

Article

Prediction of Damage Accumulation Effect of Wood Structural Members under Long-Term Service: A Machine Learning Approach

Zheng Li ^{1,2}, Duo Tao ¹, Mengwei Li ³, Zhan Shu ¹, Songshi Jing ⁴, Minjuan He ^{1,2} and Peng Qi ⁴,*

- ¹ Department of Structural Engineering, Tongji University, Shanghai 200092, China; zhengli@tongji.edu.cn (Z.L.); timbertaoduo@163.com (D.T.); shuzhan@tongji.edu.cn (Z.S.); hemj@tongji.edu.cn (M.H.)
- ² Key Laboratory of Performance Evolution and Control for Engineering Structures of Ministry of Education, Tongji University, Shanghai 200092, China
- ³ Beijing Deep Singularity Technology Co., Ltd., Beijing 100086, China; limengwei@deepsingularity.com
- ⁴ Department of Control Science & Engineering, Tongji University, Shanghai 201804, China; Jingsongshi1995@gmail.com
- * Correspondence: pqi@tongji.edu.cn

Received: 26 January 2019; Accepted: 11 April 2019; Published: 16 April 2019



Abstract: It is well known that wood structural members can stand a relatively heavy load in the short term but will gradually get weaker if the load is applied for a longer period. This phenomenon is caused by the damage accumulation effect in wood and should be appropriately considered during the design of timber structures. Although various formulation methods (also known as classical models) have been proposed to evaluate the damage accumulation effect in wood, the calibration of model parameters is very time-consuming. Our work proposes a novel method to deal with the damage accumulation effect in wood that involves the application of machine learning algorithms. The proposed algorithm considers a multi-objective optimization process with a combination of goodness-of-fit and complexity. Long-term experimental data of typical wood species are used for developing the machine learning based damage accumulation model. Compared with existing pre-formulated models, our model managed to reduce the complexity of the model structure and give sufficiently accurate and unbiased predictions. This study aims to provide a novel tool for evaluating the damage accumulation in wood structural members, and the proposed model can further support the life-cycle performance assessment of timber structures under long-term service scenarios.

Keywords: timber structures; damage accumulation model; machine learning; multi-objective optimization; long-term experiment

1. Introduction

When subjected to long-term loading at high-stress ratios, many materials show a strength reducing effect. Wood is also inevitably influenced by this strength reducing effect during long-term service. This phenomenon is caused by the damage accumulation of wood that is related to the development of internal defects of the material (e.g., initial cracks [1,2], minor splitting [3], etc.). Shu et al. [4] has proposed the analytical approach of the short-term performance of wood members with those internal defects. However, to avoid unwanted structure failure and the consequential economic losses, the damage accumulation should be appropriately considered during the design of timber structures, thus an accurate model is required to make predictions of the strength reducing effect of wood [5]. A few models have been proposed to represent the strength reducing effect of



wood, such as Gerhard's model [6], Barrett and Foschi's model [7], and Foschi and Yao's model [8]. However, all the existing models are pre-formulated and based on the empirical generalizations and simplifications, where the unknown coefficients are obtained by a regression process. Taking the Foschi and Yao's Model (Canadian model) for example, the damage model contains four independent parameters, which are actually equivalent to eight parameters considering the mean value and the standard deviation of each parameter. All the parameters including the model parameters and the short-term performance parameters of wood are modelled as independent lognormal variables, and a non-linear minimization procedure related to the quasi-newton method or genetic algorithm is used to determine the distribution characteristics for each parameter. In order to prevent the interference of the local optimal value, it is often ensured that the global optimal solution is obtained by the combination of continuous optimization of the algorithm and comparison of various algorithms. However, such a process is always time-consuming and difficult to calibrate. This indicates that the damage accumulation of wood has a strong non-linear characteristic and can hardly be interpreted based on existing physical theories or knowledge. Although relatively effective approximations have been made by the classic models, all of them have common problems, such as the model coefficients are too complicated to be interpreted by a mathematical formula, and a fixed expression of formula may never be able to take all those factors into consideration (e.g., wood species, loading history, service conditions, etc.). The limitations of classic models are mainly reflected in the different model parameters for different wood species, which means that long-term experiments on each wood species would be conducted to investigate the long-term performances of different wood species. In addition, if there is no actual wood species corresponding to the parameters to be studied, such as the value and the coefficient of variation of the short-term strength, then it is impossible to determine the parameters of the classic model itself, which would seriously affect the development of subsequent research. Even when the long-term performance of only one wood species is investigated, it is difficult to balance the calculation efficiency and calculation accuracy by the classic mathematical models. Fortunately, with the blooming computer science in the past decade, a more sophisticated model can be generated with a large number of experimental data through the machine learning technique.

