Modelling wood moisture content in outdoor conditions from measured data

Niklewski, Jonas; van Niekerk, Philip Bester; Meyer-Veltrup, Linda; Sandak, Jakub; Brischke, Christian

Published in:
IRG55 Scientific Conference on Wood Protection : Knoxville, Tennessee, USA, 19 - 23 May, 2024

2024

Document Version:
Early version, also known as pre-print

IRG/WP 24-YYXXX

THE INTERNATIONAL RESEARCH GROUP ON WOOD PROTECTION

Section Y **VVVVVVVVVV**

Modelling wood moisture content in outdoor conditions from measured data

Jonas Niklewski[1], Philip Bester van Niekerk[2], Linda Meyer-Veltrup[3], Jakub Sandak[4], Christian Brischke[5]

[1] Lund University, John Ericssons väg 1, SE-223 63, Lund, Sweden
[2] University of Goettingen, Buesgenweg 4, D-37077, Göttingen, Germany
[3] Heinz-Piest-Institute for Skilled Crafts, Wilhelm-Busch-Straße 18, D-30167 Hannover, Germany
[4] InnoeRenew CoE, Livade 6a, 6310, Izola, Slovenia
[5] Thünen-Institute of Wood Research, Leuschnerstraße 91d, D-21031 Hamburg, Germany

Paper prepared for the IRG55 Scientific Conference on Wood Protection
Knoxville, Tennessee, USA
19 - 23 May, 2024

IRG SECRETARIAT
Drottning Kristinas v. 61B
SE-114 86 Stockholm
Sweden
www.irg-wp.com

Modelling wood moisture content in outdoor conditions from measured data

Jonas Niklewski[1], Philip Bester van Niekerk[2], Linda-Meyer Veltrup[3], Jakub Sandak[4], Christian Brischke[5]

[1] Lund University, Sweden, jonas.niklewski@kstr.lth.se

[2] University of Goettingen, Buesgenweg 4, D-37077, Göttingen, Germany, philipbester.niekerk@uni-goettingen.de

[3] Heinz-Piest-Institute for Skilled Crafts, Wilhelm-Busch-Straße 18, D-30167 Hannover, Germany, meyer@hpi-hannover.de,

[4] InnoRenew CoE, Livade 6a, 6310, Izola, Slovenia, jakub.sandak@innorenew.eu,

[5] Thünen-Institute of Wood Research, Leuschnerstraße 91d, D-21031 Hamburg, Germany, christian.brischke@thuenen.de

ABSTRACT

Sustainable use of wood requires an understanding of expected service life, particularly when the material is exposed to outdoor conditions and, thus, fungal decay. Since moisture is the primary vector for fungal decay, accurate moisture prediction is a key component in service life assessment. For this purpose, the present study leverages existing measured data for linear regression of in-field moisture conditions of different wood species against climate parameters. Predictors of precipitation, relative humidity, and temperature were used in a finite distributed lag model to account for present and previous weather records. Issues of collinearity were addressed by ridge regression. The resulting model was, in general, able to describe the important features of different wood species. However, large errors were observed in certain periods, and it was hypothesized that these were related to thawing. Nevertheless, the results encourage additional effort into data-driven modelling of moisture content from measured data, and it is believed that non-linear models such as random forests and neural networks will be able to describe additional features and, in doing so, reduce the error. The study contributes to the ongoing efforts in developing effective, user-friendly, and open-source tools for performance-based service life assessment of wood. By improving our understanding of moisture content prediction in different softwoods, this research aims to enhance the reliability and sustainability of wood as a construction material.

Keywords: Moisture, linear regression, wood, species, precipitation, decay, distributed lag

1. INTRODUCTION

Service life assessment of wood in exposed environments enables informed decision-making in the design phase and is thereby key to the sustainable use of the material. The service life of wooden applications is affected by a variety of biotic and abiotic deteriorating mechanisms of different natures and outcomes. Fungal decay is particularly problematic as it eventually breaks down the material in its entirety, undermining any intended function.

The rate of fungal decay in wood is influenced by moisture content and temperature, both of which vary with the ambient climate. Moisture content, however, also depends on factors like design choices, distance to ground, the type of wood, and design detailing (Brischke et al., 2006). For instance, design flaws are often exploited in accelerated field tests to facilitate fungal decay (Meyer et al., 2016). For a detailed, quantitative approach, the temporal variation of moisture and temperature can be connected to decay development through dose-response models (Brischke and

Meyer-Veltrup, 2016). The dose model infers the relative suitability for decay development and aggregates the response over time to an equivalent exposure, which is then connected to the resistance of the material.

Natural wood has two characteristics that affect service life. First, different wood species have different inherent protective properties against fungal decay. This means that different species decay at different rates when subject to the same moisture content and temperature. Second, differences in moisture-related properties (e.g. permeability) leads to significant differences in moisture variations between species when subject to the same wetting cycles (Van den Bulcke et al., 2009). Several previous studies have set out to characterize the moisture performance of different wood products from laboratory indicators (De Windt et al., 2018; Emmerich et al., 2020). In the context of service life prediction, the distinction between moisture dynamics and inherent properties has been used to develop a general model for material resistance to fungal decay (Meyer-Veltrup et al., 2017a).

Considering the above, service life assessment includes three components: the assessment of material climate, the characterization of material resistance and a dose-response model to connect these two. Numerical models offer a sound and flexible approach to moisture content prediction as the method departs from the underlying process. However, the reliance on mathematical descriptions becomes a limitation in scenarios where phenomena are complex or not fully understood. For instance, existing numerical models developed specifically for service life assessment have limited ability to describe end-grain absorption and other wood species than spruce (Niklewski and Fredriksson, 2021). Ongoing and future research in numerical modelling, including recent studies on moisture absorption in spruce end-grain (Brandstätter et al., 2023; Buck et al., 2023; Kalbe et al., 2022) and side-grain of different species (Glass et al., 2023), holds promise for addressing these challenges. Nevertheless, the development of a general numerical model that accurately simulates wood moisture transport under conditions of cyclic exposure remains an ambitious and not yet fully realized objective. Current service life design frameworks manage this limitation by using spruce as a baseline for climate exposure, treating variations in moisture properties as a component of material resistance (Meyer-Veltrup et al., 2017a). This approach is practical, as it effectively renders the assessment of climate independently of the material. However, it also generates additional bias in service life assessment, as the wood species influences both moisture variations and inherent resistance.

