MD-Dating

Age estimation of wood via infrared spectroscopy and random forest modelling

Johannes TINTNER, BOKU University, Institute of Physics and Material Sciences, Vienna, Austria
Bernhard SPANGL, BOKU University, Institute of Statistics, Vienna, Austria
Michael MELCHER, BOKU University, Institute of Statistics, Vienna, Austria

Keywords: *"Random forest" "FTIR spectroscopy" "molecular decay"*

CHNT Reference: J. Tintner, B. Spangl, M. Melcher. 2020. MD-Dating – Age estimation of wood via infrared spectroscopy and random forest modelling. W. Börner and CHNT Organization Committee. Proceedings of the 25th International Conference on Cultural Heritage and New Technologies. DOI:xxxxxxx.

Motivation and Background

The age of artefacts is the most important property for archaeological and historical findings. Several methods have been elaborated to estimate an objects age. For organic matter radiocarbon dating has become the most important method as it can be applied on very different organic materials. The only prerequisite as a matter of principal is the presence of carbon. Anyhow, the method faces several restrictions in detail due to irregular fluctuations in the isotope concentration of the carbon isotope ^{14}C. Especially for wood, dendrochronology became a powerful alternative. Tree ring width follows the same pattern as climatic fluctuations. Given a reference chronology for a certain climatic region and a certain tree species a series of tree ring width from a wooden artefact can be fit into place. The method delivers the year of each tree ring. Especially if the wane is present, the year of trees' death can be detected. Anyhow, also this method faces several restrictions in detail. A sufficiently long tree ring series has to be present, only trunk wood can be used and in several cases the tree ring pattern does not fit into the chronology with a sufficiently high probability.

The new method of MD dating is based upon the molecular decay. This decay follows in many cases a monotonous function. Therefore, this decay function can be calibrated with a set of artefacts of known age and prediction models can be established. Until now, several models have been published for different organic materials: wood (Tintner et al., 2020a,b), paper (Trafela et al., 2007), parchment (Možir et al., 2011), and straw. Also other materials like charcoal have been investigated, but no valid models have been published so far. From an analytical point of view, infrared spectroscopy emerged as a powerful tool. From a statistical point of view, there are several options that can be used for the establishment of such prediction models. Current publications have chosen either Partial Least Squares (PLS)-regression or random forest modelling.

The talk will present the models for wood, especially for Scots pine wood, with a keen focus on the statistical modelling.

Materials and Methods

Wooden artefacts were all taken from archives of dendrochronologists in Austria, Poland, Finland, and Norway. The sample sets of the four species spruce (*Picea abies*), fir (*Abies alba*), larch (*Larix decidua*) and oak (*Quercus* spp.) consisted of Austrian material only (Tintner et al., 2020a). Scots pine (*Pinus sylvestris*) wood covered the spatial range from the Arctic zone of Fennoscandia to the Central European region of Lower Austria (Tintner et al., 2020b). The time range differed from 800 years for fir, about 2000 years for oak, 3000 years for spruce, 3500 years for larch, and 7500 years for Scots pine. Different preservation conditions were included in the sample sets. Wood from living trees, construction wood from historical buildings, waterlogged wood from cold lakes and wood from the very dry Arctic tundra was combined in the respective models. Wood from the prehistoric salt mine in Hallstatt, Upper Austria, and clay deposits near Zürich, Switzerland, had to be excluded, as the measurements displayed slightly different behaviour of the decay function. Fourier Transform Infrared spectroscopy has been applied as analytical tool. Attenuated Total Reflectance was used as the specific measuring technique. For statistical evaluation random forests models were compiled. As the predicted values y_i of the random forest model always underestimated the true year x_i for all species, especially for very old probes, we additionally calibrate the predicted years. For calibration purposes an extended exponential model was applied (details are given in Tintner et al., 2020a).