Artificial neural network (ANN) is an efficient tool in the realm of machine learning. The reason we seek to employ deep neural network in the prediction of wood damage is that it possesses two essential features that may help us to solve the confronting problems. One is the universal function approximation (UFA) [9]. Theoretically, a forward-single-hidden-layer structure containing a finite number of neurons can approximate any continuous functions, thus any physical processes can be represented by a certain type of network structure. The other feature is the way how neural networks are trained and the way they act. This allows to discover the unknown pattern by directly feeding inputs and labels to the networks, thus saving much laborious work to manually define these features. After the training process, the expected pattern will be grasped and the network should be readily generalized to other similar cases. Such acting manner provides us with a convenient and effective way of adding new input factors into consideration without much work compared with the conventional formulation method (i.e., classical models). As a consequence, the machine learning method by ANN can be used to find the relationship between influencing factors and results for a certain number of experimental data without worrying about the inherent mechanism.

In fact, machine learning has shown its power in many research fields. Except for the most commonly discussed topics on computer vision [10] and speech recognition [11], machine learning has been used in solving complex civil engineering problems. Lu et al. [12] used a double parallel model to forecast underground buried pipeline damage in MATLAB. Lautour and Omenzetter [13] presented a general method to predict seismic-induced damage via ANN describing both structure and ground motion with classical finite element model as comparisons. Machine learning method was used for the prediction of concrete creep in long-term service [14] and for durability and service life assessment of reinforced concrete structures [15]. More recently, machine learning was also introduced into building energy modeling [16] and safety assessment during constructions [17–19].

Up to the present, there is no relevant research on wood damage prediction using machine learning method. This research seeks to train a neural network model that takes material property, loading quantity, duration effect as well as temporal creep factors as input variables, expecting a damage accumulation and probability forecast as output for Hemlock lumber. Such end-to-end training procedure can give damage rate prediction and requires neither prior knowledge nor empirical formulation used in the classical models. It should be noted that the proposed machine learning model in this paper focuses on the damage accumulation of wood and final failure probability of the wood structural members, while environmental factors like temperature and humidity may be added into the model in future research.

2. Methodology

2.1. Background

The fundamental concept of ANN is based on the understanding of the way how biological neurons act and operate. Tens of thousands of cells are connected, working concurrently or sequentially through a series of superposition and activation to transform initial simple signals to a complex pattern. Such a unique mechanism unleashes its superiority of self-learning, mapping, and nonlinear functional approximations. This mechanism can capture complex interactions among input/output variables in the system without any prior knowledge about the nature of these interactions, which makes ANN exceptionally a good tool for regression, especially when used for pattern recognition and function estimation.

2.2. Algorithms

Steps of training a neural network model can be broken down into the following steps: (1) weights initialization, (2) forward propagation, (3) loss computation, and (4) back propagation algorithm that calibrates the weights. This forward-backward process will be repeated until convergence or the maximum number of iterations is reached.

To start with, Gaussian distribution N (0, 0.0001) was used to initialize the network. The initialization scheme gave the weights with a slightly different value from zero that can substantially fasten convergence, as indicated by a former study by Glorot et al. [20].

Forward propagation is one of the key steps in the neural network where simple signals are superposed and activated. As shown in Figure 1, neurons are hooked together forming a fully connected network. Neurons labeled "+1" is called bias units. As can be seen in this figure, layers L1, L2 and L3 denote the input layer, hidden layer and output layer, respectively. Moreover, x_i denotes the *i*th element in an input vector, $a_i^{(l)}$ is the activation on neuron *i* of the layer *l*, $h_{w,b}(x)$ is the final output of the network given an input vector *x* with the weight matrix of *W* and the bias vector of *b*. Each neuron can be regarded as a computational unit that takes in the input vector *x*, performs the weighted sum of inputs, and then activates the weighted sum.



