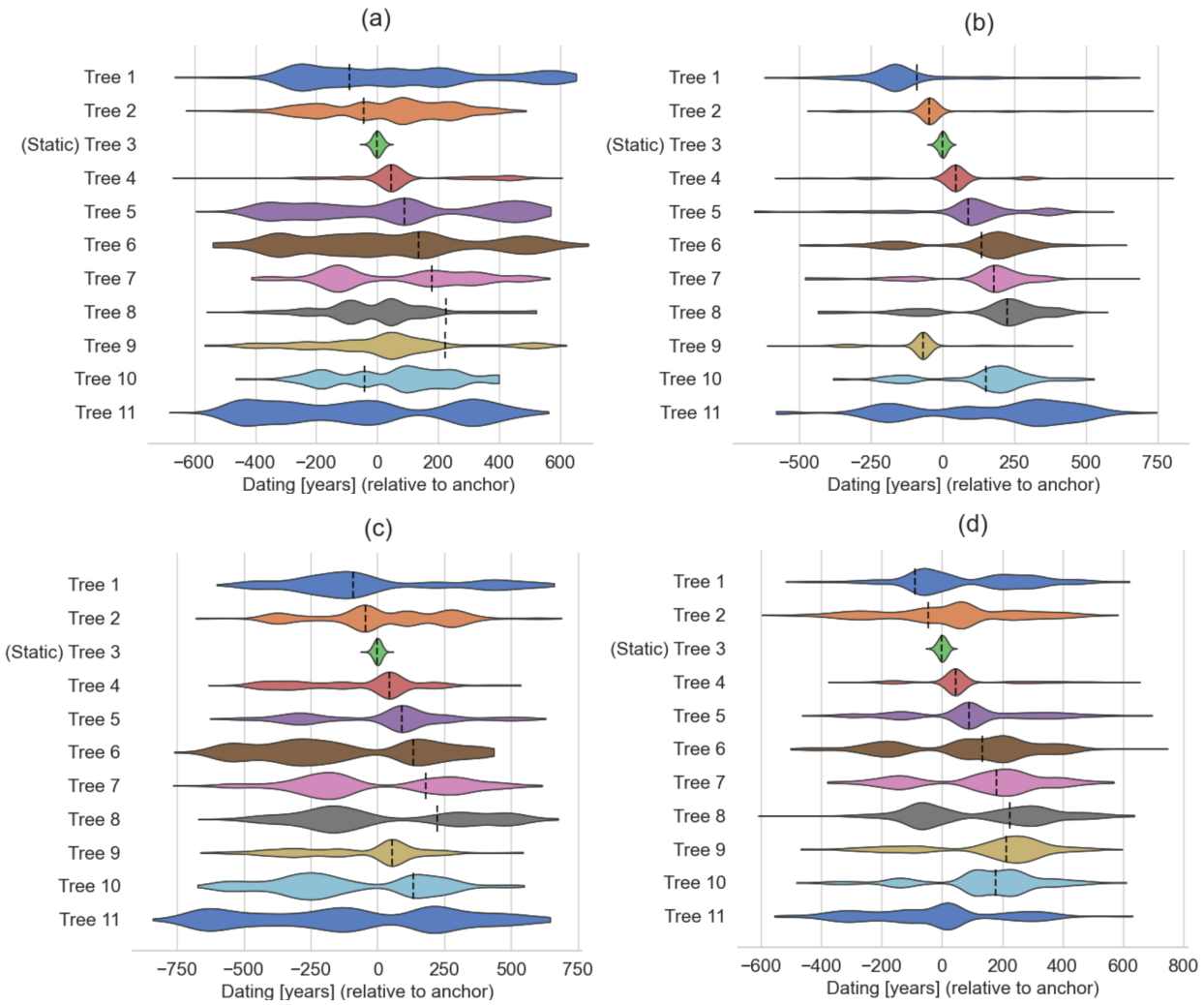

Figure 12: Showing similar results as figure 11 using the 10 series data sets.

Lastly, considering the final four generated test datasets, introducing an odd series into the 15-series datasets did not seem to have a major effect on the uncertainty of the non-static series. Perhaps least so was the 1st dataset (see fig. 10a and 13a) where the difference was most apparent. The odd series never displayed a high stability in the form of one narrow mode, as tree 4 and, to some extend, tree 15 did, but it could very well be confused for a series which 'fit in'.

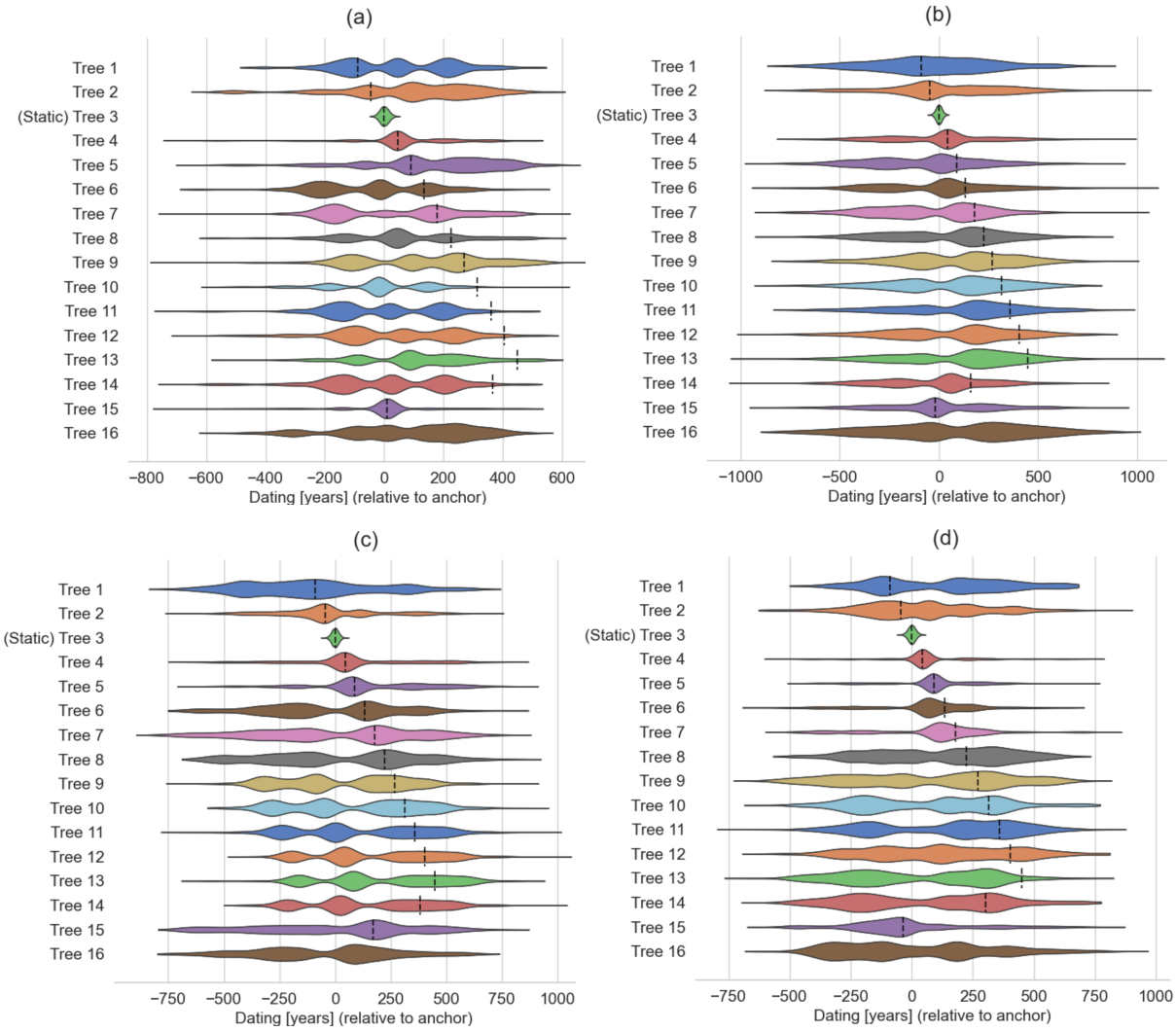

Figure 13: Showing similar results as figure 11 and 12 using the 15 series data sets.