This work presents a first attempt to model moisture variations of different softwood species using multi-variate analysis. The model does not require any explicit formulation of transport equations, or even knowledge of the underlying mechanisms. Instead, the model is simply trained from data to generalize to new data. One disadvantage is that the model needs high-quality data for training, which may be scarce. In previous work, a numerical model was instead used as a basis for training a data-driven model, effectively creating a computationally efficient metamodel for analysis of big data sets (Hosseini et al., 2023). However, since the model builds on top of the numerical model, it does not resolve any of its limitations. In this work, we address this problem by developing models from measured data instead. Measured data are obtained from Meyer-Veltrup et al. (Meyer-Veltrup et al., 2017b) and include several wood species and different variations of detailing.

The aim of this paper is to develop a simple yet useful data-driven model for moisture content prediction of different softwoods, focusing specifically on spruce and pine. The work is part of an ongoing effort to develop accurate, user-friendly, and open-source tools which facilitate performance-based service life assessment of wood. In doing so, we aim to improve the reliable use of wood in diverse applications, optimizing its potential for sustainable and enduring construction solutions.

2. METHOD

2.1 Dataset

2.1.1 Material climate

The data used in the present study originate from an experiment conducted in Hannover, Germany, (coordinates 52.395067; 9.701913) to compare different methods of accelerating decay during field durability testing. As decay is dependent on wood moisture content, accelerated test methods primarily differ in their ability to maintain favourable moisture conditions over time. The experiment in question monitored wood temperature, wood moisture content and decay development in several different wood species and several different designs over 4 years. Moisture content was monitored by resistance-type moisture sensors with insulated electrodes, and the measurements thus reflect the variation in moisture content at a specific point – usually at mid-thickness of a specimen. Decay was assessed every 6 months through pick-tests, evaluated according to the EN 252 (2015) rating from 0 (sound) to 4 (failure due to decay).

In the present study, only a subset of data is used. Specifically, we used data on horizontal and vertical boards of Norway spruce (*Picea abies*) and Scots pine sapwood (*Pinus sylvestris*). The objective was to model the moisture content variation of the horizontal boards, but vertical boards of Scots pine heartwood were used in the process. The dimensions of both types of boards were 500×100×20 mm^3 (length × width × thickness). The vertical boards were designed as board-on-board cladding, with the top end-grain being covered by a tin roof. Measurements were taken in the middle of the board thickness.

Figure 1: Experimental setup with horizontal boards (decking) and south-oriented vertical boards (cladding).

2.1.2 Weather data

Weather data were obtained from a station located about 400 m from the test-site, including hourly averages of relative humidity, temperature, wind speed, diffuse and global radiation, as well as hourly totals of precipitation. Additional 5-minute data on precipitation were obtained from a nearby weather station (Hannover-Herrenhausen, coordinates 52.3965; 9.6629) maintained by the

German Weather Service, located about 2 km from the site of the experiment. For modelling, the 5-minute data was used to calculate the relative duration of wetting, d_{wet}, defined as follows:

$$d_{wet} = \frac{1}{288}\sum_{i=1}^{288} 1_{\{p_5>0\}} \tag{1}$$

where $1_{\{p_5>0\}}$ is an indicator function which equals one when precipitation is registered within the 5 min period, and otherwise zero. The relative duration of wetting varies from 0 (no rain recorded) to 1 (rain recorded every 5-minute period). The daily precipitation collected on-site was only used to validate that the 5-minute data were sufficiently representative of the location.

In addition to the above, a few extra weather variables were obtained from a different station, located about 10 km from the experimental site (Hannover, coordinates 52.4644; 9.6779), including daily averages of snow depth, sunshine duration, and cloud cover. Due to the distance from the experimental site, these variables were deemed less reliable, and were therefore only used for diagnosing the model errors.

2.2 Model development

This study focused on exploring the dataset through simple models. For this purpose, we exclusively focus on multiple linear regression. Based on the results, more complex approaches are proposed.

2.2.1 General model

The general form of a linear regression model is:

$$y = \boldsymbol{\theta}^T \boldsymbol{X} + \epsilon \tag{2}$$

where y is the response, $\boldsymbol{\theta^T}$ is the transposed vector of coefficients to be fitted, **X** is a matrix of predictor variables and ϵ is the error term. Due to the dynamic nature of moisture variations, the response on any day depends on past events. Here, we use a simple unstructured finite distributed lag model (DLM) where **X** includes current and past values of independent variables. Note that autoregressive terms (previous values of the dependent variable) are not included. The coefficients of predictors are determined by minimizing the cost function, J(θ), which in this case describes the average squared error as well as a regularisation term:

$$J(\theta) = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}(x_i;\theta) - y)^2 + \lambda\sum_{j=1}^{m}\theta_j^2 \tag{3}$$

where $\hat{y}$ is the predicted response (for the observation x_i with coefficients θ), y is the observed response, *n* is the number of observations, *m* is the number of parameters and λ is the regularisation parameter. Ridge regularisation was used during model training to manage the bias-variance trade-off and avoid coefficient inflation in cases of multicollinearity.

Model specification

Model specification is the process of selecting a set of suitable predictors which are related to the response. The effects considered here include relative humidity, temperature, and relative duration of precipitation. However, additional predictors can be constructed by transformations and interaction which, in combination with possible lags, lead to many possible variations.

Domain knowledge of the process in question was used to guide model specification in selecting relevant variables and interpretation of results. On a macroscopic scale, the process driving moisture variations in wood is well understood, and the following guides can be used as a basis for specification:

- After extended periods without precipitation, the relationship between the surrounding climate and the moisture content is governed by the equilibrium moisture content (EMC) and the transport coefficient. As a crude approximation, these properties can in the hygroscopic range be considered similar for many softwoods (Glass and Zelinka, 2021; Hukka, 1999).
- During periods of rain, the moisture content will temporarily exceed the equilibrium moisture content. The timing of the peak is increasingly dampened and delayed with increasing depth from the exposed surface.
- After a period of rain, the moisture content will again tend towards the EMC. This process is governed by relative humidity and temperature, among other factors.