Figure 1. Illustration of a fully connected network.

Specifically, for a single unit structure, the total weighted sum of inputs to neuron *i* in layer l + 1 ($Z_i^{(l+1)}$) can be expressed as:

$$Z_i^{(l+1)} = \sum_{j=1}^n W_{ij}^{(l)} x_j + b_i^{(l)}$$
(1)

where $W_{ij}^{(l)}$ denotes the weight between input *j* in layer *l* and neuron *i* in layer *l* + 1; x_j denotes the *j*th element in an input vector with *n* elements; $b_i^{(l)}$ denotes the bias in layer *l*.

As for a supervised problem with a group labeled data $(x^{(i)}, y^{(i)})$, when $h_{w,b}(x)$ denotes the non-linear hypotheses given by neural networks under circumstances of parameter (W, b), and function f performs a specific element-wise activation that yields activated vector a, the compact notation of forward propagation depicting all neurons at once using matrix operations can be written as:

$$z^{(2)} = W^{(1)}x + b^{(1)}$$

$$a^{(2)} = f(z^{(2)})$$

$$z^{(3)} = W^{(2)}a^{(2)} + b^{(2)}$$

$$h_{w,b}(x) = a^{(3)} = f(z^{(3)})$$
(2)

Then the loss *J* can be calculated as:

$$J = \sum_{1}^{N} \left(Y - h_{w,b}(x) \right)^2 / N$$
(3)

where $y^{(i)}$ denotes the ground truth label for the *i*th sample, and *N* is the number of samples. The loss function is given by the Mean-Square-Error (MSE), and in actual use of back propagation, an additional 1/2 is multiplied upon this form to take derivative more easily. Gradients are computed by back propagation method, and the minibatch stochastic gradient descent (SGD) with momentum is implemented to calibrate weights over epochs. The SGD algorithm can be described as:

$$w_{t} = w_{t-1} - \eta \times \nabla_{w} J(w; x^{(i;i+n)}; y^{(i;i+n)}) v_{t} = \gamma \times v_{t-1} + \eta \nabla_{w} J(w) w_{t} = w_{t-1} - v_{t}$$
(4)

where η and γ stand for the learning rate and the ratio of momentum, respectively. Specifically, these two parameters are set as 0.01 and 0.5 when the weight decay parameter is taken as 1×10^{-5} in this study. w_{t-1} and w_t are the weight for the previous and current time steps, respectively. This algorithm keeps working over iterations of training epochs until convergence or the maximum iteration limit is reached.

3. Machine Learning Based Damage Accumulation Model

Wood is stronger under loads of short-term duration and becomes weaker if it is subjected to a long-term, sustained load. This phenomenon results from the time-dependent increase of strain called duration of load (DOL). Since the strength properties of wood are significantly influenced by the DOL effect, design codes for timber structures, such as the Canadian code CSA O86 [21] and the Chinese code GB 50005 [22], specify DOL strength reduction factors during the design of timber structures. It should be noted that DOL effect is an inherent material characteristic of wood instead of a biological process over time. For the long-term experiment under constant loads with a sample size over 300, the primary relationship between the time to failure and the stress ratio (i.e., the ratio of the applied stress over the short-term strength of wood) can be revealed. A damage accumulation model is in need to quantify the DOL effect and to estimate the residual strength of wood under different loading conditions. When calculating the damage state of wood for any loading histories, an approximate load model is adopted, which subdivides the load history into several time intervals with constant applied stress levels. However, complex physical mechanisms and possible chemical interactions have made it

difficult to accurately predict the DOL effect using traditional engineering formulas. To address these challenges, a machine learning method is applied to develop a practical nonlinear model for predicting the time-dependent DOL effect of wood in this study.

3.1. Long-Term Experimental Data

The database for this study consisted of five groups of long-term experimental data. The long-term experiments were conducted at FPInnovations Canada, which is a non-profit member organization that carries out scientific research and technology transfer for the Canadian forest industry. The time to failure obtained for both constant stress levels and the controlled short-term tests were ranked, and the corresponding cumulative probability distributions are plotted as shown in Figure 2. Two stress levels were chosen as the long-term constant loading levels, which is 4500 psi (31 MPa) and 3000 psi (21 MPa) corresponding to the 20th-percentile and 5th-percentile of the short-term strength distribution of the lumber, respectively. At the 4500 psi level, for example, approximately 20% of the specimens failed during the loading process, and the cumulative distribution matches quite well with that from the controlled short-term tests. Furthermore, this figure shows that a significant upward departure from the initial trend occurs from the third month with the applied constant stress levels, and such phenomenon is possibly due to the positive-skewed distributions of the influence factors for the damage accumulation model.