3.2 Real data

Using the workflow to sample real data yielded the results shown in the figures hereunder. As these real datasets contained 9 and 10 TRW series respectively, they were sampled 7000 times and a burn-in period of 4000 samples was applied to all three datasets.

The workflow initially sampled 10 living trees which all had true datings close to each other, some having the same dating. As it can be seen in figure 14, all but one tree showed relatively low uncertainty in their datings with only one dominant mode which contained the true dating. The only exception was tree 4, which displayed a significantly higher uncertainty than the other trees. Additionally, they were all observed to have additional modes, some on one side and some on both sides of the true date, which seemed to somewhat coincide with each other. Based on the modes of the marginals, trees 2, 8 and 10 were correctly dated, trees 5, 6 and 7 were dated incorrectly by 3 or fewer years late, trees 3 and 9 were dated incorrectly by 6 and 5 years late respectively, and tree 4 was dated 96 years too early.

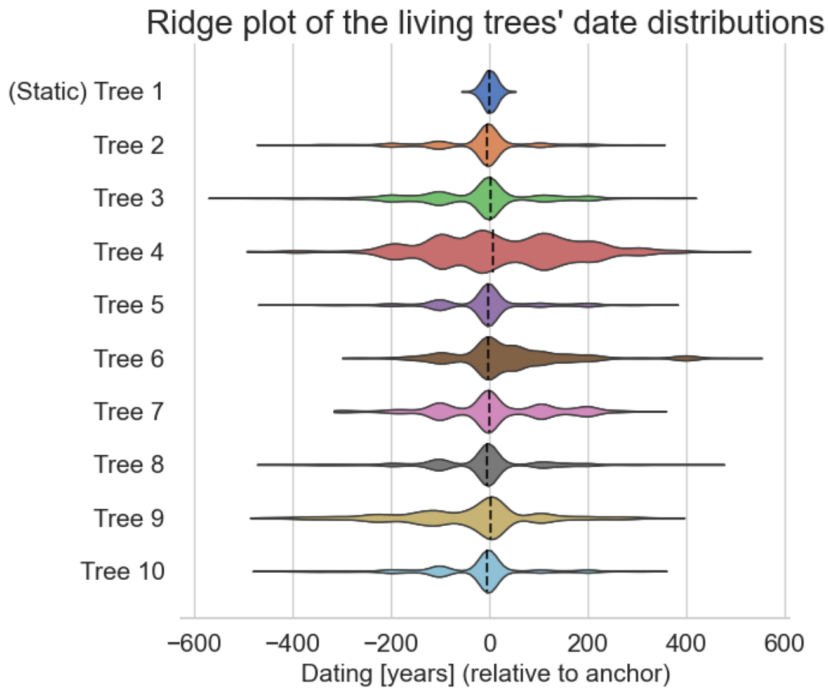

Figure 14: Showing the resulting ridge plot from sampling the living trees. True dates are shown by the dashed lines.

Visually, one might be able to make out that not only did tree number 4 prove to be the *most* uncertain series, but tree number 2 seemed to be the *least* uncertain tree (of course disregarding the static tree). They are visualised in figure 15 along with the 'sub-chronology' constructed from the remaining TRW series.

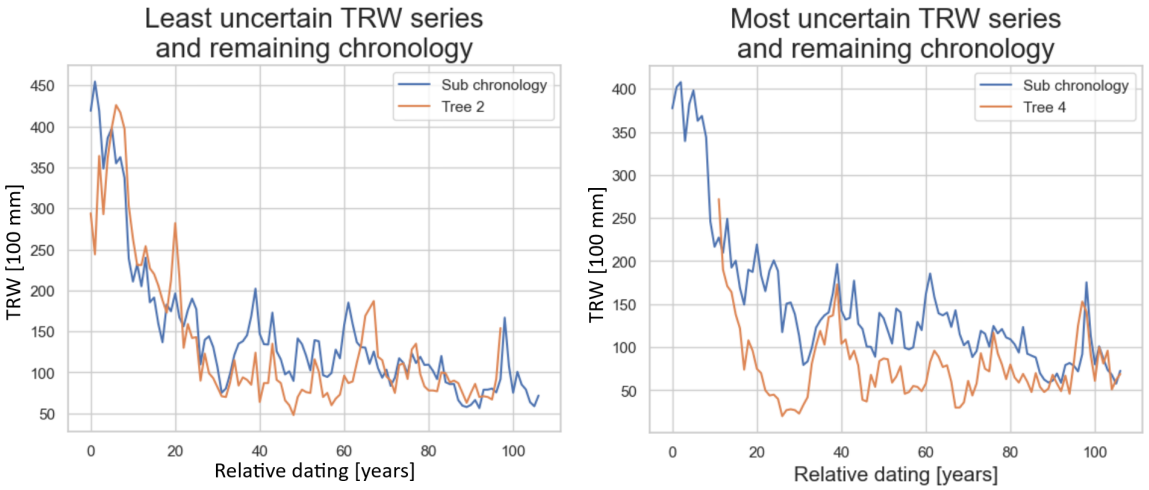

Figure 15: Showing the TRW series with *lowest* (left, tree 2) and *highest* (right, tree 4) uncertainty along with the 'sub-chronology' made from the remaining TRW series at their respective true datings. Note that the TRW series shown in orange are placed at their respective true datings relative to the *beginning* of the sub-chronology.

Moving on to the trees which had true datings potentially more spread out than the living trees. This data seemed to produce a ridge plot with modes distributed on an interval slightly wider than in the previous run, which was somewhat expected. As it can be seen in figure 16, three series in the 9-trees dataset seemed to produce multimodal distributions while the remaining five non-static trees remained very stable. Do also note how the mode of tree 6 marked with a dashed line seems to be situated in a non-dominant (2nd biggest) mode. Both the modes (most often occurring dates) and the plots were calculated and compiled using the data *excluding* the burn-in period, so this phenomenon was likely a slight visualisation issue in the package used to produce the plots.