a b

Fig. 1. a: cross sections of lightning sticks (Abies alba) 3500 years old from the salt mine in Hallstatt, Upper Austria; b: cross section of Scots pine (Pinus sylvestris) – dry storage in the Arctic Norwegian tundra, 800 years old

Results and Discussion

Modelling using random forests resulted in applicable prediction tools. Prediction quality indicated by the root mean squared error (RMSE) differed from 100 years for fir, 118 years for oak, 282 years for spruce, and 388 years for larch. Scots pine was modelled once with a time range of 7500 years leading to a prediction error of 682 years. As central European construction wood ranged only over 800 years, a second model with this time range was calculated leading to a prediction error of 92 years. These values represent the estimations of a single measurement. As commonly more than one measurement can be performed on one wooden artefact, the residual prediction error for the age of the artefact can change considerably. It has to be considered that according to the results within-sample variation cannot be described well by the molecular decay. For living trees and younger samples age predictions and actual ages correspond better. Further effort has to focus on the separation of within-sample variation from aging effects driven by preservation conditions. Restrictions for the prediction are brittle parts with obvious microbial decay and the outermost 5 mm of construction wood. The latter effect probably is reasoned by an increased access of oxygen.

For the interpretation of the results the 30 most important spectral wavenumbers used in the random forest models were assigned to band regions influenced by different molecular wood compound. It became evident that the breakdown of acetyl groups of hemicelluloses plays a crucial role in the molecular decay. For larch and Scots pine also band regions of resin acids are important. Validation procedure of the Scots pine model was performed using two sets of eight randomly selected samples each. These two sample sets were excluded from the data set, the model built up and the test set predicted afterwards. Prediction results were of sufficient quality for most of the samples. Only some predictions were far apart from the reference. Further research will have to focus on the reasons for such poorly predicted cases.

Conclusion and Outlook

Current models for wooden artefacts demonstrate the potential of molecular decay for dating purposes. Random forest models applied on infrared spectroscopic measurements result in proper prediction tools. It has to be stated explicitly that models are valid only for the preservation conditions covered by the sample set. Other conditions might lead to a different speed of the aging processes. Anyhow, in future, we can expect the approach of molecular decay to become a standard procedure in dating archaeological and historical artefacts. The method might become applicable for different organic materials and therefore of a very broad relevance. The easy and cheap measurement will foster new research questions that never have been thought about before. However, there are still several questions open that must be assessed in more detail. A detailed understanding of microbial decay under extreme conditions will help to explain, how molecular decay takes place under different preservation conditions. For a structural understanding of the decay function molecular modelling will provide the theoretical basis. Finally, also additional methods of statistical modelling will be tested.

References

Možir, A., Strlič, M., Trafela, T., Cigić, I., Kolar, J., Deselnicu, V., de Bruin, G. (2011). 'On oxidative degradation of parchment and its non-destructive characterisation and dating', Applied Physics, A 104, pp. 211-217. DOI: 10.1007/s00339-010-6108-z

Tintner, J., Spangl, B., Reiter, F., Smidt, E., Grabner, M. (2020a), 'Infrared spectral characterization of the molecular wood decay in terms of age', Wood Science and Technolgy, 54, pp. 313-327, DOI: 10.1007/s00226-020-01160-x

Tintner, J., Spangl, B., Grabner, M., Helama, S., Timonen, M., Kirchhefer, A. J., Reinig, F., Nievergelt, D., Krąpiec, M., Smidt, E. (2020b), 'MD dating: molecular decay in pinewood as a dating method', Scientific Reports, 10, 11255, DOI: 10.1038/s41598-020-68194-w

Trafela, T., Strlič, M., Kolar, J., Lichtblau, D. A., Anders, M., Mencigar, D. P., Pihlar, B. (2007). ‚Nondestructive Analysis and Dating of Historical Paper Based on IR Spectroscopy and Chemometric Data Evaluation', Analytical Chemistry, 79, pp. 6319-6323, DOI: 10.1021/ac070392t