Curve comparison in wood science: An introductory guide to common cases

Silvana Calvano †
PhD student
Department of Agricultural, Forest and Food sciences, University of Torino, Italy;
Email: silvana.calvano@unito.it

Francesco Negro *†
Associate professor
Department of Agricultural, Forest and Food sciences, University of Torino, Italy;
Email: francesco.negro@unito.it

Emma Oliver
BS student
Department of Wood Science and Engineering, Oregon State University, Corvallis, OR 97331, USA;
Email: oliveemm@oregonstate.edu

Islam Hafez
Assistant professor
Department of Wood Science and Engineering, Oregon State University, Corvallis, OR 97331, USA;
Email: islam.hafez@oregonstate.edu

Alessandro Portaluri
Full professor
Department of Agricultural, Forest and Food sciences, University of Torino, Italy;
Email: alessandro.portaluri@unito.it

(Received 27 March 2026)

Abstract: Empirical curves are essential for describing and analyzing phenomena in wood science. Although the theoretical framework is well established, the complexity of handling curves can hinder their proper application, which is likely to be more common in practice than generally assumed. This article aims to support the effective use and comparison of curves in wood science, and it is mainly intended as an accessible guide for approaching curve comparison, from the basic concepts of dependent and independent variables to the most complex of curve fitting and linearization. After a brief introduction and framework, this study addresses various key concepts, including types of curves, independent and dependent variables, linearization, and number of specimens required. These aspects are discussed through a methodological framework that can be useful in common situations. A section on avoiding common pitfalls is included as well. Proper confidence in handling curves is important not only to avoid errors, but also to enable their effective use, which can result in a more accurate description and a deeper understanding of phenomena.

Keywords: Curve comparison; Curve fitting; Wood science; Wood technology.

Introduction

Curves representing the relationship between independent and dependent variables are common in wood science. They are used to investigate a wide range of topics, from the fundamental behavior of wood to the properties of innovative products (Niemz et al. 2024).

There are several approaches to compare curves, and the statistical methods addressing the most common cases are well established (Hedderich and Sachs 2024). Nonetheless, guidance tailored to experimental conditions and data structures commonly encountered in wood science remains limited. In practice, researchers often adopt methods developed in other disciplines which may not fully reflect the specific characteristics of wood as a heterogeneous material or the testing procedures commonly adopted in wood science.

Curve comparison is not straightforward because curves may differ in shape, scale, rate of change, asymptotic behavior, or variability, and because observations collected along the same curve are often not statistically independent. The appropriate comparison method therefore depends on the structure of the

* Corresponding author
† Society of Wood Science & Technology member

data, the scientific question, and the assumptions of the selected model. An unsuitable method may obscure relevant differences or produce misleading conclusions. For this reason, researchers need a clear framework for selecting the most appropriate analytical approach for the curve-comparison problem under investigation.

To the best of our knowledge, methodological guidance specifically addressing common curve-comparison problems in wood science remains limited and fragmented. This paper aims to contribute to this area by providing an accessible framework for selecting and applying appropriate approaches. First, the theoretical background is presented, followed by examples drawn from published studies in wood science. Referring directly to examples from wood science is intended to facilitate comprehension, as it avoids translating examples from other disciplines. This paper provides an accessible framework for approaching the comparison of families of planar curves commonly encountered in wood science, drawing on established mathematical and statistical methods. This scope covers a broad range of common curve-comparison situations encountered in wood science research, from the basic concepts of dependent and independent variables to the most complex of curve fitting and linearization.

It should be noted that the aim of this work is not to provide a universal recipe, which would hardly be feasible, considering the broadness of the topic. Readers, especially postgraduate and PhD students, are encouraged to deepen their understanding of statistics and to consult appropriate statistical textbooks (see also the Limitations section). However, this article offers operational guidelines for approaching the most common case of curve comparison in wood science and for developing clear and effective manuscripts.

A brief mathematical theoretical framework

Mathematical descriptions and experimental observations in wood science are commonly expressed in the form of relationships between one or more independent variables and a response variable. These relationships between one dependent and one independent variable are typically represented graphically as curves either derived from theoretical formulations or obtained from experimental measurements. As a result, curves constitute a primary aspect of analysis across a wide range of wood science applications. Indeed, mathematical models are used to understand, predict, and optimize properties and behaviors of wood and wood-based materials. They can be categorized into several key areas, with some examples provided in Table 1.

Table 1. List of common mathematical models used in wood science, grouped per thematic areas. Examples include mathematical laws, well-established models, and models proposed in recent literature.

Area	Models	Examples
Physico-mechanical properties	**Mechanical properties**: predict wood mechanical properties based on wood characteristics like grain orientation, density, and moisture content.	Hankinson equation for wood compression strength (Eq. 10.47 in Niemz et al. 2024)
	Fracture mechanics: describe how wood fails under different loading conditions.	Linear Elastic Fracture Mechanics (LEFM) models (Chapter 9.4.7 in Niemz et al. 2024)
Moisture content and hygroscopic behavior	**Moisture diffusion**: predict how water moves through wood in different conditions.	Fick's law and its formulations (Eqs. 7.43-7.47 in Niemz et al. 2024)
	Sorption isotherms: explain how wood absorbs and desorbs moisture at different relative humidities.	Guggenheim-Anderson-de Boer (GAB) model (eq 7.11 in Niemz et al. 2024)
Thermal conductivity and fire reaction	**Heat transfer:** predict how heat moves through wood.	Fourier's law (Eq. 13.9 in Niemz et al. 2024)
	Charring rate: estimate the rate at which wood chars during a fire.	PYCIF heat-mass-pyrolysis model (Pečenko et al. 2023)
Natural durability	**Fungal growth:** predict fungal growth and wood decay rates.	Finnish mold growth model (or VTT model (Hukka and Viitanen 1999)
	Insect attack: assess risks of insect attack using population dynamics and environmental factors.	Modelling the influence of host plants on the distribution of emerald ash borer (Dang et al. 2021)
Wood processing and manufacturing	**Sawing and drying:** optimize cutting patterns and drying schedules.	Page's equation (e.g., Schuler Frantzen et al. 2026)
	Adhesives: analyze how bonding strength varies depending on parameters like wood density, resin content, etc.	Dai's internal bonding model for wood composites (Dai et al. 2008)
Sustainability	**Carbon sequestration:** estimate how much biogenic carbon trees store over time.	Chapman-Richards function (Robinson et al. 2025)
	Forest growth and yield: predict tree growth or timber yield under different climatic and management conditions.	Chapman-Richards function with mixed effects (Eslamdoust et al. 2026)
Computational and machine learning	**Finite element models (FEM):** simulate mechanical behavior and stress distributions in wood products and structures.	FEM analysis of tensile-load resistance of mortise-and-tenon wood joints (Hu et al. 2018)
	Artificial neural networks (ANN): predict wood properties based on large datasets of experimental results.	ANN to predict the quality of thermally modified wood (Bozkurt and Özbay 2026)

From a mathematical viewpoint, some models are described by differential equations, either ordinary or partial and derived from first principles (fundamental axioms or postulates), while others are derived from empirical functions. These models usually rely on specific parameters that serve to tune them. In most cases, their solutions can be expressed as functions of one or two independent variables. Therefore, they can be graphically represented by planar curves or surfaces. These models usually depend on one or more parameters that characterize the corresponding family of functions. Such parameters are generally estimated or calibrated from experimental data, whereas the independent variables constitute the arguments of the model.