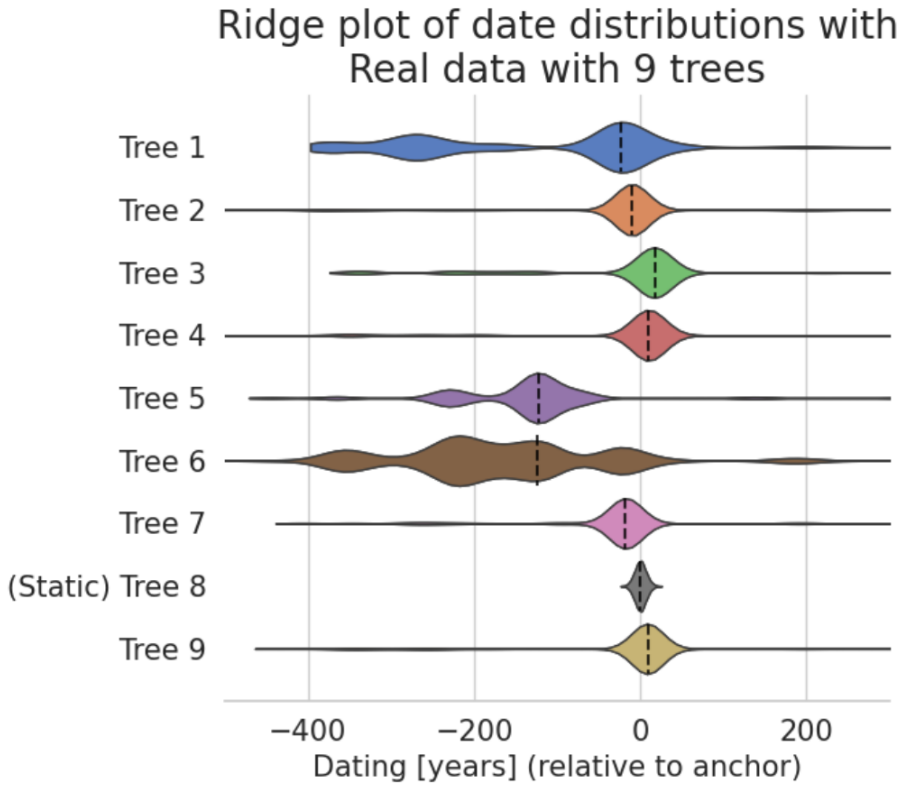

Figure 16: Showing the resulting ridge plot from sampling the dataset containing 9 real TRW series. The dashed lines here show the *modes* and not the *true datings*.

Having sampled the 9-series dataset and extracted the dates which occurred the most often, the subsequent chronology was constructed and compared to the chronology compiled from the same data using the current human-aided methods. In figure 17 the 9 TRW series, which date distributions and modes are shown in figure 16, are shown at their respective datings found at their respective mode.

In figure 18 the final 'modal' chronology produced using the datings suggested by the workflow can be seen. In the same plot the chronology based on the same data, but produced by a scientist with the current method of constructing chronologies, can be seen as well for comparison. In figure 18, the human made chronology has been shifted such that they overlap in the most 'fitting' way for better comparison.

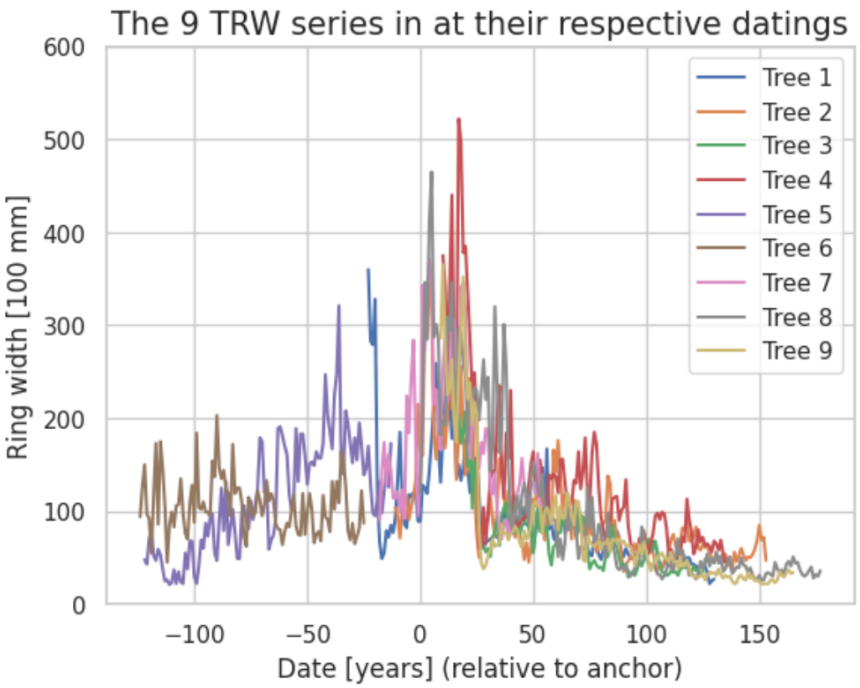

Figure 17: Showing the 9 real TRW series at their respective datings found at the modes of their respective distributions (shown in fig. 16).

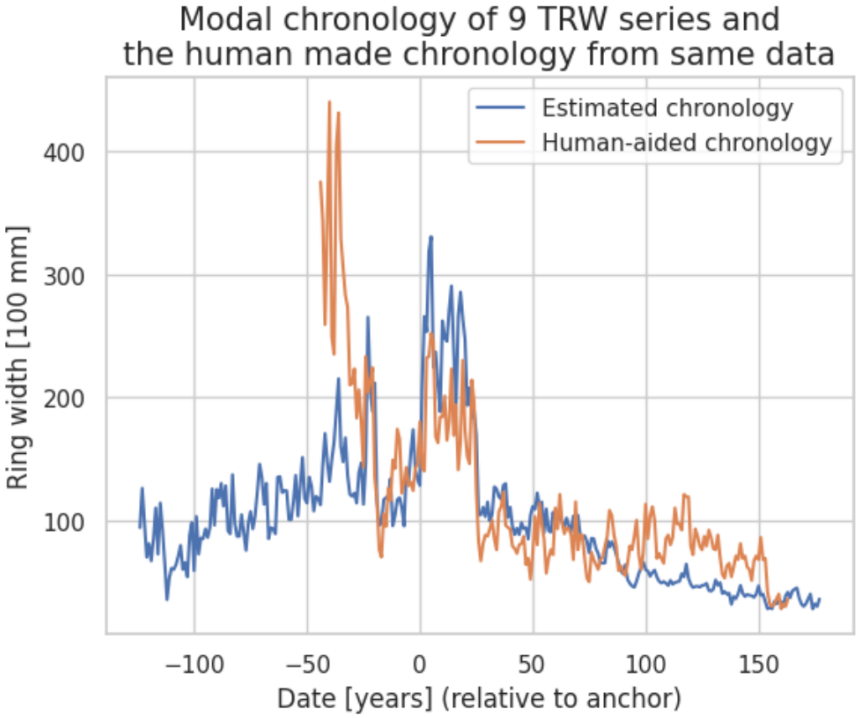

Figure 18: Showing the chronology derived from the results from the workflow as well as the chronology derived by current human-aided means from the 9-series data. Note that for the purpose of comparison they have been shifted such that they overlap where they best fit together.