Figure 2. Short-term and long-term experimental results.

Although it would be ideal to conduct the experiments as long as possible, most experimental groups were unloaded after three years due to a realistic compromise. The surviving specimens were tested for failure under ramp loading to assess the reduction in short-term strength induced by the period of constant loading. This process is illustrated in Figure 3 where the strength of the control curve is shown versus cumulative probability with normalized rank. The short-term test curve is shown as a solid line while the dotted red line represents the results from a long-term test with a constant load level. It can be observed that the residual strength of the specimens reduced due to the DOL effect. For the survived pieces that subsequently failed in a short term test, the weakest pieces in that group did not match their counterparts on the control curve but that involved only three to five pieces. After that, the residual strength of long-term specimens would match up with its counterpart in the control sample.



Figure 3. The process of assessing strength reduction.

The whole database is split into a training dataset and a testing dataset for training and evaluating the network with a ratio of 0.8 to 0.2. The training dataset is randomly shuffled to break down possible sequential dependencies. In general, all inputs need to be dimensionless and standardized to analogous distribution to reduce the difficulty of training. A Z-score function that achieves a deviation of 1 and a mean of 0 can satisfy these requirements and was used in this study, and Tanh activation functions [23] were employed to all hidden layers within the network.

3.2. Model Implementation

The input variables used in the machine learning method were selected according to the selection of parameters in the model proposed by Foschi et al. [8]. The proposed damage accumulation model considers the effect of the diversified wood species that are also treated as network input. After investigating into the inputs form and the condition of long-term experimental data, the selected input variables are as follows: distribution of short-term strength (i.e., wood property that consists of mean value and standard deviation), current applied stress, the variation trend of applied stress, logged load duration, and time since initial loading (i.e., logged hour). All input series and output labels were normalized properly as aforementioned. The structure of the model and an illustration of load-relevant inputs (1, 2, 3) and output are shown in Figure 4. The left column shows the input concerns, including wood property, loading status and time-relevant variants. When all these data are fed into our neural network model, it is expected that the proposed model will be able to generate the increment of damage percentage as output signals through the fully connected network.



Figure 4. Diagrammatic sketch of input and output on the damage process prediction network model.

One possible understanding of the parameter calibration process in our model is that the model tries to reveal the essence behind the random phenomenon of creep in wood. Such training process of parameter calibrating actually establishes a potentially deterministic relationship between a given input and the target, which would reflect how variables (e.g., damage accumulation) are interdependent (e.g., material properties, load duration, etc.) and attribute it to a causal relationship.

It is commonly acknowledged after 2015's ImageNet competition that a deeper structure will have better performance under mild assumption that the training data is of good quality (i.e., the data should be sufficient and well distributed). In this study, the network consisted of 10 hidden layers except for the essential input layer and the output layer, and each layer in the network was assigned with an adjustable scale factor (namely the number of neurons) specifying layer width for later evaluation. One hidden layer in the middle was set to have a scale factor twice the value of other scale factors. The hidden layers were activated by TanH, and the output layer was activated by SoftPlus [13] considering label value and the damage process value should be in the range of 0 to 1. Such nonlinear activation functions allowed the networks to compute nontrivial problems using only a small number of nodes. Dropout was exploited with a random dropout rate of p = 50%. This was enlightened by the design of networks as presented in Reference [10]. Dropout is a random process of disabling neurons in a layer with a chance of p, so the structure relies less on the output of the nodes in the previous layer. This is a well-known method for regularization to reduce over-fitting. However, dropout was not used to test time prediction in this study as they will undermine the consistency over runs.