The simplest case is that of functions depending on a single variable. Their graphs can be represented in a Cartesian reference frame xOy by plotting the dependent variable on the y-axis and the independent variable on the x-axis. A typical example in wood science is the stress-strain diagram shown in Figure 1. The experimental relation may be represented either as stress as a function of strain or, conversely, as strain as a function of stress. Exchanging the two axes replaces each point (ε, σ) with (σ, ε); geometrically, this corresponds to reflecting the plotted curve across the bisector $y = x$, provided that the same graphical scaling is used. In standard mechanical testing, stress is usually plotted on the y-axis and strain on the x-axis, particularly in displacement- or strain-controlled tests. This convention, however, should not be interpreted as establishing a causal relation: the choice of axes, the control mode of the testing device, and the physical dependence between stress and strain are distinct aspects of the experiment.



Figure 1. Experimental tests are usually set up to have strain as the independent variable and stress as the dependent one.

Types of curves

In practice, curves encountered in wood science differ primarily in the way they are represented and derived. From an analytical perspective, this leads to a limited number of recurring curve types, which differ in terms of continuity, parameterization, and suitability for statistical treatment. Three main types of curves are typically encountered in wood science as described below and in Figure 2.

1. *Infinite series of points based on mathematical models*. An empirical model function can be visualized as an infinite series of points, in which for each value of x, a corresponding value of y can be calculated. Over time, research in wood science has defined well-established mathematical models that describe several behaviors of wood and derived products. Empirical curves are commonly used in wood science, usually when comparing experimental data with expected behaviors.
2. *Finite series of points fitted by a curve* (see also curve fitting section). When data points are collected by experimental measurements, a curve can be fitted to these points to describe the phenomenon under investigation. When the functional

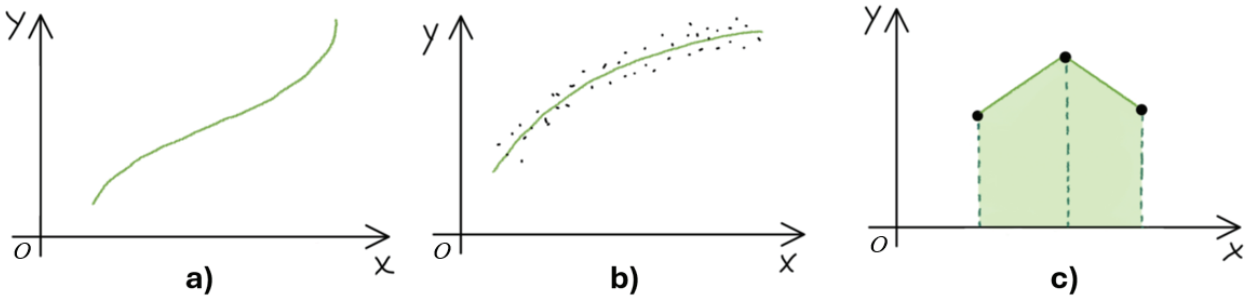


Figure 2. Three curve types typically encountered in wood science: (a) an infinite series of points coming from a mathematical model (e.g., sorption isotherm models depicting wood equilibrium moisture content, EMC, vs. relative humidity); (b) a curve fitting to a finite series of points (e.g., shape change vs. wood EMC, where the points are measured data and the line depicts the fitted curve); and (c) a finite series of points for which a fitting is not proposed, for various possible reasons; the points, which may be shown or not, are connected by broken or smooth lines (e.g., shear resistance in bonding quality test vs. glue spread rate).

dependence between the independent and dependent variables is known, the curve is fitted accordingly using a suitable class of functions, usually depending upon one or several parameters. When the relationship is not known, fitting different types of curves can be attempted. The class of candidate fitting functions should not be selected solely by analogy with similar datasets or by the number of available observations. For any finite dataset, infinitely many functions can be constructed that interpolate all observed points exactly. Therefore, a good numerical fit does not by itself identify an appropriate model. The choice of the fitting function should instead be guided primarily by the theoretical, physical, or biological model underlying the process under investigation, together with the expected qualitative behavior of the response, such as monotonicity, asymptotic limits, convexity, or the presence of saturation. The amount and distribution of the data remain important because they determine how reliably the model parameters can be estimated and how adequately the model assumptions can be checked. However, even a large dataset cannot compensate for the use of an inappropriate model family.

3. *Series of points for which a fitting is not proposed*. In many cases, researchers may choose not to fit a curve to a series of points. For example, curve fitting should be avoided when a mathematical model associated with the phenomenon under investigation is not known. Another case is when the resulting fitting, even if desirable, does not adequately represent the data. In other cases, data may be just too scarce to propose a curve fitting. It is also important to note that polynomial functions, particularly those of third degree or higher, may fit the observed data very well. However, unless the underlying process is known to be polynomial in nature, such a fit should be regarded only as an empirical approximation and not necessarily as the most accurate, interpretable, or physically meaningful representation of the phenomenon.

When curve fitting is not applied, it is quite common to connect points by lines. The resulting broken line can also be smoothed to transform it into a smooth curve, even though an analytic formal expression is not associated with it. Despite the lack of a formal expression, it is possible to calculate parameters related to the curved or broken line. For example, the trapezoidal rule enables the calculation of the area under a broken line: the area under each segment of the line is a trapezoid, and the sum of the areas of the n trapezoids is the total area under the curve.

The distinction between these curve types is not only descriptive, but it directly affects the choice of appropriate statistical methods. The following sections build on this classification to discuss the selection of independent and dependent variables and the extraction of meaningful quantities for curve comparison.

Figure 3 depicts the process to approach curve comparison. Provided that it is not intended as a universal recipe, it provides a reference methodology that applies to several common cases in wood science. In addition, the curve fitting section reports a specific example from wood science to show how the methodological flow chart applies in practice and to give an example of the elements considered in a specific case study. The remaining part of the article follows the conceptual workflow of Figure 3, which also indicates the sections of the text that correspond to the numbers in the figure, as well as pitfalls to be consulted alongside the process (see Table 2 for section numbers key).

Choosing explanatory variables and model factors

In a mathematical model, an independent variable is an argument of the function, while the dependent variable is the corresponding response. In statistical applications, independent variables are also commonly referred to as explanatory variables or predictors. Their values may be set by the experimental protocol, as in the case of treatment temperature, or recorded during the study, as in the case of time or initial moisture content (MC).