4 Discussion

4.1 Generated data

Having applied the workflow to various types of data, both artificially generated and real TRW data, the results were somewhat mixed. Considering the results from the generated data, which in theory should display little uncertainty, acted in a somewhat unexpected manner. The fact that they were generated such that they should fit together with high correlation meant that a relatively high stability across all series was expected. In the cases of the 5-series test datasets, the uncertainty of some series was furthermore higher than expected considering the relatively low number of series, although the true datings were captured in the most dominant modes most of the time. The tendency of some series displaying relatively high uncertainty was observed across all the dataset sizes, and seemed to be amplified by the number of series involved. This was expected as more combinations of datings were possible, and so more local optima were expected to be present. A pattern did however arise between some series, showing multiple modes in their marginal distributions which hinted towards internal correlations between the given subset of series. An example hereof is shown in figure 10a, where trees 4, 5, 6 and possibly 7 have very similar multimodal marginal distributions. This suggested that these series were correlated and hinted at which internal lag (that is, their dating relative to each other) their correlation was highest. This can furthermore be seen in figures 9a (trees 2 and 10, and trees 8 and 9), 9b (trees 5 and 8) 10c (trees 9-14) to mention a few examples. It was thereby sometimes possible to make out which series shared high correlation with each other. It was hence possible to gain insight in which series could be matched, even if their distributions did not obviously capture their respective true datings.

Part of the uncertainty issue seen in even the smallest generated datasets can of course originate from the test chronologies having too constant variation on some intervals, where an underlying pattern is obscured by normally distributed noise. This is however possible to encounter in real data as well, and so it was thought to be indications of series having relatively high noise and low underlying signal resulting in high uncertainty. As figure 2 shows, some intervals of the generated series do contain more of what appears to be random noise than clear signal (e.g. comparing years 200-400 or 800-1000 with a relatively clear signal, to years 600-800 with slightly less clear signal). This clarity/unclarity of a signal is however not a problem, as the generated underlying chronologies *are* the signal and were made to simulate real underlying chronologies. This means, that if there are periods where the signal is rather 'flat' and easily obscure-able by noise, TRW series from this period are naturally prone to be more uncertain.

Another possible reason why the uncertainty in marginal distributions of the generated data is higher than expected might lie in the lack of convergence. As it can be seen in figure 4, only one of the three selected dates (4d) seemed to have converged to a reasonable degree. This can also be deduced from figure 7a as tree 5 seemed to be the most stable non-static tree. This might indicate that the burn-in period was not over yet when the sampling ended. Having analysed the evolution of the trees' datings, as shown in figure 4b-d, using a large number of samples could have helped gain a better insight into the convergence time of the different series. However, due to time constraints, this additional analysis was not possible to

carry out. Considering the graphs in figures 4, 5 and 6 the evolution of the LOO correlations apparently reached a limit very soon after commencing the sampling, and the selected test series' datings seemed to either converge right away (see fig. 4c and 6c) or never converge to a single apparent date. Some series appeared to simply be 'easier' for the workflow to cross-date while others apparently were more uncertain, which was in line with the goal of this project.

Introducing a series into the dataset which did not fit in with the other series, it was in most cases rather difficult for the workflow to place it at one specific dating, which was expected. However, an effect this 'rouge' series seemed to have was introducing uncertainty to the other non-static series. This was also to some degree expected, as the odd series would sometimes be included in the LOO correlations of the remaining series, and sometimes not. This would then make it tougher for the workflow to settle on datings (i.e., converge) and thereby return high uncertainty for some series. It is however worth mentioning, that even when an odd series was introduced, one could in some cases still make out which series seemed to be correlated with each other despite high uncertainty. Take figure 11a as an example, where trees 1, 2 and 4 have very similar marginal distributions, and although they seem to have high uncertainty, their shape can be distinguished from that of the odd series. Although this was the case in some examples, other datasets did not yield equally distinguishable distributions. Figure 13b is a good example hereof, where the odd series (tree 16) looks like it shared correlation with trees 7 through 13. In the case of the 10-series datasets 1 to 3 the odd series very closely resembles what it was expected to look like, having a long distribution with no clearly defined modes. It would however not be easy to distinguish odd series from the remaining ones in the cases of figure 12a, c and d, as there is little resemblance between the remaining series' distributions. In datasets 1 and 3 (figure 12a and c) the odd series seemed to have a worsening effect on the distributions' stability of the non-odd series, whilst in the case of dataset 4 (figure 12d) seemingly being able to 'blend in' with the remaining series. One might notice how in figure 12d the odd series (tree 11) resembles tree 2, suggesting that these series are intercorrelated. The example of dataset 2 (figure 12b) is especially intriguing as the odd series seems to have *stabilised* the remaining series, when comparing figures 9b and 12b. This might be an oddity, yet it is still fascinating how these results end up this different. It might be the result of a 'lucky' initialisation of dates, as everything else was kept the same, including sample number and burn-in period. Another possible reason as to why the non-odd series seemed more unstable when the odd series was introduced, is that, as the number of series had increased, the number of samples until convergence (considering an 11-series dataset *without* odd data) should increase as well.

Across the majority of the workflow runs, both including and excluding the odd series, the series' true datings' proximity to the anchor series appeared to have an impact of their stability. If a series' true dating was nearest or second nearest, it seemed to have a more stable marginal distribution compared to the other series. Why this appeared to be the case is unknown, but as the workflow seemed to prefer shorter chronologies over longer ones would be in the favour of series close to the anchor series.

4.2 Real data

Applying the workflow to the type of data it was intended to work on - real TRW data series - yielded varying results. Feeding the workflow data which were 'stacked' seemed to work as intended, cross-dating 9 of the 10 series correctly within 6 years, and 7 within just 3 years. It was furthermore determined, that tree number 4 was the most uncertain tree which did seem to be likely when comparing tree 2 and tree 4 in figure 15. Tree 4 seems to have a period from its initial year to around year 30 where the steep down-slope is rather smooth compared to the same part in the sub-chronology. This was just a visual inspection of course, but it did suggest that the workflow's output is reliable.

In the case of the data comprising of TRW series with potentially more scattered datings (9-series dataset) most of the series displayed rather high stability, only three out of the eight non-static trees showing more than one mode. Furthermore, only tree 6 seemed to have uncertainty of noteworthy magnitude. However, when comparing the resulting chronology to the human made equivalent, it was clear that tree 6 was very likely not the only tree 'out of place'. Considering the length of the workflow's chronology and the alignment of the peaks and the valley between relative years -50 and 0 (fig. 18), tree 5 (fig. 17) was very likely also dated too early. Aside from this, it is not difficult to see where the two chronologies share similarities, suggesting that the workflow likely was on the right path. It is however apparent from this example that the workflow needs improvement in one or more ways to be able to produce chronologies better resembling the human made ones. This of course does not go without mentioning, that removing or moving just one TRW series to a different dating can influence the final chronology quite substantially. This is somehow apparent from comparing the sub-chronologies in figures 15a and b (blue graphs), noticing the difference in scale along the y-axis and how some peaks and valleys differ. This possible misplacement of trees 5 and 6 in in 9-series dataset might therefore be the main reason why the two curves in figure 18 also look so different after year 50.