3.3. Optimum Model Selection

The optimum model is expected to have: (1) good performance and (2) efficient model structure with proper parameters. Therefore, our final machine learning based damage accumulation model (i.e., M-D model) should be able to deal with the complexity of training parameters and to provide accurate predictions. The performance of the proposed method was evaluated via the prediction accuracy using the coefficient of determination (*R*). The model complexity was evaluated via the number of layers or nodes/neurons (denoted as scaling factor in this study). Meanwhile, for the model-selection referred to hyper-parameter tuning, cross-validation was taken to decide the best training parameters such as minibatch size and learning rate for momentum SGD method.

It should be noted that all the training parameters were investigated independently during the analytical process. Figure 5 shows the training process during experiments of cross-validation. The trace of training historical loss is recorded on the left and the current best is plotted on the right. The "X-axis" denotes the logged number of the epoch (i.e., each epoch iterates over the whole dataset). During the model selection experiments, the maximum epoch number was set to 10,000. In Figure 5, subplot a) in an upper row shows the training process performance when evaluating the variation of learning rate effect of the momentum SGD algorithm, whilst both the fixed minibatch size and structure scale factor are set to 25. It is noted that when the learning rate is set to be lower than 1×10^{-4} , the optimization speed is relatively low as indicated by the orange and the red line. By setting the learning rate higher than 1×10^{-2} , it may result in a strong oscillation of the learning curve as indicated by the blue line. Therefore, the learning rate of 1×10^{-3} was a reasonable choice towards the best model training result in our specific case.



Figure 5. Model-selection: (**a**) algorithm learning rate; (**b**) training data minibatch size; (**c**) model structure scale factor.

Similarly, the minibatch and structure scale factor have also been studied, as shown in Figure 5. From subplot b), it is noted that the loss drops faster for a smaller size minibatch. However, the minibatch with the size from 5 to 15 shows a bumpier result in the cost function. In terms of the maximum number of epochs, the training process oscillates with large noise and even fails to converge. The reason regarding the noise is that there were some hard samples which caused gradient oscillations to the cost function. We can see that the larger batch size is, the more it will behave like the full-batch method with a smoother learning curve. Yet learning with a bigger batch is slower and may be quite inefficient when dealing with big training data. Another important issue to be noticed is that there's an obvious local minimum region in the learning curve. When the loss decreases to about 0.5, a plateau prevents the demanding model to find a deeper relation to describe the material damage. Fortunately, the yellow line with a minibatch size of 25 seems to have a great potential of skipping such a local minimum. Therefore, compromising over training stability, efficiency and effectiveness, the minibatch size of 25 is a proper choice in this study.

The network scale factor was evaluated to decide the network size. With experience within structural engineering, a model giving the coefficient of determination *R* greater than 0.95 will satisfy the lowest requirement for the DOL effect prediction. We gradually increased the number of nodes in each layer until the model behavior met such requirement. This is a reverse-pruning method with the pruning techniques to trim the network size. It was operated by removing nodes from the network during the training by identifying and eliminating unnecessary/redundant nodes that would not noticeably affect the performance of the network. Figure 5c demonstrates that the ability of network is increased when the network size (scale factor) with the scale factor larger than 5, and the model performance achieves a threshold of 0.95. The growing structure of network increases the computational spending of the training process; thus, the size of the net is expected to be as simple as possible in terms of efficiency. In the case where the scale factor is 25, the network can avoid trapping into a local minimum. Meanwhile, the curve shows an expected downward trend and yields an error loss of less than 0.02.

3.4. Cross-Validation of the Model

In order to understand the stability of the model, explore the generalization ability of the model and evaluate the accuracy of the model prediction in practice, the cross-validation method was used to verify the model performance and to serve as the basis for the final model selection. In consideration of the small data set, 50 times of random selection were adopted to examine whether the model correctly learned most of the patterns from the data and was not overly disturbed by noise in the data. The results showed that the deviation and variance of the model were both kept at a lower value.

Fifty groups of partition experiments were conducted to access the comprehensive performance of different training processes on different data sets through a certain number of iterations of training. The analysis results are shown in Figure 6. It is noted from Figure 6a that due to the different division of the training set and the test set, the loss value of the objective function shows some differences in the first 15 iterations of training. However, the loss function gradually gets converged as the model training continues. Specifically, the error prediction loss of each group of trainings on the test set is shown in Figure 6b (each point in the box diagram represents an experiment): the mean loss is 0.05923, the standard deviation is 0.01166, and the variance is 0.00014. It is believed that the final effects were acceptable with satisfactory performance of the parametric training. Moreover, the validation results showed that the verification set and the training set of the data satisfy the independent and identical distribution. The final training effect of the model is independent of the number of divisions.