Describing these dynamics with a linear regression model is challenging. First, the influence of relative humidity and temperature depends on the presence of precipitation. Second, the dynamics of the measured moisture variations depend on the nature of the measurement system and depth. For example, the moisture content on the surface will increase immediately as rain lands on the wood surface, but it can take several hours for the response to be recorded at the depth of measurement (Niklewski, 2018). Similarly, for gravimetric measurements the response will be immediate but depend on the dimension of the wood in question.

Nevertheless, guided by the points above we can exploit the fact that, everything else being equal, the moisture content of rain-exposed wood will consistently exceed the moisture content which would occur under sheltered conditions. The difference between these two cases specifically captures the effect of rain. We can then describe the moisture content of rain-exposed wood as the sum of two terms:

$$u_t = u_{t,1} + u_{t,2} \tag{4}$$

where $u_{t,1}$ is the moisture content of sheltered wood and $u_{t,2}$ is the difference between rain-exposed and sheltered wood, as shown in Fig. 1. This type of decomposition is useful when using simpler models with limited data for training, as it helps to both construct and evaluate the model. In addition, using a single model as a baseline for several wood species adds robustness and consistency to the predictions. For example, under dry conditions, all wood species with a common model for $u_{t,1}$ will tend towards the same value.

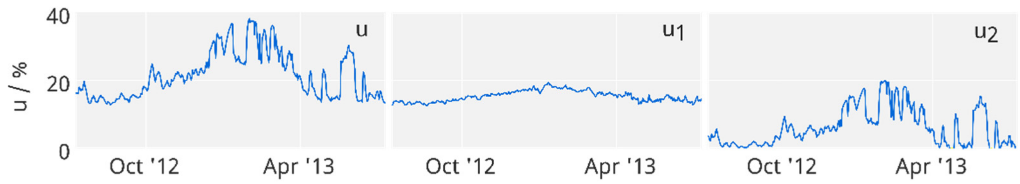

Figure 2: The moisture content, u, decomposed into a sheltered part, u_1, and a rain-induced part, u_2.

The main drawback of this decomposition, in addition to potential bias, is the need for two separate datasets: one set of measurements for rain-exposed wood specimens and a parallel set for sheltered conditions. A sheltered reference was not included in the present dataset. Therefore, a vertical south-oriented specimen of Scots pine heartwood was used as a proxy for sheltered conditions.

Due to the relatively low permeability of Scots pine heartwood, its vertical orientation, and the fact that measurements were recorded at a depth of 10 mm, it was reasonable to assume that the influence of rain was small. Additional support is provided by previous experiments, where the difference between vertical exposed wood specimens and actual sheltered wood specimens was small (Isaksson and Thelandersson, 2013). Most importantly, it can be assumed that the moisture content of the horizontal boards of other softwood specimens (spruce and pine sapwood) will consistently exceed that of the selected reference.

For modelling the inherently complex dynamics of rain-induced moisture variations, we use a simplified approach where precipitation is included in all terms, either alone or with interaction to other variables, and without intercept. As a result, a rain event will increase the model output over n periods (where n is the maximum number lag), whereafter it will default to zero. Hypothetically, the number of relevant lags will then govern the period of drying after a rain event.

2.2.2 Collinearity and autocorrelation

Multicollinearity occurs when two or more predictors are correlated and leads to inflated variances of the estimated coefficients, making them unstable and unreliable for interpretation and prediction. In this study, there were two potential sources of multicollinearity: correlation between different weather variables and correlation between lagged versions of the same variable (autocorrelation). There are several ways to deal with collinearity in regression, such as dimensionality reduction (principal component analysis or partial least squares regression), regularisation techniques (lasso or ridge) or iterative selection (forward or backwards). Simpler techniques involve dropping or merging collinear variables.

Initial screening indicated a moderate positive correlation between precipitation duration and relative humidity ($\rho = 0.41$), where rainy days usually measure higher relative humidity. Temperature and relative humidity showed moderate negative correlation (ρ =-0.43), whereas temperature and precipitation only exhibited minor correlation (0.09). Further, near-constant amplitudes in the autocorrelation function (ACF) of temperature indicate very high autocorrelation at small lags, which is expected. Relative humidity also exhibits high autocorrelation, but precipitation did not.

Due to the correlation between relative humidity and precipitation, initial modelling attempts showed that the importance of rain was negligible - regardless of wood species. This is counterintuitive since the main difference between wood species is, in this context, their response to rain. Instead, differences between wood species were erroneously modelled by inflating the coefficients for relative humidity. While this approach resulted in the smallest cost (error) for the specific dataset, it also meant that the resulting model did not generalize well to datasets with long dry periods.

Here, the collinearity was dealt with by (1) dropping relative humidity as a predictor of $u_{t,2}$ and (2) estimating the coefficients of $u_{t,2}$ with ridge regularisation. By omitting variables in the development of $u_{t,2}$, the model is expected to have systematic bias which will, in turn, affect the coefficient estimates and their inference statistics. However, we instead focus exclusively on minimizing the validation error, rather than statistical inference of the predictors of the final model.

2.2.3 Cross validation

Cross validation is done during model training to tune hyperparameters, avoid overfitting and estimate the model error. The dataset is typically split into a validation set and a training set. The model coefficients are estimated from the training set, and the corresponding error is calculated by comparison against the validation set. In the case of time-series analysis, it is advantageous to have the validation set include a full season, i.e. a year of data, as the bias may vary between seasons. This is, however, problematic since the available data is limited in time, and using a full year of data for validation would limit the amount available for training. This problem can be addressed by k-fold cross validation, where the full dataset is used for both training and validation. However, classical k-fold cross validation is not recommended for time-series.