4.3 Workflow adjustments and improvements

Throughout the process of developing this workflow, several ideas and principles were attempted and discarded for various reasons. Conversely, the workflow could be improved upon in different ways with some suggestions being presented in the following part. Some suggestions were not implemented due to time constraints and some were not fully conceptualised, and they are therefore left for a potentially improved future version of the workflow.

When sampling the nine proposed dates for the individual trees the initial percentiles were the 1st, 5th, 10th, 90th, 95th and 99th, which were later changed in order to be able to select closer to d_k^* as well as from the tails further away from d_k^*. The wildcard positions or 'scouts' from the bee colony analogy were originally set to 'search' the interval of $[d_k^* - L/2, d_k^* + L/2]$ but was increased to search on an interval of L on each side of d_k^*. This was done in order to always include all possible positions where the series $\bar{y}_k$ could overlap with the remaining chronology. This would not necessarily be the case if the smaller interval were used. These scouts could be further improved/optimised by restricting their intervals to the relevant search space only, such that they would not return suggestions which would result in zero overlapping years.

The time constraints previously mentioned were primarily caused by a combination of workflow inefficiency and a large amount of data to be sampled. This is another challenge which was attempted solved in various ways, mostly by simply optimising the code by vectorisation. One specific measure which was considered and attempted would have not only increased the sampling efficiency, but also based the sampling of date vectors on the correlation between each pair of TRW series. The method would have sampled every $\bar{d}_k$ at once from a multivariate normal where $\bar{\mu}$ and $\mathbf{\Sigma}$ would be updated. Here $\bar{d}_{k-1}$ would be used as the mean vector $\bar{\mu}$, and instead of $\mathbf{\Sigma}$ a correlation matrix potentially consisting of the cross-correlation between each pair of series given their internal lag (or difference in datings) would be used. This could have potentially allowed for sampling of dates where series which had high correlation with each other would 'follow' each other and not potentially be split up. This method was given an attempt but ultimately discarded due to time constraints and as it did not provide better results than the method already developed and described in the methods section.

The definition of s_k which would determine the breadth of the foragers' search space was also changed and adapted slightly. The final expression was settled on as it would provide a reasonably wide search space around d_k^* scaled by the parameter L. It is of course debatable whether s_k should be scaled by L as it ultimately restricts the maximum *and* minimum size of the search space around any given d_k^*. The function for s_k could be assigned to be just univariate and give the same variability and magnitude of standard deviation to any series regardless of dataset size. It might however in this case be difficult for two matching series which, have been initiated far apart, to 'meet' and match if they are part of a large dataset.

Some additional measures were employed in the workflow, such as randomising the dates at convergence, removing gaps and selecting an anchor series as the other series' point of reference. These were all implemented to solve problems which arose as a result of how the sampling went on, but two were considered temporary 'band-aid'-like solutions, as they introduce some issues themselves.

The first one introduced was attempting to escape local optima by detecting convergence in $\langle r \rangle$ and then create some artificial 'jitter' by selecting the dating vector not based on LOO correlation but randomly among the 9 proposals. The issue at hand was early convergence and originated from an earlier version of the workflow, but upon working on and gradually adjusting the method the issue rarely occurred. It was kept in place to avoid this issue if it were to occur as it would not influence the sampling if it did not happen. This does however introduce two new matters; one could use probability to choose between proposed dates instead of selecting between these randomly, and one would prevent convergence towards the true datings if one actively escapes optima. **Firstly**, one could, as in the ABC method, choose between proposed 'food sources with a *probability* determined by their quality, and not simply choose the one with the *highest* quality. This could either be done *every* time a tree's dating is chosen or whenever the average correlation seems to have converged. The first option makes the check for convergence redundant and simplifies the method slightly, while the second option would make the workflow faster as the acceptance probabilities of the 9 dates would not be calculated every sample. **Secondly**, when actively escaping optima when converged one could see it as having sampled one optimum from the population of optima, whereafter another (potentially) different optimum would be explored. If the correlation

would converge often, then with enough samples the dates would end up converging towards their global optimum more often than not. This would result in similar convergence of the dates to what is shown in figures 4d and 6d, where the dates find back to their optimal date if they happen to stray away. Additionally, one might think that implementing this would 'jitter' when converged would also somehow make removing the burn-in period redundant. While the possibility of getting a completely new arrangement of the series, like when the dating vector is initiated, is present, it is rather low being $(\frac{2}{9})^K$ or approximately 0.05% using 5 series. Furthermore, if one were to use the probabilistic method of determining which dating to select the probability of getting a completely 'random' arrangement equivalent to the initiation would likely be even lower. That is, the arrangement in d_k^* selected in the case $\langle r \rangle$ has converged is very unlikely completely random, but is most likely close to being a slightly different arrangement of d_{k-1}^*. What this means for the burn-in period and how this could be changed will be discussed once the next additional measure has been discussed.

Removing gaps was perhaps one of the most controversial measures taken when developing this workflow. This is because it changes the datings of the TRW series by moving them back/forwards in time *without* having sampled and tested these new dates. Furthermore, suppose a series has been stable at a relative date of e.g. 700 whilst being separated from another group 100 years into the past. If the series connecting the group and the series at $d = 700$ suddenly moves to $d = 50$, the series at $d = 700$ is suddenly assigned a new date at $d = 600$. This is problematic as it in this situation creates multiple modes in an otherwise stable series. Despite potentially causing these issues, not having gap removal to some extent also proved to be troublesome. If gaps were not removed regularly individual series or pairs which fit well together would start wandering off in either a positive or negative direction. This resulted in a chronology with multiple gaps, even in generated datasets where no gaps were supposed to exist. This was considered a bigger issue as removing gaps returned more accurate results than leaving them in place. An alternative to this constant gap removal could be only doing it in the beginning of the sampling e.g. during the first 10-20% of the samples and then allow the datings to evolve according to their LOO correlation. This would further allow series which do not fit in with the others to wander off to either side and not disrupt the sampling of the other series. This 10-20% initial part of the samples could thereby function as the burn-in period, but appropriate analysis would have to confirm the validity of this fraction. As it was seen in the results from the 5-series datasets removing the proposed burn-in period happened to also remove relevant data (comparing figures 7b and 8) suggesting that the burn-in period was perhaps too large. Seeing the impact on the results of one 5-series dataset does beg the question of whether as much relevant data was removed from any other samplings and thus worsened the quality of those results. This effect of the workflow first finding the true datings and then seemingly *diverging* towards other datings might of course be counteracted by implementing some of the suggestions discussed here.

The final addition was anchoring the longest series such that it would have a constant dating and act as an internal reference point for the other series. This was done as having no internal reference could in theory allow the whole chronology to move e.g. 20 years into the future, adding 20 years to all datings, without any *internal* changes occurring. This is why the chronology needed to be anchored to a single point in time. The choice of it being the longest series was an arbitrary choice which could most likely be improved upon. Suppose that the anchor series happens to be an odd series which does not fit in with any

of the remaining data, all the remaining series would then likely display a high uncertainty compared with the anchor series. Although the remaining series might display very similar uncertainty patterns in their marginal distributions suggesting high correlation, it might be wiser to choose the anchor series more carefully. Nevertheless, the uncertainty of the anchor series is not directly assessed and is only indirectly visible through the stability (or instability) of the other series. One proposed solution could be to calculate the cross-correlation matrix between every pair of series, and the series with the highest average cross-correlation could perhaps more 'safely' be anchored as the reference point. This would furthermore decrease the effect of essentially leaving the uncertainty of the anchor series unaddressed, as this series is likely the least uncertain one if its average cross correlation with any other series is the highest.