Figure 6. Cross-validation analysis: (**a**) validation loss over 50 times of training; (**b**) model performance at the last cross-validation epoch.

3.5. Optimum Model Performance

After investigating several important parameters in model selection, the final M-D model was trained with two phases with parameter settings listed in Table 1. The training process of the final optimum model was done first on Phase 1 that begins from initialization with a relatively bigger learning rate of 1×10^{-3} and momentum of 0.5 and then switched to Phase 2 with a smaller learning rate of 1×10^{-5} and momentum of 0, respectively.

Phase Label	Algorithm	Learning Rate	Weight Decay	Momentum	Minibatch Size	Scale Factor	Epoch
Phase 1	SGD	0.001	1×10^{-5}	0.5	25	25	10,000
Phase 2	SGD	0.0001	1×10^{-5}	0	25	25	20,000

Table 1. Configurations for training optimum model.

The final performance on training and testing data is illustrated in Figure 7. The different case parameters in the figure represent the specimens with different strength characteristics and tested under different loading. For example, the case "2000 (6936, 0.41)" means that the mean value and coefficient of variation of the short-term strength for the specimen is 6936 psi and 0.41, respectively, and the specimen is loaded under 2000 psi. As can be seen in Figure 7a, the colorful scatters represent the predicted damage increments given by our model over time with respect to different cases. The gray and red points in the right plot represent the final damage accumulation procedure over the whole experimental period. At last, the performance of the model, with variant inputs such as wood species and loads, to predict lumber load damage reaches an *R*-value of 99% based on the test dataset.

Moreover, three experiments were conducted to further demonstrate the model's ability to generalize other cases as shown in Figure 8.

In Figure 8a,b are new load cases with a load of 3500 psi and an initial strength of N (6936, 0.41) and N (7092, 0.2), where N (μ , σ^2) stands for Gaussian Distribution with the mean μ and the derivation of σ , and c) is a new material case with a load of 3000 psi and an initial strength of N (7014, 0.305). The corresponding results are shown in Figure 8 with process damage on the left. All these data are modified from presently available data. As illustrated in Figure 8, for an unprecedented load of 3500 psi, the model performs excellently with both types of lumber. For example, in case of b), the effect of an unprecedented load of 3500 psi on the wood of initial strength N (7092, 0.2) is within the range of impact caused by a load of 3000 and 4000 psi. Concurrently, regarding real experiments of psi 3000 and N (7092, 0.2) as a reference (as marked), it is found that a decrease in material strength to N (7014, 0.305) would deteriorate the load tolerance of wood, which will also be true in the real cases. These facts demonstrate that the M-D model has acquired a few meaningful aspects to predict wood damage behavior.



Figure 7. Best M-D model performance on the whole dataset: (**a**) learning output for damage increments; (**b**) cumulative damage result given by M-D model.



Figure 8. Experiments on model capacity of generalization: (**a**) new load case with a load of 3500 psi and an initial strength of N (6936, 0.41); (**b**) new load case with a load of 3500 psi and an initial strength of and N (7092, 0.2); and (**c**) new material case with a load of 3000 psi and an initial strength of N (7014, 0.305).

4. Comparisons with Conventional Models

The influencing factors of DOL effect in wood varies and sometimes are highly interactive, which makes it difficult to predict damage accumulation for different wood species with classic engineering models. Meanwhile, many meaningless yet important parameters remain to be calibrated through a whole set of fatigue procedure [16]. Whereas, with the state-of-the-art computational paradigm of the deep learning model, it is not necessary to consider the various source of inputs, complex physical or chemical interaction mechanism when making the lumber damage prediction. Compared with traditional methods, the ANN-based model proposed in this study can tolerate relatively inaccurate, noisy or incomplete data [10]. The proposed model is not susceptible to outliers due to its better filtering capability and better adaptability in most cases. Figure 9 shows the comparisons between the proposed M-D model and the two conventional models (Gerhard's model and Foschi and Yao's model). It is noted that Gerhard's model is not fitted to follow the experimental trends. Nevertheless, based on the proposed model, a rather precise result (R > 99%) can be achieved with the above-mentioned superiorities, and it is better or at least equivalent to the results provided by the commonly-used classical models.