For this purpose, we use a particular type of expanding window cross validation split commonly used for time series. As shown in Fig. 3, the training set expands forward in each fold, followed by a fixed size of trailing validation data. In addition to being a sound method given the context, the difference in error between folds provides additional information on the effect of training batch size. This can be useful for balancing model complexity against data size, effectively avoiding overfitting. The process of optimizing model hyperparameters and simultaneously estimating the error is called flat cross validation. For simple models with relatively few parameters, as in this study, flat cross validation is generally acceptable (Wainer and Cawley, 2021). The alternative is to use nested expanding window cross validation. However, we still use a hold-out set of data to compare the final model against. The third year of data is however not ideal for estimating the error since the effect of structural changes caused by decay and other weathering may begin to affect the measurements.

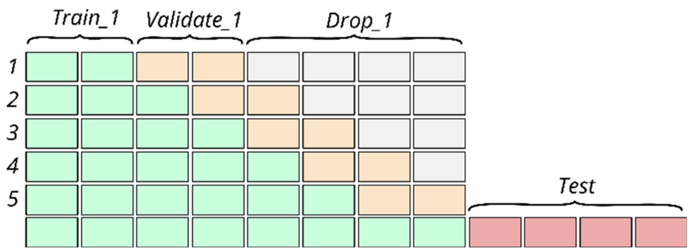

Figure 3: Expanding window cross validation with 5 folds. The training data (green) expand in each fold while the size of the trailing validation set is fixed. Note that the trailing data (without rain) have not been included here.

3. RESULTS AND DISCUSSION

3.1 Model specification

3.1.1 Sheltered wood

The model describing sheltered wood was fitted against temperature and relative humidity. The terms considered were temperature (T), relative humidity (H), polynomial transforms and interaction. The equilibrium moisture content, a non-linear transformation of relative humidity and temperature, $f(H,T)$, calculated according to Glass and Zelinka (2021), was also considered. Initial screening showed that polynomial transforms did not improve accuracy, and these were therefore dropped. Further, the equilibrium moisture content was marginally better than relative humidity. Using equilibrium moisture content and temperature, grid search was used to determine an appropriate number of lags. In each coefficient estimate, ridge regularisation was used to balance the coefficient estimates.

Fig. 4 shows an example of coefficient estimation with three lags per variable and varying degrees of penalty, λ. Note that the minimum validation error is obtained with balanced coefficients ($\lambda \approx 160$), and that the coefficient estimate of present temperature changed sign (from positive to negative) during the process. Regularisation did not reduce the validation error significantly but was necessary to obtain stable coefficient estimates. The optimal penalty was, in general, only non-zero when several lags of temperature were included. In contrast, no local minimum was reached when omitting temperature and including several lags of relative humidity, implying that no regularisation was needed in this case.

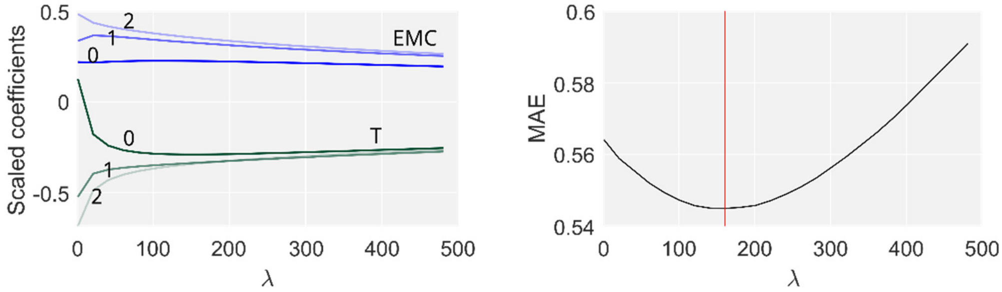

Figure 4: Ridge regression, showing shrinking coefficients (left panel) and validation error (right panel) with increasing penalty, λ.

The left panel in Fig. 5 shows the validation error obtained from the exhaustive grid search, with the red square highlighting the selected model. The right panel shows the resulting moisture content variation after training the model against the full (validation) dataset. The performance increased significantly by including both variables in the model (see first row and first column). However, the value of adding additional variables beyond a single contemporary term of T and a single lagged term of EMC was small.

The grid search did not indicate a local minimum with respect to validation error, since the error marginally decreased with an increasing number of parameters. In this situation, the best model is usually one with relatively few parameters. The choice can be guided by an information criterion, such as Akaike Information Criterion (AIC), to balance model complexity against goodness-of-fit, but this method is not traditionally used in combination with regularisation. Instead, we selected the model with a single (contemporary) interaction term and the corresponding number of lagged variables of relative humidity resulting in the least validation error.

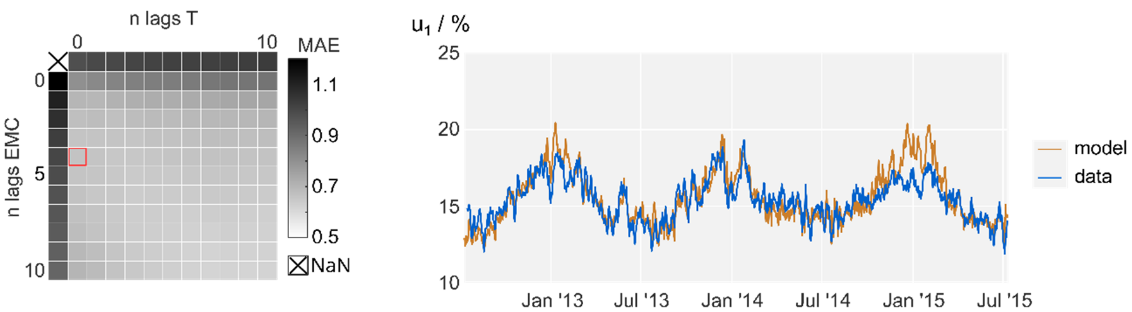

Figure 5: Validation error of different variable combinations (left panel) and the selected model trained and compared against the full dataset (right panel).