4.4 Other considerations

When producing the test datasets, a method known as AutoRegressive Moving Average (ARMA) was initially used to generate the underlying test chronologies. It was however found, that chronologies generated using ARMA contained too much noise for the workflow to properly cross-date the generated test series. That is, the uncertainty of these series were much higher than expected due to the lack of a clear underlying pattern. This is why the currently applied method was utilised instead, where every point was generated from a normal distribution centred around the previous point. The fact that there now was more of an underlying pattern with some moderate noise seemed to make it easier for the workflow to find the correct dates of the generated test series. It is of course debatable to what extend generated data should be 'easy' for the workflow to work with. The general assumption was that the data had to be easy enough to correctly cross-date, to discover if the workflow was *underperforming*, but noisy enough to challenge the workflow and be able to determine weak points in the code. That is if the workflow failed to correctly cross-date data which could easily be done by a person it would be clear that something fundamental likely had to change. Furthermore, if it would *correctly* cross-date 'easy' data it would be impossible to say whether the workflow was 'perfect' or, more likely, just good enough to work with simple datasets. More importantly, it would be impossible to determine *how* good the workflow's performance was, that is *where* between 'perfect' and 'just good enough' it would be. If on the other hand the dataset would be too hard, the workflow would most likely fail to successfully cross-date the series, and the same problem would arise - is it *almost* perfect, *far from* perfect or somewhere in between? It was therefore determined that the method which produced test chronologies with more visible underlying patterns were most suitable for workflow testing.

A measure which was *not* taken in this study was data normalisation and de-trending. As mentioned in the methods section de-trending and normalising the data before analysis is not uncommon and is part of calculating some analytical statistics. This could have been done in this workflow as well in order to potentially make it easier for it to cross-date real data, but was not done due to the human-aided method being compared to in this project not utilising either transformation method. Another reason why de-trending was not utilised here is that different tree specimens can display different ring growth trends, even within the same species from the same stand due to factors like competition and disturbance (Cook

and Kairiūkštis, 1990). It would therefore be unsafe to remove any constant trend (e.g. an exponential or linear trend) from the data. The simulated data is however sampled directly from a generated underlying chronology with only normally distributed noise added to them. That is, no additional underlying trend was added to the test series, meaning the workflow was tested using data which in principle were de-trended. A piece of advice to users of this workflow would therefore be to consider de-trending the data according to local factors before subjecting them to the workflow. It was furthermore thought that if the data were to be de-trended and/or normalised automatically part of the underlying population signal could be removed as well, which could make it more challenging for the workflow to correctly cross-date the TRW series.

Moving on from potential additional measures and on to some final comments on the workflow process. In the calculation of the standard deviation s_k, negative r_k-values were regarded *more* favourable (resulting in smaller standard deviation) than an r_k-value of 0. This might somehow seem counter-intuitive, as later on when accepting/rejecting samples high positive average correlation values are always considered favourable, but it all came down to the goal of the s_k-function. This function served the purpose of controlling the size search area around d_k^*. A key principle here was, that if the LOO correlation of series k at d_k^* was close to -1, a high LOO correlation should in principle be achievable by shifting that series relatively few years back or forwards in time. This is why negative values of r_k were regarded equally as positive values. This might of course not be optimal, and a more detailed function weighting high correlations higher than negative values of same magnitude might be more beneficial.

The keen reader and experienced user of MCMC sampling might have noticed how rejected samples are *re-sampled* instead of keeping the previous sample and moving on to the next. This was a detail originally overlooked as no major issues arose from using this method (repetitive rejections slowing down the sampling being the main concern). There is however no doubt, that the probability of converging faster would be higher if one were to *keep* the previous sample if no better is found, and then move on to the next sample. This minor mishap was unfortunately realised too late in the process and there was no time for running the samplings once again due to the amount of data which would have to be re-sampled. For the sake of testing the difference, the change was made and the first 5-series dataset was sampled, but the results did not show any major improvements. The most uncertain series (tree 1, 2 and 4) did not have as many modes as in figure 7a, but the accuracy of tree 1 was lower than in the original run (see fig. 19 in the appendix).

One final thing worth mentioning about this workflow is that it has not been tested or trained on real data which *had* gaps. That is, data consisting of TRW series which are supposed to be clustered in two or more groups with at least one 'empty' year separating them. This was not part of the scope in this project as an assumption when developing this workflow was that the data which is fed to the workflow is known to (or at least suspected to) have some overlapping throughout. If there were to be a gap (or multiple) in the data the results would likely show which series happen to be correlated with each other as discussed previously. However, since this has not been tested it cannot be said with certainty.

5 Conclusion

In this project a workflow based on a statistical sampling method to build chronologies was compiled. Although it is yet to be perfected, the results did suggest that this approach can lay the foundation to a new tool to be used prior to chronology building. This workflow was applied to both artificially generated and real tree ring data, and was used to develop probabilistic chronologies as well as estimate the uncertainty of the constituent data. When compiling chronologies of real data, the estimated datings were either compared to the trees' true datings, or the resulting chronology was compared with the respective chronology developed using human-aided methods. It was proven partially successful in estimating the true datings of artificially generated test series, especially when applied to smaller datasets. It was additionally proven partially able to detect series which were more strongly correlated with each other, even when their true datings were not correctly estimated. It was furthermore found that if an odd series, which did not 'fit in' with the remaining series, was introduced that most of the resulting distributions either remained unchanged or displayed a higher degree of uncertainty. Despite this, some series continued to display high correlation with each other, by having strikingly similar marginal distributions. The odd series were in most cases showing high uncertainty in the form of wide irregular marginal distributions. They were however in some cases able to 'blend in' with the remaining series as these also showed similarly wide marginals. Whether this odd series was present or not, it was in the majority of the cases, at least to some extent, apparent that the series with a true dating closest to the anchor were the most stable series.

When feeding the workflow *real* tree ring data it produced results of fairly good quality. In the case where the living trees, the workflow performed rather well rarely estimating incorrectly by more than 5 years from the truth. In the case where data were likely more spread out, the workflow seemed to struggle with cross-dating 2-3 of the series. This was seen when the estimated and human made chronologies were compared. The workflow is therefore, at its current stage, likely *not* suitable for chronology building for the sake of comparison, but *did* show potential in suggesting datings and producing TRW graphs for comparison and inspiration. Furthermore, it showed potential in pointing out which series, if any, have high uncertainty, as well as which series might be more strongly correlated with each other. A piece of advice for the user of this workflow is to appropriately de-trend the data before applying it, to compensate for growth patterns and possible competition captured in the data. Implementing the suggested improvements or not, it has without doubt potential of saving time for the scientist, by automatically providing this insight into the data.

6 Acknowledgments

Först och främst vill jag tacka mina handledare Ullrika Sahlin och Johannes Edvardsson för all hjälp, vägledning och inspiration ni har gett mig.