This article focuses primarily on curve-comparison problems involving one or two explanatory variables. These may be continuous predictors or categorical factors with two or more levels. For example, wood density may be treated as a continuous predictor to estimate wood swelling (Pulgar and Riesco 2024, see below) or as a categorical factor, e.g., with “low”, “medium”, and “high” levels (Nugroho et al. 2025). More complex experimental designs may also include blocking variables, covariates, repeated-measures factors, or random effects. Although it is possible to design studies with more than two independent variables, models involving several explanatory variables and their interactions may require more complex experimental designs and statistical methods. Since this article is intended as an introductory guide, it focuses mainly on situations involving one or two primary explanatory variables, while acknowledging that additional blocking factors, covariates, repeated-measures structures, and random effects may also be present.

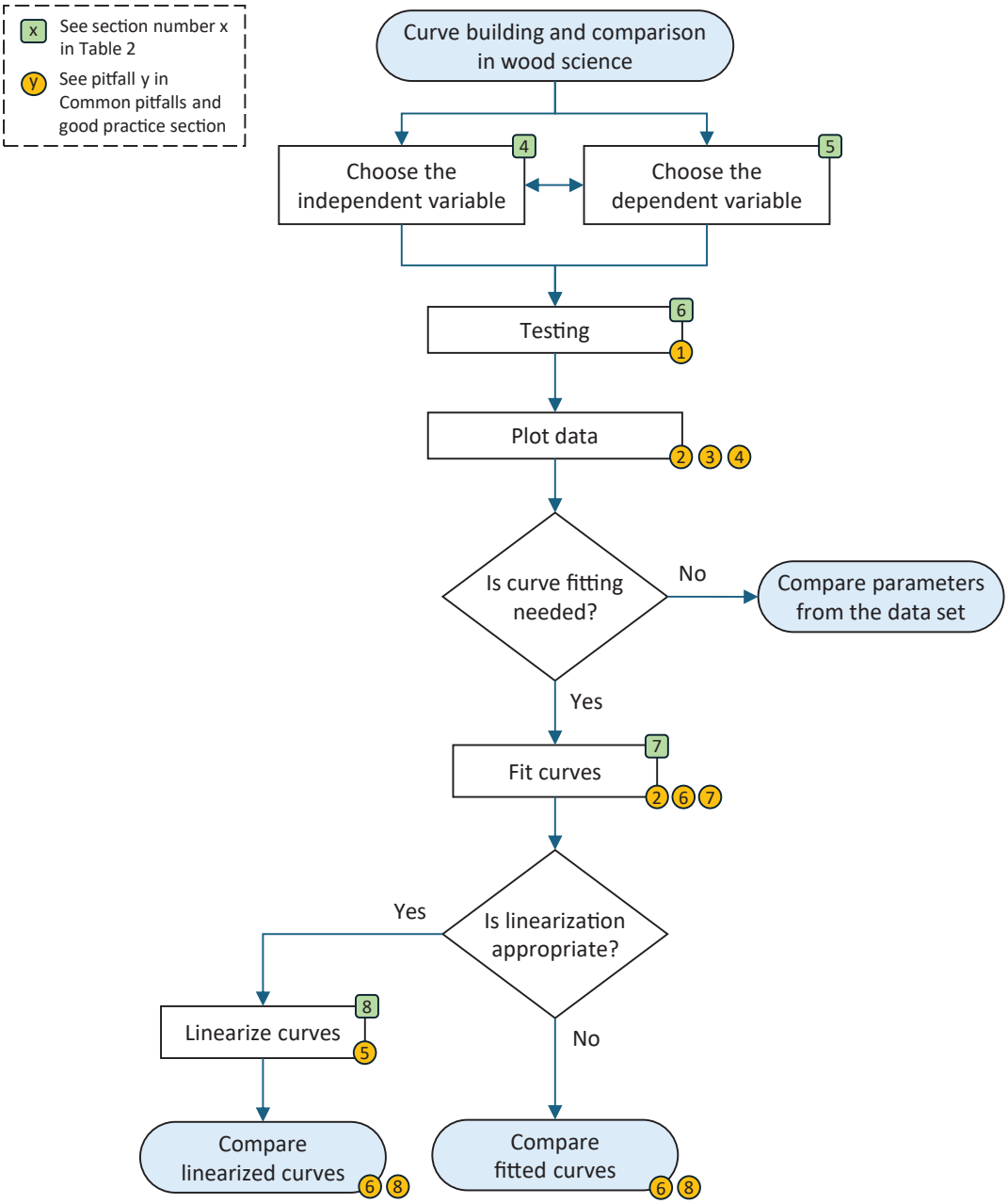


Figure 3. Methodological flow chart to approach the most common cases of curve comparison in wood science (see Table 2 for section numbers key).

Table 2. Key to section numbers shown in green squares in Figure 3 (the numbers in orange circles refer instead to pitfalls in the Common pitfalls and good practice section).

Section title in paper	Corresponding section number in Figure 3
Introduction	1
A brief mathematical theoretical framework	2
Types of curves	3
Choosing explanatory variables and model factors	4
Choosing the dependent variable	5
The number of specimens	6
Curve fitting	7
Linearization	8

The main categories of explanatory variables and model factors commonly encountered in curve comparison are listed below, along with representative examples from wood science.

Continuous predictors are numerical variables that can take values on a continuous scale. They are commonly used in regression models to describe how a response changes as the predictor varies. For example, in wood drying, wood density may be treated as a continuous predictor to estimate swelling and shrinkage values (Pulgar and Riesco 2024).

Categorical factors classify observations into two or more distinct groups or levels. Their effects may be investigated using methods such as ANOVA when the relevant assumptions are satisfied. In wood science, a typical example is the comparison of the mechanical properties of wood from different clones or species (e.g., Calvano et al. 2026a, Nyobe et al. 2024).

Blocking variables are categorical variables that are not the primary focus of the study but may contribute systematic variation to the response. Experimental units are grouped into blocks so that this source of variability can be accounted for separately from the treatment effect. For example, in a randomized complete block design on wood physico-mechanical properties, wood provenance can be included in the statistical model as a blocking factor rather than removed from the analysis (Chen et al. 2026).

Covariates are usually continuous variables that are not the main object of investigation but may influence the response. They are included in the statistical model so that comparisons among the groups of primary interest can be adjusted for their effect. In ANCOVA, for example, a covariate may be used to compare treatment groups at a common or adjusted value of that variable. For example, Abe et al. (2016) investigated the light conducting properties as a method to identify two wood species from archeological artifacts. Wood density was set as the covariate because it affects light conduction, and wood species as the independent variable.

Repeated-measures factors arise when the same experimental units are observed repeatedly at different time points or under different conditions. The resulting observations are generally correlated and therefore cannot be treated as independent replicates. Repeated-measures ANOVA or, more flexibly, mixed-effects models may be used to account for this dependence. For example, Zhang et al. (2025) investigated the decay of paper birch and red pine wood by measuring the mass loss of specimens collected from the same in-field logs (the experimental units) over a period of five years.