The fact that temperature has a strong positive effect on performance is noteworthy, as temperature in this range does not have a significant effect on the equilibrium moisture content (Glass and Zelinka, 2021). Moreover, while temperature has previously been shown to influence the moisture content in outdoor conditions, this effect is usually much smaller than relative humidity (Brennan and Pitcher, 1995). One possibility is that the temperature is affecting the moisture content readings via its effect on the electrical resistance. However, this is unlikely as the calibration curve was calibrated for different temperatures, and wood temperature was measured during the test. Alternatively, it could be that the temperature carries some information about the long-term trend in relative humidity.

It can be noted that the history of relative humidity (6 days) is short relative to previous studies, where it has been suggested that monthly averages can be used to determine (global) moisture content (Brennan and Pitcher, 1995). This can also be compared to a previous study based on numerical data, where it was found that a history of 10 days was sufficient to describe moisture content in boards (20 mm) in a wide variety of climate conditions (Hosseini et al., 2023). It is possible that the model would be more robust and generalize better by increasing the number of lagged terms of relative humidity. In the present study, this would increase the validation error by an insignificant amount.

Qualitatively, and except for the third winter, the model performance is satisfactory. Since the linear nature of the model is unable to capture the effect of hysteresis, an absolute error of less than ±0.5% should not be expected. Signs of large systematic errors can, however, be observed during the first and last winter seasons, where the model underestimates the peak in moisture content. Interestingly, previous numerical simulations have indicated similar discrepancies during the winter months (Niklewski et al., 2016). We were unable to conclusively find the reason for this discrepancy. It is noteworthy that the relative humidity of the third winter was on average lower than previous years, with vastly fewer hours exceeding 99%.

Speculatively, the higher measured moisture content during winter could be caused by the difference between weather records (used in the analysis) and the microclimate at the wood surface. A comparison between the records of air temperature and wood temperature revealed that the daily maximum wood temperature exceeds the air temperature during the warm months and approximately equals the wood temperature during the colder months. This difference is explained by surface heating by solar radiation. Conversely, the minimum daily wood temperature was generally lower than the air temperature due to radiative cooling. As a result, water may condense at the cooler wood surface and increase the moisture content.

3.1.2 Exposed wood

Modelling the exposed component of the moisture variation is more challenging. First, the dynamics of liquid water transport in wood is less well understood compared to diffusion in the hygroscopic range. From a modelling point of view, precipitation as a predictor is also associated with a large degree of uncertainty. Fortunately, it exhibits much less autocorrelation than relative humidity and temperature.

The decomposition into sheltered and exposed wood stipulates that the response should tend to zero after a longer period without precipitation. A simple baseline model can therefore be specified from lagged predictors of precipitation and without intercept. The lack of intercept imposes a constraint on the model and therefore introduces additional bias, i.e. including an intercept would reduce the error.

The left panel of Fig. 6 shows the reduction in validation error against maximum lag order and the right panel shows the moisture content variation (corresponding to 10 lags) together with the snow depth over the same period. No regularisation was used for coefficient estimates, which is equivalent to setting λ in Eq. 2 equal to zero. The coefficient weight decreased with increasing lag, except for the first lag that was consistently smaller than the second. This can be explained by the delayed and dampened response in the wood core when the surface is subjected to a sudden increase in moisture content. Previous studies on similar specimens under controlled conditions suggest that the delay between the rain event and the peak in moisture content could be in the order of 10-20 hours at a depth of about 10 mm (Niklewski et al., 2018). Since the time of day when the rain events occur is unknown, it is likely that the main impact of a rain event will be recorded the following day. Similar results have been obtained from numerical studies (Hosseini et al., 2023).

While there are significant discrepancies between the simple baseline model and the measurements, as seen from the average validation error of ~2.6% moisture content, the predicted effect of precipitation did capture many of the measured features. The timing and width of the distinct peaks are quite accurate, but their amplitudes are not consistent. During the winter seasons, the measured moisture content is generally sustained at an elevated level, which is not adequately captured by the model. Interestingly, the largest discrepancy is again observed in the third winter where the model underestimates the increase in moisture content.

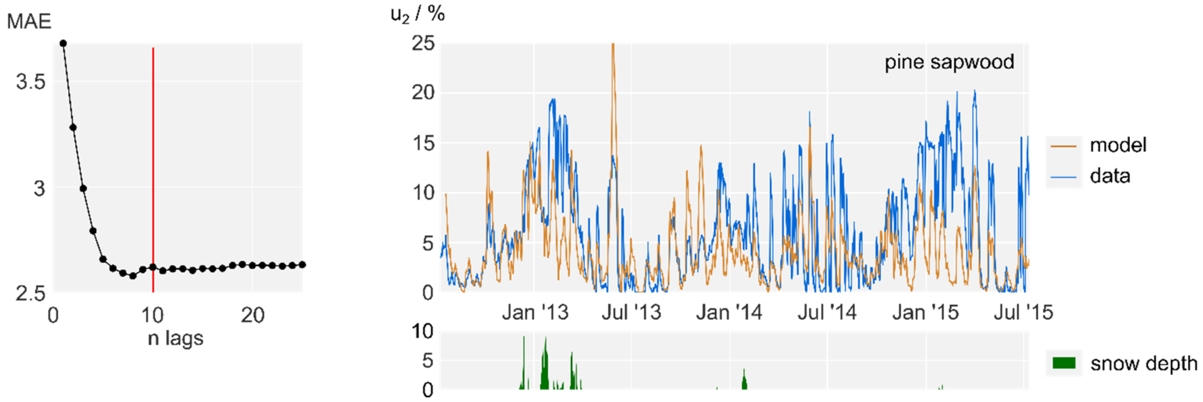

Figure 6: The mean absolute validation error (MAE) for varying number of lags (left panel), and the comparison between measured and predicted (rain-induced) variation in moisture content (right panel). The variation of snow-depth over the same period is included.

Qualitatively, there seems to be some indication of the model underestimating the moisture content during winter and vice versa, which would point towards temperature being an important but omitted variable. During detailed inspection of residuals, it was observed that the model tends to underestimate the moisture content specifically when the temperature increases from negative to positive, which could point towards the effect of thawing. In some periods, specifically from the end of January to early April 2013, the measured moisture content peaked during thawing (as indicated by decreased snow cover). Specifically, on the 28th of January, the moisture content increased by 8% points without any records of rain (although a rain event did occur the day prior). This period was then followed by a period of fluctuating snow cover and frequent rain, during which the moisture content was sustained at a very high level.