Figure 9. Comparison of different prediction models.

Some factors including the environment temperature and humidity and the material density were still not considered in the machine learning model. In general, the material density of wood is considered to be proportional to its short-term strength. Moreover, it should be noted that many of the parameters shown in the formulation of the Foschi and Yao's model are of no clear physical meaning and just the mathematical parameters introduced to simply fit the complex long-term performance curve. The machine learning model only considers the physical meaning parameters so as to exclude interference from unrelated variables on the model mechanism.

5. Conclusions

Neural networks build interactions between input features and automatically generate the best-fit model. With the help of computer science techniques, this study proposed a novel machine learning method for complex civil engineering problem (i.e., time-dependent damage accumulation of wood).

To simultaneously optimize objectives of the best performance and the model complexity, model structure selection and parameter calibration experiments were carried out with consideration on goodness-of-fit. In terms of model design, learning rate and training batch size are important to be set properly so that weights optimization algorithm can overcome difficulties during the training process, and the scale of the model structure may directly limit the capacity of net learning pattern.

With the final optimum model setting indicated by model selection experiments, our prediction model achieves R score of 0.9938. Moreover, when this M-D prediction model is generalized to unprecedented load or over different wood property, it can still give out some meaningful indicative references on unseen cases. Besides, the proposed method avoids the fatigue process of parameter calibration, and meanwhile, it can take other potentially influential factors into consideration by just enlarging the dimension of input features. Though there are still some drawbacks for such machine learning model, such as the inability to consider all the wood species and any time interval, it is believed that the model has certain application prospects and provides a reference for subsequent research. This is due to that the experimental data used for the model training process has covered a wide range of wood species from the perspective of the mean value and coefficient of variation of the wood strength. In terms of wood structure design, the most potential application of such machine learning model is to obtain the load duration influence coefficient under different load conditions and different live/constant load ratios through reliability analysis. Such coefficient would make a direct influence on the conversion of wood properties from the standard value to the design value under the long-term effect of load, resulting in a more effective and economical design index of wood structures considering the whole life cycle. Moreover, the potential for predicting the long-term damage accumulation is mainly reflected in the ability to further study the damage process within each individual wood member.

This study tries to provide a tool for quantifying the damage accumulation effect of wood structural members and can further support the life-cycle performance assessment of timber structures on a larger scale.

Author Contributions: Conceptualization, Z.L., M.H. and P.Q.; Methodology, Z.L. and P.Q.; Validation, M.L. and P.Q.; Investigation, D.T. and Z.S.; Resources, Z.L. and M.H.; Data curation, M.L. and D.T.; Writing—Original draft preparation, Z.L., M.L. and D.T.; Writing—Review and Editing, P.Q. and S.J.; Visualization, M.L. and D.T.

Funding: The project received funding from the National Natural Science Foundation of China (Grant No. 51608376 and Grant No. 51878476), Shanghai Municipal Education Commission—Gaofeng Program (Grant No. 0200121005/049), Science and Technology Commission of Shanghai Municipality—Shanghai Sailing Program (Grant No. 17YF1420400), and Shanghai Education Development Foundation and Shanghai Municipal Education Commission—Chenguang Program (Grant No. 16CG20).