Random effects represent sources of variation associated with groups, subjects, or clusters whose observed levels are regarded as a sample from a broader population. They are commonly included in mixed-effects models to account for correlation among observations belonging to the same group and to estimate variability among groups. For example, when comparing wood properties, tree identity may be included as a random effect because wood specimens from the same tree are likely to be more similar to one another than specimens taken from different trees (e.g., Adutwum et al. 2023).

Finally, these categories describe different aspects of a statistical model and are not always mutually exclusive. A variable may be continuous or categorical in mathematical terms, while also having a particular role in the experimental design. For example, time may be treated as a continuous predictor (e.g., to estimate the shape change of hygromorphic wood bilayers, see Table 4), but measurements taken over time on the same specimen can also form a repeated-measures structure (e.g., Zhang et al. 2025 cited above). The appropriate classification therefore depends both on the nature of the variable and on the role it plays in the scientific question and statistical model.

Choosing the dependent variable

Dependent variables are quantities measured to assess changes resulting from the independent variables. They are traditionally thought of as a recorded variable in a dataset. As stated in the mathematical theoretical framework section (above), the y-variable, graphed on the y-axis, is the dependent variable for data with one independent variable. Although correct and often very useful, this is a narrow definition of the concept. Dependent variables, in fact, can include any quantity that adheres to the definition above, including curve parameters.

Curve parameters are unknown coefficients estimated during the fitting procedure. Once estimated separately for each experimental unit, their values may subsequently be treated as response variables when comparing groups. For example, if the same quadratic model is fitted to each specimen, the resulting specimen-specific estimates of its coefficients may be statistically compared across treatments. So, when fitting curves based on the function $y(x) = ax^2$, the a coefficient operates as a parameter that is adjusted to find the best curve. Once finalized, it can be used as a dependent variable to compare different curves based on the same function.

A key question in curve comparison is which dependent variable(s) should be considered. Several curve parameters may be of interest, such as the maximum or the average value,

the slope at a specific point, the area under the curve (AUC), etc. It is therefore necessary to determine which parameter is more useful to investigate. Five main situations can be found in practice in wood science, as outlined in Table 3 and Figure 4.

Finally, another option is to linearize data, which consists in transforming the relationship between variables to make it linear (see the Linearization section, which lists three common cases of data linearization). Linearization is performed to simplify data interpretation, to improve model outputs, or to meet statistical assumptions. Linearized relationships can then be compared by analyzing differences in regression coefficients (slopes) or intercepts. This is typically achieved using ANCOVA, as a general linear model that combines ANOVA and regression.

Table 3. Parameters to be considered when analyzing curves in wood science. The examples listed are not intended to be exhaustive, nor to describe how to conduct data analysis. Instead, they provide a prompt reference for visualizing some practical applications of the theoretical cases.

n.	Parameters being considered	Notes	Examples in wood science
1	One point	Information from the curve is lost. The analysis is reduced to one-dimensional data.	Brinell hardness of densified birch and aspen wood (Fig. 5 in Altgen et al. 2026) Bending strength of glue laminated timber beams (Fig. 7 in Martins et al. 2025).
2	Two or more points	Information from the curve is partially lost. Comparison between points is possible, but the trend between points is not considered.	Mechanical properties of wood plastic composites measured every day over 8-d water immersion (Fig. 10 in Rezakalla 2025) Mass loss values depending on wood moisture content to assess the moisture dynamics of brown rot degradation (Fig. 9 in Belt et al. 2026)
3	A section of the curve	Considering a section of the curve can be chosen by convention, as in standards that prescribe how to calculate the modulus of elasticity (MOE). It can also be chosen because of particular interest, such as, for example, focusing on the behavior of wood in the first 2 h of testing.	Near-infrared spectroscopy to determine MoE in static bending of Mongolian oak wood (Fig. 5 in Liang et al. 2018) Fracture evolution stage of four wood species under compact tension load (Fig. 6 in Hu et al. 2021)
4	The entire curve	The entire curve is considered, whereas single points are not, or are not the focus. This approach is particularly suitable for curves with similar paths.	Deformation of hygromorphic bilayers made from different woods over 24-h conditioning cycles (Fig. 3 in Calvano et al. 2026b) Time-temperature effect on plywood and oriented strand board creep deflection (Fig. 4 in Miyamoto and Sinha 2025)
5	Two or more of the previous approaches	Taking into account both punctual parameters (cases 1 and 2 from this Table) and those related to the curve (cases 3 and 4) provides a broader picture of the phenomenon under study. In some cases, it may be necessary to assess the relationship between the parameters considered.	- Flame spread evolution and final value for treated Japanese cedar wood subjected to a 10-min Steiner tunnel test (Fig. 3 in Jang and Park 2026). - Bending force at proportionality limit and at maximum limit in bending strength test of beech and aspen wood (Figs. 3 and 5 in Svoboda et al. 2017)

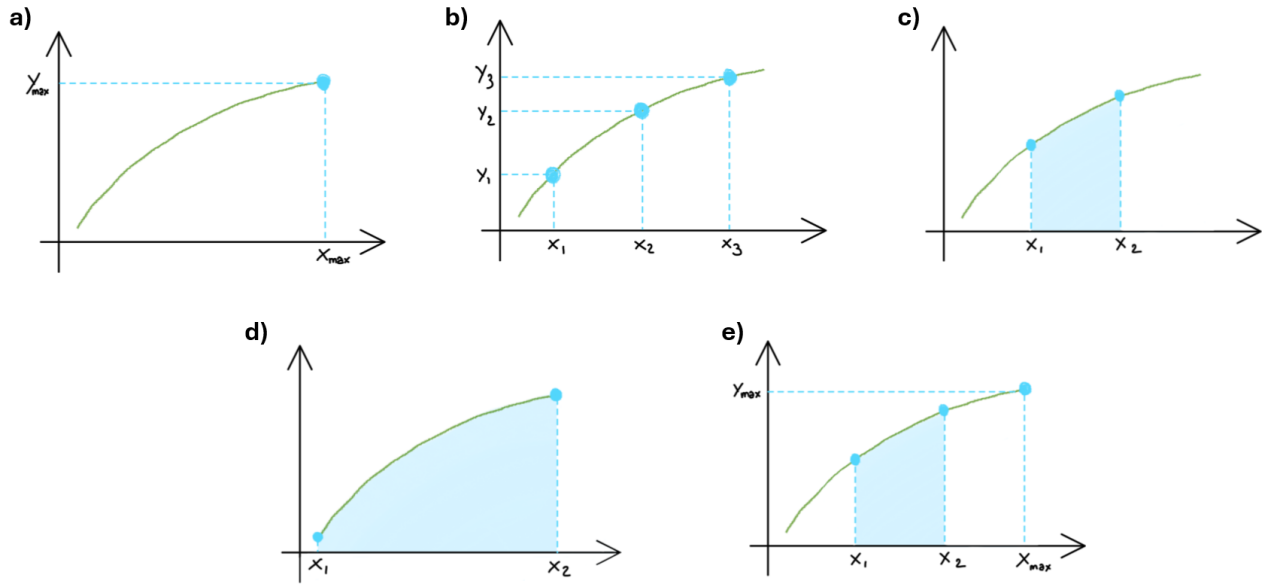


Figure 4. Parameters considered when analyzing curves, according to Table 3: (a) one point (type 1 in Table 3); (b) two or more points (type 2); (c) a section of the curve (type 3); (d) the entire curve (type 4); and (e) two or more of the previous approaches (type 5).