我还想对我女朋友杨将将说谢谢：谢谢你的支持、你的鼓励，还有你一直在我身边为我加油。你知道我永远都会这样对你。

Jeg vil også sige mange tak til min familie og mine venner i Danmark; min mor Birgitte, min far Lars, min lillesøster Dea, til Martin og til Rune for jeres støtte og opmuntring, og for altid at være der når jeg har brug for jer.

Finally, I would like to thank my friends David, Madou, Ruben, Evelina, Stig, Adam, Xin, Ming-Heng, Honia, Simona, and the entire Abisko-gang for your support and solidarity throughout the project and the entire program.

7 References

Akbari, R., Mohammadi, A., & Ziarati, K. (2010). A novel bee swarm optimization algorithm for numerical function optimization. *Communications in Nonlinear Science and Numerical Simulation*, *15*(10), 3142–3155. https://doi.org/10.1016/j.cnsns.2009.11.003

Baillie, M. G. L., & Pilcher, J. R. (1973). A Simple Crossdating Program for Tree-Ring Research.

Bauer, B. (2018). How tree rings tell time and climate history | NOAA Climate.gov.

Buras, A. (2017). A comment on the expressed population signal. *Dendrochronologia*, *44*, 130–132. https://doi.org/10.1016/j.dendro.2017.03.005

Buras, A., & Wilmking, M. (2015). Correcting the calculation of Gleichläufigkeit. *Dendrochronologia*, *34*, 29–30. https://doi.org/10.1016/j.dendro.2015.03.003

Cook, E., & Kairiūkštis, L. (Eds.). (1990). *Methods of dendrochronology: Applications in the environmental science*. Kluwer Academic Publishers ; International Institute for Applied Systems Analysis.

Eckstein, D., & Bauch, J. (1969). Beitrag zur Rationalisierung eines dendrochronologischen Verfahrens und zur Analyse seiner Aussagesicherheit. *Forstwissenschaftliches Centralblatt*, *88*(1), 230–250. https://doi.org/10.1007/BF02741777

Edvardsson, J., Christensen, K., Jensen, J. O., Linderson, H., & Baittinger, C. (2024). Exploring the potential for a 9000-year tree-ring chronology consisting of subfossil oak material from southern Scandinavia. *Dendrochronologia*, *88*, 126268. https://doi.org/10.1016/j.dendro.2024.126268

Fritts, H. C. (1976). *Tree rings and climate*. Academic Press.

Hollstein, E. (1980). *Mitteleuropäische Eichenchronologie: Trierer dendrochronologische Forschungen zur Archäologie und Kunstgeschichte*. P. von Zabern.

Holmes, R. L. (1983). Computer-Assisted Quality Control in Tree-Ring Dating and Measurement. *Tree-Ring Bulletin*.

Karaboga, N. (2009). A new design method based on artificial bee colony algorithm for digital IIR filters. *Journal of the Franklin Institute*, *346*(4), 328–348. https://doi.org/10.1016/j.jfranklin.2008.11.003

Rinn, F., & Heidelberg. (2011). TSAP-Win Time Series Analysis and Presentation for Dendrochronology and Related Applications - User Reference.
Version 4.64 for Microsoft Windows.

Robert, C. P., & Casella, G. (2010). *Monte Carlo statistical methods* (2. ed., softcover reprint of the hardcover 2. ed. 2004). Springer New York.

Stoller-Conrad, J. (2017). Tree rings provide snapshots of Earth's past climate.

Student. (1908). The Probable Error of a Mean. *Biometrika*, *6*(1), 1. https://doi.org/10.2307/2331554

Wigley, T. M. L., Briffa, K. R., & Jones, P. D. (1984). On the Average Value of Correlated Time Series, with Applications in Dendroclimatology and Hydrometeorology. *Journal of Climate and Applied Meteorology*, *23*(2), 201–213. https://doi.org/10.1175/1520-0450(1984)023<0201:OTAVOC>2.0.CO;2

Wilson, R., Anchukaitis, K., Briffa, K. R., Büntgen, U., Cook, E., D'Arrigo, R., Davi, N., Esper, J., Frank, D., Gunnarson, B., Hegerl, G., Helama, S., Klesse, S., Krusic, P. J., Linderholm, H. W., Myglan, V., Osborn, T. J., Rydval, M., Schneider, L., … Zorita, E. (2016). Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. *Quaternary Science Reviews*, *134*, 1–18. https://doi.org/10.1016/j.quascirev.2015.12.005

8 Appendix

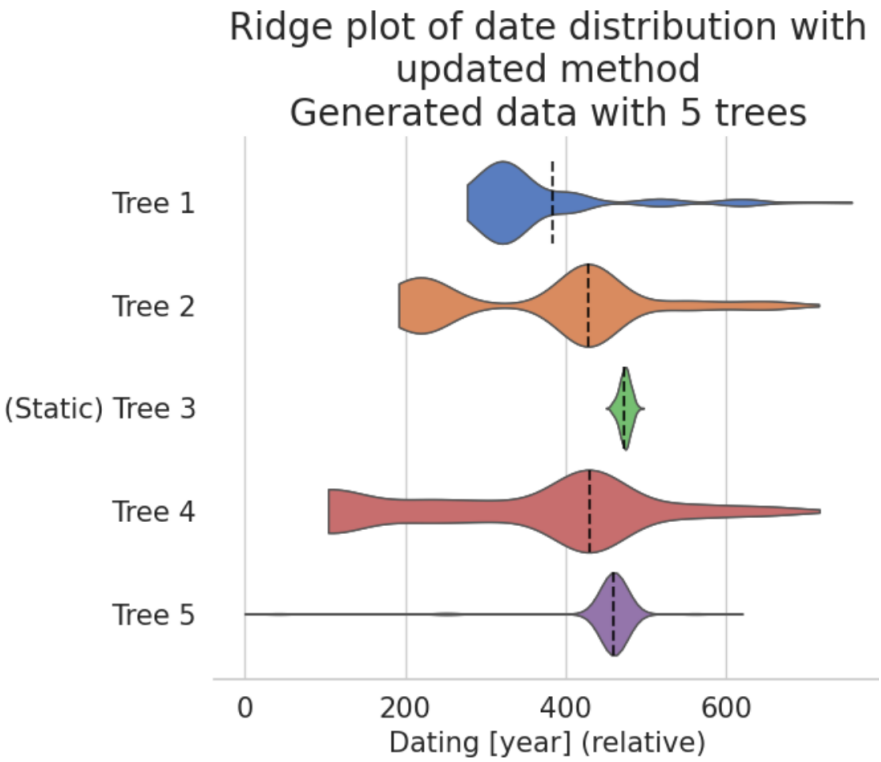

Figure 19: Showing the marginal distributions after running 3500 samples (and removing the burn-in period) on dataset 1 with 5 generated series. The method has been updated such that if a sample is *rejected* the previous sample is 'duplicated' and saved as the 'new' sample, and the sampling continues. The distributions in this figure are to be compared with the ones in figure 7a.