Acknowledgments: The authors gratefully acknowledge National Natural Science Foundation of China (Grant No. 51608376 and Grant No. 51878476), interdisciplinary research fund by Shanghai Municipal Education Commission—Gaofeng Program (Grant No. 0200121005/049), the Shanghai Sailing Program (17YF1420400) of Science and Technology Commission of Shanghai Municipality and the "Chenguang Program" (16CG20) of Shanghai Education Development Foundation and Shanghai Municipal Education Commission for supporting this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. He, M.; Zhang, J.; Li, Z. Influence of cracks on the mechanical performance of dowel type glulam bolted joints. *Constr. Build. Mater.* **2017**, *153*, 445–458. [CrossRef]
- 2. Zhang, J.; He, M.; Li, Z. Mechanical performance assessment of bolted glulam joints with local cracks. *J. Mater. Civ. Eng.* **2018**, *30*, 04018094. [CrossRef]
- 3. Shu, Z.; He, M.; Li, Z. Estimating inelastic drift demands for wood portal frame systems considering accumulated damages. *Struct. Des. Tall Spec. Build.* **2018**, 27, e1440. [CrossRef]
- 4. Shu, Z.; Li, Z.; Yu, X.; Zhang, J.; He, M. Rotational performance of glulam dowel type bolted joints: Experimental investigation and analytical approach. *Constr. Build. Mater.* **2016**, *19*. [CrossRef]
- 5. Booth, L.G. The strength testing of timber during the 17th and 18th centuries. J. Inst. Wood Sci. 1964, 13, 5–30.
- 6. Gerhards, C.C. Time-related effects on wood strength: A linear cumulative damage theory. *Wood Sci.* **1979**, *11*, 139–144.
- 7. Barrett, J.D.; Foschi, R.O. Duration of load and probability of failure in wood Part 1: Modelling creep rupture. *Can. J. Civil Eng.* **1978**, *5*, 505–514. [CrossRef]
- 8. Foschi, R.O.; Folz, B.; Yao, F. *Reliability-Based Design of Wood Structures, Structural Research Series*; Rep. NO. 34; Department of Civil Engineering, University of British Columbia: Vancouver, BC, Canada, 1989.

- 9. Balázs, C.C. Approximation with Artificial Neural Networks. Master's Thesis, Faculty of Sciences, Eötvös Loránd University, Budapest, Hungary, 2001.
- Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Image net classification with deep convolutional neural networks. In Proceedings of the International Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; Curran Associates Inc.: New York, NY, USA, 2012; pp. 1097–1105.
- Hinton, G.; Deng, L.; Yu, D.; Dahl, G.E.; Mohamed, A.; Jaitly, N.; Senior, A.; Vanhoucke, V.; Nguyen, P.; Sainath, T.N.; et al. Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups. *IEEE Signal Process. Mag.* 2012, *29*, 82–97. [CrossRef]
- 12. Lu, S.; Liu, T.; Chen, Y. Application of artificial neural network to modeling of pipeline damage prediction. *Int. Forum Inf. Technol. Appl.* **2009**, 499–502.
- 13. Lautour, O.R.D.; Omenzetter, P. Prediction of seismic-induced structural damage using artificial neural networks. *Eng. Struct.* 2009, *31*, 600–606. [CrossRef]
- 14. Gandomi, A.H.; Sajedi, S.; Kiani, B.; Huang, Q.D. Genetic programming for experimental big data mining: A case study on concrete creep formulation. *Autom. Constr.* **2016**, *70*, 89–97. [CrossRef]
- 15. Taffese, W.Z.; Sistonen, E. Machine learning for durability and service-life assessment of reinforced concrete structures: Recent advances and future directions. *Autom. Constr.* **2017**, *77*, 1–14. [CrossRef]
- 16. Naganathan, H.; Chong, W.O.; Chen, X. Building energy modeling (BEM) using clustering algorithms and semi-supervised machine learning approaches. *Autom. Constr.* **2016**, *72*, 187–194. [CrossRef]
- 17. Bloch, T.; Sacks, R. Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models. *Autom. Constr.* **2018**, *91*, 256–272. [CrossRef]
- Alwasel, A.; Sabet, A.; Nahangi, M.; Haas, C.T.; Abdel-Rahman, E. Identifying poses of safe and productive masons using machine learning. *Autom. Constr.* 2017, 84, 345–355. [CrossRef]
- 19. Poh, C.Q.X.; Ubeynarayana, C.U.; Goh, Y.M. Safety leading indicators for construction sites: A machine learning approach. *Autom. Constr.* **2018**, *93*, 375–386. [CrossRef]
- 20. Glorot, X.; Bengio, Y. Understanding the difficulty of training deep feedforward neural networks. *J. Mach. Learn. Res.* **2010**, *9*, 249–256.
- 21. Canadian Standards Association (CSA). *Engineering Design in Wood, CSA O86-09*; Canadian Standards Association (CSA): Mississauga, ON, Canada, 2009.
- 22. Standards Press of China. *Standard for Design of Timber Structures, GB/T50005-2017;* Standards Press of China: Beijing, China, 2017.
- 23. Weisstein, E.W. Hyperbolic Tangent; MathWorld: Champaign Country, IL, USA, 2008.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).