The number of specimens

The number of specimens depends on the specific case under investigation. There are theoretical methods for calculating it depending on the expected variability, e.g., ISO 3129 (2019). In practice, two operational approaches are generally accepted in wood science. First, the number of specimens is set based on other research studies that are taken as a reference. Second, the number is chosen based on the prescriptions of technical standards, like EN 326-1 (1994) for wood-based panels. In both cases, it is not rare to test a higher number of specimens to make the statistical analysis more robust. When these operational criteria are used, the expected variability should also be considered. A higher number of specimens is generally required for highly variable materials such as solid wood, whereas a lower number may sometimes be adequate for more homogeneous engineered wood products. Ideally, however, the number of specimens should be justified by considering the expected effect size, variance, desired statistical power, number of treatment groups, and repeated-measures structure. Importantly, multiple observations collected along the same curve do not increase the number of independent experimental units.

Curve fitting

Preliminary considerations on curve fitting have already been discussed in Types of curves, particularly in the sections "Finite series of points fitted by a curve" and "Series of points for which a fitting is not proposed." Some additional aspects that should be considered when fitting and evaluating a curve are discussed below.

A fundamental step in curve fitting is the analysis of residuals, defined as the differences between the observed values and the corresponding values predicted by the fitted model. Residual analysis helps assess whether the selected model adequately represents the phenomenon under investigation. Ideally, residuals should be randomly distributed around zero and contain no systematic patterns. Trends, curvature, changes in variance, clusters, or unusually large residuals may indicate model misspecification, non-constant variance, dependence among observations, or the presence of influential data points. In such cases, the model assumptions, the selected functional form, or the experimental structure should be reconsidered.

Another important step is the evaluation of goodness-of-fit, which describes how closely the fitted curve represents the observed data. Several complementary measures may be used. The coefficient of determination, (R^2), represents the proportion of variability in the response accounted for by the model and generally ranges from 0 to 1, with values closer to 1 indicating a closer fit. However, (R^2) should be interpreted cautiously, particularly for nonlinear models, and should not be used as the sole criterion for model selection. Adjusted (R^2) penalizes the inclusion of additional predictors and is therefore useful when comparing models of different complexity within an appropriate regression framework. The root mean squared error (RMSE) is the square root of the mean squared residuals and measures the typical magnitude of the prediction errors in the units of the response variable. Lower RMSE values indicate a closer fit, provided that the models are evaluated using the same dataset and response scale.

Confidence intervals for the fitted parameters should also be reported. They quantify the uncertainty associated with the estimated coefficients. Narrower confidence intervals generally indicate greater precision, whereas wider intervals suggest that the available data do not determine the parameter accurately. However, precise parameter estimates do not by themselves demonstrate that the selected model is scientifically appropriate.

Researchers should also consider whether the selected model and its parameters have a clear physical or biological interpretation. Purely empirical functions, such as high-order polynomials, may reproduce the shape of the observed data without explaining the mechanisms responsible for that shape. By contrast, models containing physically interpretable parameters, such as the unit shrinkage coefficient of wood, may support a more meaningful scientific interpretation of the phenomenon. This consideration is especially important when fitted parameters are subsequently compared among experimental groups. The same model should adequately describe all groups before its estimated parameters are compared.

Finally, overfitting, excessive smoothing, and the use of models without physical justification are discussed in the common pitfalls section. These issues should be considered together with residual analysis, goodness-of-fit measures, parameter uncertainty, and the scientific meaning of the selected model.

Table 4 reports an example of a study based on curve fitting, representing one of the most complex scenarios covered in this article. This example complements Table 3 (on parameters to be considered), Figure 3 (on the methodological flowchart), and this section on curve fitting.

Linearization

In the workflow of Figure 3, linearization is an optional step considered after a plausible model has been selected.

Table 4. Example from a recent study (Calvano et al. 2026b) based on the methodological framework shown in Figure 3.

Activity	Example
Choose the experimental design, including selection of the independent and dependent variables and review of similar studies in literature.	The study investigates the shape change of hygromorphic wood bilayers over time. The active layer is made with two wood species (independent variable n. 1), and bilayers are bonded using two adhesives (independent variable n. 2). The shape change over time (dependent variable) is compared using two-way ANOVA (2x2). The experimental design is similar to that of a study taken as a reference.
Conduct the experimental testing.	Four different bilayers are tested. Ten specimens per bilayer are produced. The shape change of each specimen, in mm, is measured at regular intervals over 24 h.
Plot the series of points for each specimen.	The series of points obtained for each specimen are plotted in a Cartesian reference frame.
Perform the curve fitting for each specimen.	Curve fitting is performed according to a function (model) widely accepted in literature. Residuals are randomly scattered around the zero line. The goodness-of-fit is assessed by (R^2). As a result, a curve is associated with each of the 10 specimens per panel.
Is linearization needed?	Linearization is not needed. The fitted curves are compared.
Measure the values of the dependent variable(s) for each curve.	The area under the curve (AUC) is calculated for each specimen. This results in 10 specimen-level AUC values for each bilayer type. These values are retained as the experimental observations used in the statistical analysis; group means are reported only as descriptive summaries.
Perform the statistical analysis.	Boxplots are drafted to have a visual representation, but they are not included in the article. A two-way ANOVA is conducted on the specimen-level AUC values to assess the effects of wood species, adhesive type, and their interaction.
Compare the results of the statistical analysis.	Non-significant and significant differences are discussed.

Linearization is one of several approaches available for analyzing nonlinear relationships and should be considered when the transformation produces an approximately linear relationship and the assumptions of the resulting statistical model are satisfied. Its main advantage is that parameters of a nonlinear model may become the slope and intercept of a straight line, facilitating graphical diagnostics and, when assumptions are met, statistical comparisons among experimental groups. For example, consider a family of empirical quadratic models of the form

$$y(x) = ax^2 + bx \quad [1]$$

where (a) and (b) are parameters estimated from the data and used to tune the model. Both parameters contribute to the shape of the curve: (a) controls its curvature, whereas (b) controls its linear component and, in particular, its slope at ($x = 0$). Therefore, comparing the fitted curves only through (a) would neglect the contribution of (b) and could lead to an incomplete or misleading interpretation. A rigorous comparison should instead consider the parameter vector ((a, b)), together with the possible statistical dependence between the two estimates.