It should be noted that the above observation is not evidence of causality between snowmelt and increased moisture content, but rather an indication of some correlation between thawing and increased moisture content. The effect of thawing and freezing, or phase change of water in general, on the wood temperature has been observed in previous studies. In general, when the wood moisture content is above cell wall saturation and the ambient temperature is increased from subzero to positive, then the wood temperature remains at zero during melting (Brischke and Rapp, 2008; Klement et al., 2021). It is unclear how the transition from water to ice affects the electrical resistance and thus the apparent moisture content.

In the context of durability, the importance of precise moisture content estimates decline as the temperature becomes close to freezing, as the conditions for decay fungi become unfavourable. In fact, dose-response models often assume that the rate of decay is zero on days where the minimum wood temperature is below zero. From a problem-oriented perspective, it could therefore be rational to apply lower weights to measurements on days with subzero (or near-zero) temperatures. This could be achieved via weighted least squares.

3.2 Precipitation as a predictor

The model for $u_2(t)$ is based on lagged versions of wetting duration. Wetting as a predictor performed better than daily precipitation amount, both in terms of validation error and qualitative assessment of the moisture variation. Wetting duration is however calculated from 5-minute precipitation data, which are generally not easily accessible for prediction. It is therefore of interest to investigate how the error changes when using readily available precipitation input.

Most locations in Europe can access hourly precipitation data, either through weather stations or reanalysis like ERA-5 (Muñoz Sabater, 2019). Like 5-minute data, hourly precipitation data can be used to estimate the daily duration of wetting. Fig. 7 shows the measured hourly precipitation against the number of 5-minute periods where precipitation was registered. The data is obtained from the year 2022, measured over about 1.000 stations in Germany. Note that the whiskers extend to non-outlier data, but the 25th and 75th percentiles include all data. The data shows that, for example, when rain is recorded in 8 separate 5-minute periods over the same hour, then 50% of the time the station recorded between 0.19 and 0.74 mm of total precipitation, with a median value of 0.36.

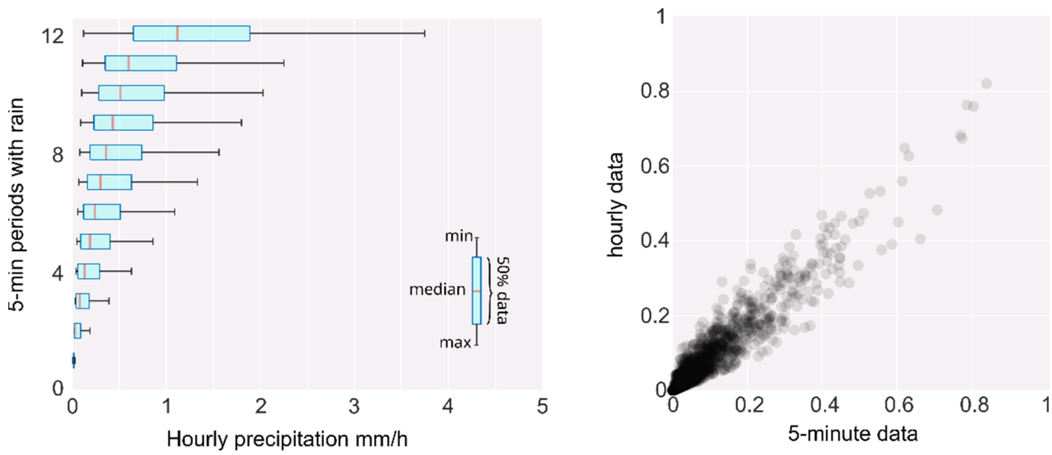

Figure 7. Number of 5-minute periods with recorded precipitation and the hourly amount of precipitation in the same hour, based on 2022 data from about 1.000 stations in Germany (left panel) and the daily wetting duration (relative) calculated from 5-minute and hourly data, respectively, with data from the present study (right panel).

Substituting the relative duration of wetting for amount of daily precipitation increased the validation error from 2.6 to 3.5. The error could be reduced to approximately 2.9 by estimating the duration of wetting from hourly data, which was done by setting a threshold of 1 mm hourly precipitation to define an hour of complete wetting (based on the left panel in Fig. 7) and then interpolating linearly between 0 and 1 mm. The comparison between wetting duration calculated from hourly data and 5-minute data is shown in the right panel of Fig. 7, using data from the present experiment. The comparison indicates that raw daily rain data is not an ideal predictor of the response in moisture content, and the duration of wetting should be estimated from the highest available resolution of the available data.

Several interactions between precipitation (amount) and other variables were tested, however none of them significantly reduced the validation error.

3.3 Final model

For modelling different species, the reference model u_1 is added to the model u_2 which is trained for each species separately, using the same model structure as for Scots pine Sapwood. The resulting moisture content variation for three different wood species, including Norway spruce, Scots pine sapwood and Beech are shown in Figure 8, where the later was included to test the approach against a low-durability hardwood. For all species tested, the inference on accuracy follows those previously discussed for Scots pine sapwood. The model captures the general features and differences in response to rain between wood species, but the accuracy varies over the time-period.

The maximum and minimum moisture content measured on each day (in the three replicate specimens) gives an indication of variation between the triplet specimens. This variation varied from negligible, in the case of vertical Scots pine heartwood and sapwood, to pronounced for Beech and Scots pine sapwood. In most cases, it could be seen that each single specimen was quite consistent in its rank compared to the others, i.e. each specimen consistently measured lower or higher than the others. In some cases, individual specimens drifted in ranking over time, going from the lower range to the upper range. This drift will naturally lead to a drift in the average as well, to which the model is fitted, meaning that the variance in the response will change over time, violating the assumption of stationarity.