9 Popular scientific summary

When most people see or hear about tree rings, they likely think of the fact that a tree's age can be estimated by counting the number of rings visible on its trunk when cut down. However, when one analyses these rings a bit more carefully and measures their thickness, they can help solve mysteries about past climate and growth conditions. What's particularly interesting about tree rings is that they grow thicker during warm years with favourable conditions, and remain relatively thin during colder and less favourable years. This means that if one had a tree so thick that it would have hundreds or even thousands of rings, one could make estimates of how much colder or warmer the past was compared to today. Even though most trees do not live nearly that long, it is still possible to mimic this phenomenon by using *multiple* trees that lived during different time periods. In practice, scientists carefully measure the width of each ring in multiple tree samples and record these as time series. They then compare these time series in pairs to determine whether and where they 'fit' together, that is, where peaks and valleys align. This process is called *cross-dating*, as it involves determining the trees' relative 'dates', i.e., the years in which each tree's TRW time series begins and ends in relation to others. One can then gradually add more and more trees to this time series, building it further and further back in time, resulting in a time series (known as a *chronology*) equivalent to having a sample from one very old tree.

However, the process of cross-dating many trees' TRW series is time-consuming, even for experienced scientists, and addressing the uncertainty of individual TRW series is also part of this time-consuming process. This presents an opportunity to develop a tool that can assist scientists by suggesting possible datings of each TRW series. It would further assist by pointing out which TRW series appear to fit together, and which, if any, do not. This is precisely the challenge that this project is based on and seeks to address. Principles will be used from different statistical approaches like Artificial Bee Colony and Markov Chain Monte Carlo sampling. The resulting output of such a tool will be in different forms, depending on the user's preferences. These include visualisations of the marginal distributions of individual trees' estimated dates and graphical plots of their TRW time series at those dates. This may aid scientists by providing estimated relative dates between trees, along with an output to compare their own results against. This tool is thereby meant to *assist* the scientists and not *replace* them.

10 Populærvidenskabligt resumé

Når de fleste enten ser eller hører om træers årringe, er deres første tanke sandsynligvis, at man kan vurdere et træs alder ved at tælle dets ringe. Analyserer man ringene mere nøje og måler deres tykkelse, kan de være med til at opklare mysterier om fortidens klima og vækstforhold. Det interessante ved træers ringe er, at de i varmere og mere gunstigt vejr vokser sig tykkere end de gør i koldere vejr. Det vil sige, at hvis man havde et træ så tykt og gammelt at det havde hundrede- eller tusindvis af årringe, ville man kunne sige noget om hvor meget koldere eller varmere fortidens klima var sammenlignet med i dag. Selvom langt de fleste træer ikke lever nær så lang tid, er det stadig muligt at efterligne dette fænomen ved at bruge *flere* træer fra forskellige tidsaldre. I praksis gøres dette ved at nøje udmåle tykkelsen af hver årring fra adskillige træer (årringetykkelse forkortes ofte til TRW efter det engelske udtryk 'Tree Ring Width'), hvorefter hvert træs TRW-data logges i en tidsserie. Disse tidsserier sammenlignes derefter parvist for at undersøge om og hvordan de 'passer sammen', det vil sige hvor mønstre som toppe og dale stemmer overens. Denne proces kaldes for *krydsdatering*, eftersom man fastslår træernes relative 'dateringer', altså hvilket år et træs TRW-tidsserie starter og slutter i forhold til et andet træs. Flere og flere træer kan derved kobles sammen i denne såkaldte *kronologi*, hvilket resulterer i én lang tidsserie, som havde man med et gammelt træ at gøre.

Denne krydsdateringsproces er dog tidskrævende, selv for eksperter, og vurdering af usikkerheder mellem individuelle årringeserier er også en del af denne tidskrævende proces. Dette giver anledning til udviklingen af et værktøj, der kan hjælpe forskere ved at foreslå sandsynlige dateringer af hvert træs TRW-serie. Derudover ville værktøjet samtidig udpege hvilke, om nogle, træer, der ikke 'passer ind' med de andre. Denne opgave er netop, hvad dette projekt forsøger at løse. Løsningen vil gøre brug af statistiske metoder og principper fra Artificial Bee Colony (kunstig bikoloni) og Markov Chain Monte Carlo-prøvetagning. Produktet resulterende fra dette værktøj vil være af forskellige typer, afhængigt af brugerens præferencer. Mulighederne vil omfatte visualisering af træ-datiringenes marginalfordelinger og grafer, som viser TRW-tidsserierne givet deres respektive dateringer. Dette ville potentielt kunne hjælpe forskere, ved automatisk at kunne vurdere træers relative dateringer, og give et sammenligningsgrundlag til deres egne resultater. Dette værktøj er dermed ment som en *hjælp* til forskere, og vil altså langt fra kunne *erstatte* dem.

11 Populärvetenskaplig sammanfattning

När de flesta ser eller hör talas om träds årsringar, är deras första tanke troligen att man uppskatta ett träds ålder genom att räknar ringarna. Om man skulle analysera ringarna mer noggrant och mäter deras tjocklek, kan de bidra till att reda ut mysterier kring tidigare klimat och växtförhållande. Det intressanta med trädringar är, att de växer tjockare under varma och gynnsamma förhållanden än under kalla. Det innebär att om man hade ett träd så tjockt och gammalt att det innehöll hundra- eller tusentals årsringar, skulle man kunna säga något om hur mycket värmare eller kallare klimat var tidigare jämfört med idag. Trotts de allra flesta träd inte lever så länge än, är det dock möjligt att imitera detta fenomen genom att använda *flera* träd från olika tidsperioder. I praktiken görs detta genom att noggrant mäta tjockleken av varje årsring från flera träd (årsringstjocklek förkortas ofta till TRW efter det engelska uttrycket 'Tree Ring Width'), varefter varje träd får sin TRW-data loggad som en tidsserie. Dessa tidsserier jämförs därefter med parvis för att ta reda på om och hur de 'passar ihop', det vill säga där mönster av toppar och dalar överensstämmer. Denna process kallas för *korsdatering*, eftersom man bestämmer trädens relativa 'dateringar', det vill säga vilket år en träds TRW-tidsserie börjar och slutar i förhållande till en annans. Fler och fler träd kan sedan kopplas samman i denna *kronologi*, vilket resulterar i en lång tidsserie, som om man hade haft ett mycket gammalt träd.

Denna korsdateringsprocess är dock tidskrävande, även för experter, och osäkerhetsbedömning mellan individuella årsringsserier är också del av processen. Detta ger upphov till utvecklingen av ett verktyg som kan hjälpa forskare genom att föreslå sannolika dateringar av varje träds TRW-serie. Dessutom skulle verktyget samtidigt identifiera vilka träd, om några, som inte 'passar in' med de andra. Detta är precis vad det här projektet försöker att lösa. Lösningen kommer att använda statistiska metod och principer från Artificial Bee Colony (artificiell bikoloni) och Markov Chain Monte Carlo-sampling. Produkten resulterande från detta verktyg kommer att vara av olika typer, beroende på användarens preferens. Möjligheterna inkluderar visualiseringar av träd-dateringarnas marginalfördelningar samt grafer som visar TRW-tidsserierna utifrån deras respektive dateringar. Detta skulle potentiellt kunna hjälpa forskare genom att automatisk utvärdera träds relative dateringar, och ge en jämförelsegrund för deras egna resultat. Detta verktyg är därmed tänkt som et *hjälpmedel* för forskare och kommer inte på något sätt att *ersätta* dem.