When a simpler representation is desirable, the model can be linearized, for $x \neq 0$, by dividing both sides by x:

$$\frac{y(x)}{x} = ax + b \quad [2]$$

In the transformed coordinates, (a) and (b) become, respectively, the slope and the intercept of a straight line. This representation may facilitate the simultaneous comparison of both parameters using linear-model methods, provided that the transformation is appropriate for the data and that the assumptions on the transformed residuals are verified. In this view, ANCOVA can test differences in slopes through the interaction between the predictor and the group, and differences in intercepts through the group effect.

Linearization should not be applied automatically. The transformation must be defined over the observed range and should yield an approximately linear relationship with independent residuals and reasonably constant variance. Because transformations can alter the error structure and the relative weighting of observations, direct nonlinear regression is generally preferable when assumptions are better satisfied on the original scale, when asymptotic parameters must be estimated from the same data, or when uncertainty is required on the original measurement scale. The following subsections illustrate three common transformations.

Linearization of an exponential model

A standard exponential-type model has a solution of the following form:

$$y(x) = y_0 + C\, e^{-kx} \quad [3]$$

where x is a real number and C at [3] is a positive constant. By setting $Y(x) = \ln(y(x) - y_0)$ and $q = \ln C$, the previous equation reduces to:

$$Y(x) = -k\, x + q \quad [4]$$

which is the Cartesian equation of a line, in the $\widehat{XOZ}$-plane, having slope $-k$ and y-intercept q. Thus, by changing the parameters k and q, a two-parameter family of lines in a semi-logarithmic representation is obtained. Implementing this change of coordinates in the dataset provides a linear dependence between the variable x and the variable Y, as at [4].

This transformation requires $y(x) > y_0$ and treats y_0 as known or fixed independently; for a decreasing exponential curve, $k > 0$. Statistically, the slope estimates $-k$ and the intercept estimates $q = \ln C$. When several treatments or material groups are compared, ANCOVA can test whether their decay rates (slopes) or initial amplitudes (intercepts) differ. This requires independent observations and an adequate residual structure on the transformed scale. Repeated measurements from the same specimen should instead be analyzed with a repeated-measures or mixed-effects model. If y_0 must be estimated from the same observations, direct nonlinear regression is usually preferable.

Linearization of a super-exponential model

A different treatment is deserved for the super-exponential model, starting with

$$y(x) = y_0 + C\, e^{-kx^n} \quad [5]$$

where C, k, and n are positive constants, and $x \geq 0$. The following chain of inequalities holds:

$$y_0 < y(x) \leq y_0 + C \quad [6]$$

By setting $Y(x) = \ln\left(\frac{y(x) - y_0}{C}\right)$ and by considering the previous inequalities, one obtains Y(0) = 0, whereas for x > 0, $Y(x) < 0$. By setting $Z(x) = \ln(-Y(x))$, which is well defined for x > 0 because $Y(x) < 0$; by also setting $X = \ln x$, the previous super-exponential model reduces to the straight-line having Cartesian equation given by:

$$Z(X) = q + n\,X \quad [7]$$

where $q = \ln k$. Thus, the relationship is linear on a log-log scale. Once again, a two-parameter family of straight lines is obtained, where the parameters are the slope and the y-intercept.

The slope is n and the y-intercept is $q = \ln k$, so group differences in curve shape and rate can again be assessed by comparing slopes and intercepts. Particular care is required near $x = 0$ and when $y(x)$ approaches y_0, because the logarithms can amplify measurement error. If y_0 or C is estimated rather than known, its uncertainty propagates into the transformed variables; direct nonlinear regression may therefore be more reliable.

Linearization of a logarithmic-type model

A standard logarithmic-type model has a solution of the following form:

$$y(x) = y_0 - k\ln(x + 1) \quad [8]$$

where $x \geq 0$ and k is any nonzero constant. By setting $Z(x) = y(x) - y_0$ and $X(x) = \ln(x + 1)$, the previous equation reduces to a linear equation in the Cartesian $\widehat{XOZ}$-plane; namely:

$$Z(X) = -k\,X \quad [9]$$

Since $X \geq 0$, the previous equation defines a one-parameter family of half-lines through the origin of the transformed $\widehat{XOZ}$-plane, parametrized by k and lying in the right half-plane.

Here, $-k$ is the slope. If y_0 is known, the regression may be constrained to pass through the origin; if y_0 is estimated from the data, an intercept should be included or the original model should be fitted directly. For comparisons among groups, a group-by-X interaction tests differences in k.

In summary, linearization is useful when it yields an interpretable straight-line representation and an acceptable residual structure. It should be regarded as one possible branch of the workflow in Figure 3, not as a compulsory replacement for nonlinear fitting.

Common pitfalls and good practice

Although statistical methods for curve analysis and comparisons are well established, their practical application may still lead to misleading results if data are processed or interpreted improperly. This is particularly relevant in wood science, where experimental variability and complex material behavior require careful handling of data. While some issues, such as data scarcity and overfitting, were discussed earlier in the Types of curves section, the following points highlight additional pitfalls frequently encountered in practice and outline good practices that help ensure reliable and meaningful curve comparison.

1. Treating repeated points from the same specimen as independent observations should be avoided. This may be tempting when sample size is limited, but it violates the basic assumption of independent observations. It creates several distortions like artificially increased sample size, incorrect standard errors, narrower confidence intervals, and altered p-values. One example is that of collecting multiple spectra at different points of the same specimens in Fourier transform infrared spectroscopy (FTIR) analysis

and treating them as independent replicates. This risk may arise in various fields of wood science, especially when dealing with a limited number or small-sized prototypes or samples. This may occur for various reasons, including the constraints of the experimental design, the need to find a balance with the availability of partner industries, and the need to limit costs. Repeated observations need not be avoided, but their dependence must be accounted for in the statistical analysis. Measurements obtained from the same specimen should not be treated as independent replicates; depending on the experimental design, repeated-measures or mixed-effects models may instead be required.

2. Inappropriate data processing can compromise the results. For example, excessive smoothing of experimental curves is another frequent pitfall. Smoothing techniques can help reduce measurement noise, but applying them excessively may distort important features of the curve, such as local maxima, inflection points, or sudden changes in slope. As a result, the smoothed curve may no longer represent the actual behavior of the material. For example, sound absorption of wood-based panels may be measured at 2 Hz intervals by testing devices, and showing the data only for octave-thirds (63, 125, 250, 500, 1000, 2000, 4000, 8000 Hz) may result in overlooking relevant absorption peaks. Additionally, in FTIR spectra of wood and wood-based materials, excessive smoothing can broaden or reduce the intensity of absorption bands, merge closely spaced peaks, and obscure weak shoulders associated with subtle chemical changes. Smoothing should therefore be applied cautiously and only when it does not alter the relevant characteristics of the data.