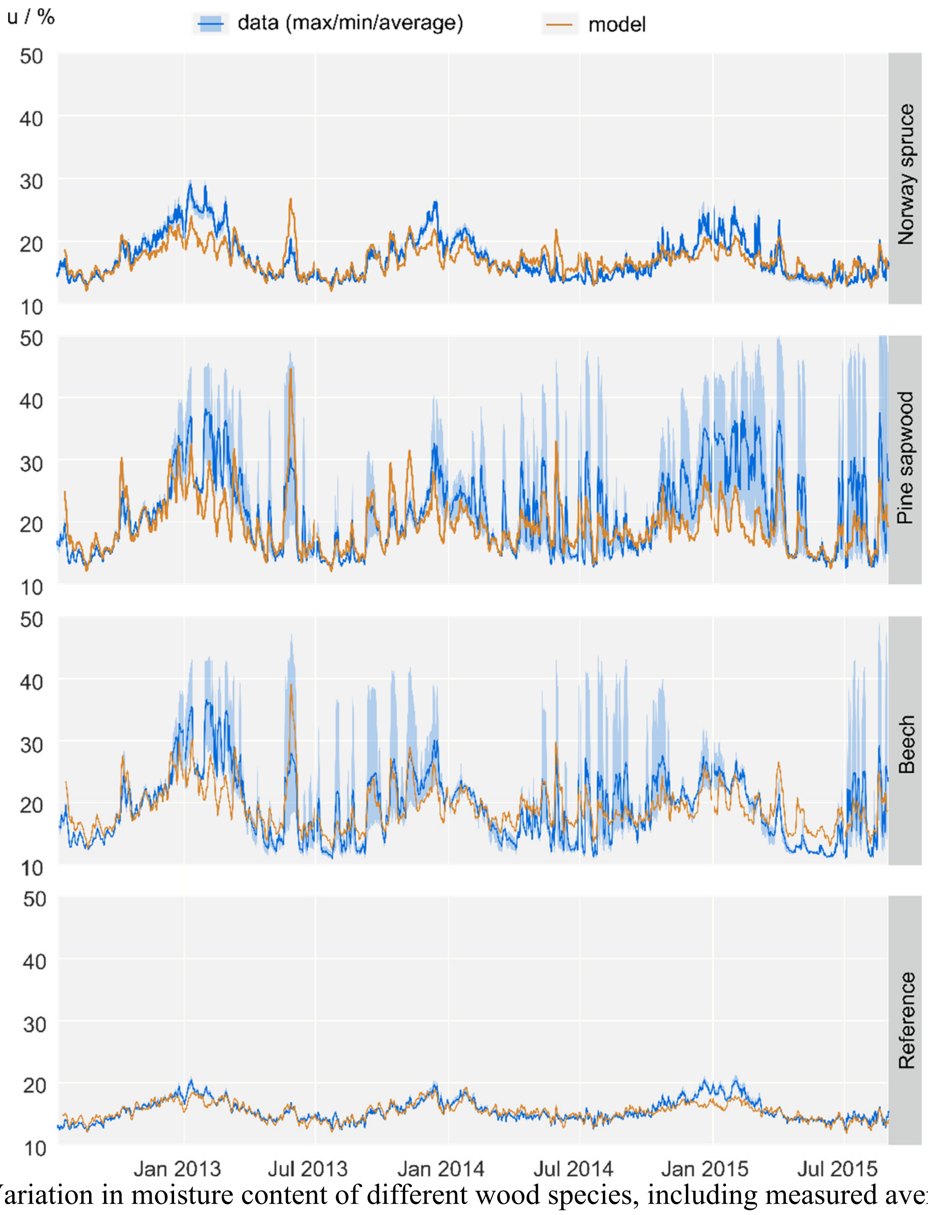

Figure 8: Variation in moisture content of different wood species, including measured averages (solid blue), measured min/max (shaded region) and the predicted response (solid orange). The reference model, which was trained against vertical Scots pine heartwood, is included at the bottom.

3.4 Residual analysis

Analysis of model residuals can provide some insight into model performance beyond validation error. Ideally, the residuals have a mean of zero and constant variance (homoskedasticity), are independent (no autocorrelation) and are normally distributed. The model can be used for prediction regardless, but failing to meet these assumptions can make statistical inference tests biased, indicate underlying problems in model specification or lead to bias in coefficient estimates. The residuals of $u_2(t)$ clearly exhibit both heteroskedasticity (increasing variance at higher values) and autocorrelation. This is not surprising since the model is simplified and the effect of one or several omitted variables and non-linearity is included in the error term. In general, the nature of the residuals encourages more advanced modelling attempts specifically suited for time-series analysis, such as including endogenous variables, and non-linear modelling.

3.5 Uncertainties

The model developed in the present study uses a vertical board made of high-density pine sapwood as a proxy for a sheltered reference. This could have been avoided by simulating the sheltered reference numerically or by using training data from a different experiment. However, these alternatives would likely introduce other types of issues via numerical or experimental bias. For future testing of moisture performance in outdoor conditions, we would recommend including sheltered specimens in the experimental design. Sheltered references enable the increased moisture content due to wetting, which is the critical component for durability applications, to be isolated and analysed.

Accurate resistance-type measurements can be obtained, given adequate calibration, in the hygroscopic range. Beyond this range, the increased electrical resistance from increasing moisture content is subtle. With detailed calibration, as done in the present dataset, measurements can be extended into the over-hygroscopic range. However, under varying climate temperature the accuracy of such measurements is questionable. In the lack of calibration in the over-hygroscopic range, measurements are normally censored at an upper value, approximately corresponding to cell-wall saturation. The characteristics of the measurements should be considered. In the present study, for example, it would be appropriate to reduce the estimated error when both predicted and measured values are above cell-wall-saturation. Specific methods should be considered for censored data.

CONCLUSIONS

This study set out to describe the measured variation in moisture of softwood from weather data, through multiple linear regression. For model specification, we used signal decomposition and ridge regularisation. The results show the following:

- Wood moisture content of rain exposed softwood could be modelled through superposition of sheltered wood and a rain induced term.
- Through this process, measurements from different types of softwood (here: Norway spruce and Scots pine sapwood), which represent very different permeability to water, could be described with reasonable agreement.
- Error (residual) diagnostics showed problems with heteroscedacity and systematic error, likely stemming from omitted variables and non-linear effects, and possibly structural changes to the wood specimens over time.

While the model developed in this study was not able to exhaustively describe all features of the measured data, the fact that a simplistic model was able to describe much of the variance certainly encourages further analysis with more sophisticated methods. In future research, we will attempt to improve performance through more advanced non-linear and/or autoregressive modelling techniques. In addition, structured distributed lag models will be tested for more robust coefficient estimates.

ACKNOWLEDGEMENT

The work was funded by Formas (Swedish research council for sustainable development) [2021-02053]. Weather data was, in part, provided by the Institute of Meteorology and Climatology, Leibniz University Hannover. Holger Schilke is acknowledged for providing the data.

4. REFERENCES

Brandstätter, F, Kalbe, K, Autengruber, M, Lukacevic, M, Kalamees, T, Ruus, A, Annuk, A, Füssl, J (2023): Numerical simulation of CLT moisture uptake and dry-out following water infiltration through end-grain surfaces. *Journal of Building Engineering* **80**, 108097.

Brennan, G, Pitcher, J (1995): Equilibrium moisture content variations of timbers commonly used in Western Australia. *CALMScience* **2**(1), 25–50.

Brischke, C, Bayerbach, R, Rapp, A O (2006): Decay-influencing factors: A basis for service life prediction of wood and wood-based products. *Wood Material Science and Engineering* **1**, 91–107.
Brischke, C, Meyer-Veltrup, L (2016): Modelling timber decay caused by brown rot fungi. *Materials and Structures* **49**, 3281–3291.

Brischke, C, Rapp, A O (2008): Influence of wood moisture content and wood temperature on fungal decay in the field: observations in different micro-climates. *Wood Science and Technology* **42**, 663–677.

Buck, D, Wallentén, P, Sehlstedt-Persson, M, Öhman, M (2023): Moisture-and mould-resistance: multi-modal modelling leveraging X-ray tomography in edge-sealed cross-laminated timber. *Materials & Design* **230**, 111967.

De Windt, I, Li, W, Van den Bulcke, J, Van Acker, J (2018): Classification of uncoated plywood based on moisture dynamics. *Construction and Building Materials* **158**, 814–822.

Emmerich, L, Brischke, C, Sievert, M, Schulz, M S, Jaeger, A C, Beulshausen, A, Humar, M (2020): Predicting the outdoor moisture performance of wood based on laboratory indicators. *Forests* **11**, 1001.

EN-252 (2015): Field test method for determining the relative protective effectiveness of a wood preservative in ground contact.

Glass, S V, Boardman, C R, Farkas, N, Lazarcik, E Q, Zelinka, S L (2023): Improved engineering model for water absorption in softwoods. *Science and Technology for the Built Environment* 1–15.

Glass, S V, Zelinka, S L (2021): Moisture relations and physical properties of wood. Wood handbook: wood as an engineering material. Madison, WI.

Hosseini, S H, Niklewski, J, van Niekerk, P B (2023): Moisture prediction of timber for durability applications using data-driven modelling. In: *World Conference on Timber Engineering (WCTE 2023)*. Oslo, Norway, pp. 3808–3815.

Hukka, A (1999): The Effective Diffusion Coefficient and Mass Transfer Coefficient of Nordic Softwoods as Calculated from Direct Drying Experiments. *Holzforschung* **53**, 534–540.

Isaksson, T, Thelandersson, S (2013): Experimental investigation on the effect of detail design on wood moisture content in outdoor above ground applications. *Building and Environment* **59**, 239–249.

Kalbe, K, Kalamees, T, Kukk, V, Ruus, A, Annuk, A (2022): Wetting circumstances, expected moisture content, and drying performance of CLT end-grain edges based on field measurements and laboratory analysis. *Building and Environment* **221**, 109245.

Klement, I, Vilkovskỳ, P, Vilkovská, T (2021): The influence of wood moisture content on the processes of freezing and heating. *Applied Sciences* **11**, 6099.

Meyer, L, Brischke, C, Preston, A (2016): Testing the durability of timber above ground: a review on methodology. *Wood Material Science & Engineering* **11**, 283–304.

Meyer-Veltrup, L, Brischke, C, Alfredsen, G, Humar, M, Flæte, P O, Isaksson, T, Brelid, P L, Westin, M, Jermer, J, (2017a): The combined effect of wetting ability and durability on outdoor performance of wood: development and verification of a new prediction approach. *Wood Science and Technology* **51**, 615–637.

Meyer-Veltrup, L, Brischke, C, Källander, B (2017b): Testing the durability of timber above ground: evaluation of different test methods. *European Journal of Wood and Wood Products* **75**, 291–304.

Muñoz Sabater, J (2019): ERA5-Land hourly data from 1981 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS) [data set].

Niklewski, J (2018): Durability of timber members: moisture conditions and service life assessment of bridge detailing. PhD Thesis. Lund University.

Niklewski, J, Brischke, C, Frühwald Hansson, E, Meyer-Veltrup, L (2018): Moisture behavior of weathered wood surfaces during cyclic wetting: measurements and modeling. *Wood Science and Technology* **52**, 1431-1450.

Niklewski, J, Fredriksson, M (2021): The effects of joints on the moisture behaviour of rain exposed wood: a numerical study with experimental validation. *Wood Material Science & Engineering* **16**, 1–11.

Niklewski, J, Fredriksson, M, Isaksson, T (2016): Moisture content prediction of rain-exposed wood: Test and evaluation of a simple numerical model for durability applications. *Building and Environment* **97**, 126–136.

Van den Bulcke, J, Van Acker, J, De Smet, J (2009): An experimental set-up for real-time continuous moisture measurements of plywood exposed to outdoor climate. *Building and Environment* **44**, 2368–2377.

Wainer, J, Cawley, G (2021): Nested cross-validation when selecting classifiers is overzealous for most practical applications. *Expert Systems with Applications* **182**, 115222.