3. It should also be noted that spectroscopic curves often require additional transformations of raw signals to improve comparability between samples and reduce instrumental effects such as baseline drift or differences in signal intensity. Common procedures include baseline correction and normalization of the spectra. Although these treatments can facilitate comparison between measurements, they may also modify the shape or relative intensity of spectral features if applied improperly. For example, aggressive baseline correction can remove meaningful peaks, while inappropriate normalization may amplify noise or distort peak ratios. These transformations should therefore be applied carefully and consistently to all samples. In addition, the procedures used should be clearly described in the methods section so that the analysis can be reproduced and correctly interpreted.

4. Curves should be plotted in a way that allows readers to easily understand the data. Overcrowded plots with too many curves, unclear legends, or indistinguishable line styles can make interpretation difficult. Showing representative curves together with variability indicators (e.g., error bands or boxplots of derived parameters) often improves clarity. In addition, curves should be presented using consistent units and scales. Differences in axis scaling, unit conversion, or normalization methods can create misleading visual impressions. When multiple curves are compared in the same figure, identical axes and clearly defined units help ensure a fair comparison. In wood science, it is common to investigate the properties of several different statistical groups, for example, in terms of wood species, type of engineered product, adhesive, etc. The risk of overcrowding graphs is therefore often present: when needed, graphical ways to split the plots and increase the readability of figures should be considered.

5. Another common issue concerns the interpretation and presentation of error bars. These are frequently used to show variability in experimental data, but their meaning is not always clearly stated. In particular, it is often unclear whether the bars represent the standard deviation, the standard error of the mean, or confidence intervals. Since these quantities describe different aspects of variability, failing to specify them can lead to incorrect interpretation of the results. In addition, overlapping error bars are sometimes interpreted as evidence that two curves are not significantly different. This assumption is not generally valid, because statistical significance depends on the specific statistical test and sample size rather than on visual overlap alone. A good practice is to clearly report what the error bars represent and to use formal statistical tests to support conclusions about differences between curves. Similarly, failing to report confidence intervals or uncertainty bands hides the accuracy of the results and encourages over-interpretation, especially when differences seem large enough at first glance.

6. Using linearization without recognizing that transformations can alter error structure and bias parameter estimates. Linearization, in fact, alters the error structure of the data and can lead to bias in estimates. Therefore, reasons for applying linearization and the deriving limitations should be clearly mentioned.

7. Comparing fitted parameters without checking whether the same model is appropriate for all groups can lead to misleading results because different groups may be governed by different relationships. In such cases, the analysis may reveal statistical differences, when the actual difference lies in the fact that the model fits some groups only. For example, engineered wood products bonded with different

adhesives may react differently to environmental conditions because thermoplastic and thermosetting adhesives behave according to different models.

8. Using high-order polynomials without physical justification, just to get nice fittings, may lead to several drawbacks, including overfitting. For example, polynomials may very well fit the data and oscillate widely between them at the same time. Also, polynomials may return inconsistent predictions, even slightly outside the range of the measured data.

9. Fitted curves are sometimes used to estimate values outside the range of experimental data. This can be risky because the model may not represent the physical behavior of the material outside the measured interval. In wood science, properties often change nonlinearly with variables such as MC, temperature, or strain, so extrapolated values may be unreliable. For example, the modulus of elasticity behaves differently within the proportionality limit (the maximum stress at which the stress-to-strain ratio remains linear) and above it, where different types of non-linear behavior occur. Similarly, dimensional changes in wood are proportional to MC change below the fiber saturation point (FSP). Above the FSP, further changes in MC generally produce negligible or no additional swelling or shrinkage. Interpretations should therefore be limited to the range of observed data unless the model has a clear theoretical basis.

Limitations

The purpose of this work is to provide an introductory reference for dealing with the most common cases of curve comparison in wood science. Therefore, this article has a deliberately focused scope. Accordingly, it does not aim at replacing the study of statistics at the postgraduate and PhD student level, nor at replacing the more comprehensive content of statistical handbooks. In particular, several statistical concepts are mentioned briefly here. This is because full discussion of these concepts would require more space, and would be better suited to a book chapter or a book. Similarly, the flowchart in Figure 3 is intended as a general guideline for the most common situations, while deepening into the various elements is left to the development of the research being conducted.

Conclusion

Curves are fundamental in wood science for describing the relationships between independent and dependent variables. However, understanding and effectively comparing curves is not always straightforward, which can limit their practical application. The methodological flowchart included in this article provides an operational reference to tackle the key decision points in curve comparison, including choosing independent and dependent variables, fitting curves, applying linearization if appropriate, and comparing curves. Common pitfalls, including excessive smoothing, overfitting, and failure to assess the meaning of the parameters considered, are also listed as a practical tool for performing sound research.

While open questions in curve comparison mainly concern advanced or application-specific cases, a broader methodological issue is the incomplete reporting of model selection, diagnostics, dependence structures, and uncertainty. The framework proposed here may help authors describe these aspects more clearly in the materials and methods sections of their papers, thereby improving the reproducibility and comparability of curve-based studies in wood science.

A proper understanding of curves goes beyond simply avoiding errors or following standard procedures. Instead, it enables a more proactive approach to research which allows researchers to handle curves of varying complexity with confidence. This, in turn, allows for a deeper and more precise description and comprehension of the phenomena under investigation, ultimately contributing to the advancement of knowledge in wood science.

References

Abe H, Watanabe K, Ishikawa A, Noshiro S, Fujii T, Iwasa M, Kaneko H, Wada H (2016) Simple separation of *Torreya nucifera* and *Chamaecyparis obtuse* wood using portable visible and near-infrared spectrophotometry: differences in light-conducting properties. J Wood Sci 62:210–212. https://doi.org/10.1007/s10086-016-1541-z

Altgen M, Altgen D, Pasca D, Zimmer K (2026) Temporary property changes and moisture-driven recovery in thermo-mechanically densified birch and aspen wood. Eur J Wood Wood Prod 84:35. https://doi.org/10.1007/s00107-026-02400-w

Adutwum JO, Sakagami H, Koga S, Matsumura J (2023) An application of mixed-effects model to evaluate the role of age and size on radial variation in wood specific gravity in teak (*Tectona grandis*). J Wood Sci 69:9. https://doi.org/10.1186/s10086-023-02076-3

Belt T, Treu A, Möttönen V, Altgen M (2026) Moisture dynamics during *Coniophora puteana* brown rot degradation of Scots pine sapwood. Int Biodeterior Biodegrad 209:106296. https://doi.org/10.1016/j.ibiod.2026.106296

Bozkurt Ö, Özbay G (2026) Artificial neural networks as a decision-support system for predicting the quality attributes of thermally modified wood. Eur J Wood Prod 84:125. https://doi.org/10.1007/s00107-026-02478-2

Calvano S, Bombieri A, Rizza D, Bergante S, Chiarabaglio PM, Cremonini C, Negro F, Zanuttini R (2026a) New poplar clones from conventional and agroforestry plantations in Northern Italy: preliminary results on peeling attitude and properties of solid wood and plywood. Forests 17:130. https://doi.org/10.3390/f17010130

Calvano S, Negro F, Goli G, Togni M, Cremonini C, Zanuttini R (2026b) Active behavior of poplar wood bilayers in response to changes in relative humidity. J Build Eng 118:115087. https://doi.org/10.1016/j.jobe.2025.115087

Chen F, Peng J, Jiang Y, Shi L, Ma J, Ma XQ, Cao Z, Li M (2026) Multi-trait evaluation and selection of elite *Cunninghamia lanceolata* provenances for high-quality, large-diameter and knot-free timber. Ind Crops Prod 241:122868. https://doi.org/10.1016/j.indcrop.2026.122868

Dai C., Yu C, Jin J (2008) Theoretical modeling of bonding characteristics and performance of wood composites. Part IV. Internal bond strength. Wood Fiber Sci 40(2):146–160.

Dang YQ, Zhang YL, Wang XY, Xin B, Quinn NF, Duan JJ (2021) Retrospective analysis of factors affecting the distribution of an invasive wood-boring insect using native range data: the importance of host plants. J Pest Sci 94:981–990. https://doi.org/10.1007/s10340-020-01308-5

Eslamdoust J, Vasiliauskas S, Little T, Ashiq MW (2026) Height–diameter modeling of tree species in boreal and mixed forests using a mixed-effects approach and stand-level variables. For Sci Technol 22(1):217–230. https://doi.org/10.1080/21580103.2025.2571222

EN 326-1 (1994). Wood-based panels - Sampling, cutting and inspection - Part 1: Sampling and cutting of test pieces and expression of test results. European Standards. https://www.en-standard.eu/bs-en-326-1-1994-wood-based-panels-sampling-cutting-and-inspection-sampling-and-cutting-of-test-pieces-and-expression-of-test-results/

Hedderich J, Sachs L (2024) Applied Statistics. Methods using R. Springer Berlin, Heidelberg. 1023 p. https://doi.org/10.1007/978-3-662-70074-7

Hu W, Guan H, Zhang J (2018) Finite element analysis of tensile load resistance of mortise-and-tenon joints considering tenon FIT effects. Wood Fiber Sci 50(2):121–131. https://doi.org/10.22382/wfs-2018-014

Hukka A, Viitanen HA (1999) A mathematical model of mould growth on wooden material. Wood Sci Technol 33:475–485. https://doi.org/10.1007/s002260050131

ISO 3129 (2019). Wood - Sampling methods and general requirements for physical and mechanical testing of small clear wood specimens. International Standards. https://www.iso.org/standard/74839.html

Jang ES, Park HJ (2026) Horizontal flame spread of flame-retardant-treated Japanese cedar (*Cryptomeria japonica*) exterior siding material. Wood Fiber Sci 58(1):53–59. https://doi.org/10.22382/wfs-2026-06

Liang H, Zhang M, Gao C, Zhao Y (2018) Non-destructive methodology to determine modulus of elasticity in static bending of *Quercus mongolica* using Near-Infrared Spectroscopy. Sensors 18:1963. https://doi.org/10.3390/s18061963

Longford NT (2010) Analysis of Covariance. In: Peterson P., Baker E., McGaw B. International Encyclopedia of Education (Third edition). Elsevier, Amsterdam, pp. 18–24. https://doi.org/10.1016/B978-0-08-044894-7.01304-X

Martins C, Ferreira C, Monteiro SRS, Dias AMPG (2025) Mechanical properties of sweet chestnut sawn wood and glulam performance. Eur J Wood Wood Prod 83:112. https://doi.org/10.1007/s00107-025-02266-4

Miyamoto BT, Sinha A (2025) Time-temperature effects on early-stage primary thermal creep of plywood and oriented strand board (OSB) at elevated temperatures. Wood Fiber Sci 57(1):15–26. https://doi.org/10.22382/wfs-2025-02

Niemz P, Teischinger A, Sandberg D (2024) Springer Handbook of Wood Science and Technology. Springer Nature Switzerland, Cham. 2069 p. https://doi.org/10.1007/978-3-030-81315-4

Nugroho WD, Anjayani GD, Tampubolon AF, Pradhana AW, Cipta H, Tiyasa NP, Yamagishi Y, Kudo K, Rahman MH, Nakaba S, Funada R (2025) Stem gravitropism and tension wood formation in three tropical woody species with different wood densities. J Wood Sci 71:23. https://doi.org/10.1186/s10086-025-02196-y

Nyobe CJ, Nyobe NS, Nkibeu JB, Oum Lissouck R, Ayina Ohandja LM (2024) Effect of slope of grain on mechanical properties of some tropical wood species. Wood Mater Sci Eng 20(3):448–494. https://doi.org/10.1080/17480272.2024.2356047

Pečenko R, Hozjan T, Huč S (2023). Modelling charring of timber exposed to natural fire. J Wood Sci 69:19. https://doi.org/10.1186/s10086-023-02091-4

Pulgar JA, Riesco G (2024) Implementing linear mixed effects models to enhance estimation of the dimensional stability of wood of *Laurus nobilis* L. For Syst 33(2):e05. https://doi.org/10.5424/fs/2024332-20854

Robinson N, Drever CR, Gibbs DA, Lister K, Esquivel-Muelbert A, Heinrich V, Ciais P, Silva-Junior, CHL, Liu Z, Pugh, TAM, Saatchi S, Xu Y, Cook-Patton SC (2025) Protect young secondary forests for optimum carbon removal. Nat Clim Chang 15:793–800. https://doi.org/10.1038/s41558-025-02355-5

Rezakalla AI (2025) The impact of water exposure on the mechanical properties of a wood-plastic composite. Wood Fiber Sci 57(4):205–212. https://doi.org/10.22382/wfs-2025-18

Schuler Frantzen B, Silveira EA, Cruz Lamas G, de Oliveira TF (2026) Intermittent versus continuous kiln drying of hardwood boards: comparable-stage analysis of energy use, kinetics, quality, and climate impacts. Biomass Bioenerg 213:109443. https://doi.org/10.1016/j.biombioe.2026.10944

Svoboda T, Gaffová Z, Rajnoha R, Šatanová A, Kminiak R (2017). Bending forces at the proportionality limit and the maximum - technological innovations for better performance in wood processing companies. BioResources 12(2):4146–4165. https://doi.org/10.15376/biores.12.2.4146-4165

Hu W, Liu Y, Li S (2021) Characterizing mode I Fracture behaviors of wood using compact tension in selected system crack propagation. Forests 12:1369. https://doi.org/10.3390/f12101369

Xu X, Pan R, Chen R (2021) Combustion characteristics, kinetics, and thermodynamics of pine wood through thermogravimetric analysis. Appl Biochem Biotechnol 193:1427–1446. https://doi.org/10.1007/s12010-020-03480-x

Zhang Y, Peng Z, Song Z, Schilling JS (2025) Repeated measures of decaying wood reveal the success and influence of fungal wood endophytes. mSystems 10:e00382-25. https://doi.org/10.1128/msystems.00382-25