Sound & Vibration 2026, 60(2), 3869.
doi: https://doi.org/10.59400/sv3869

Article

Nonlinear damping identification using extended continuous wavelet transform and long-short term memory: Application to a spur gear pair system

Nourhaine Yousfi[1,*], Fatma Mejdoub[2], Ali Akrout[1], Lassaad Walha[1], Mohamed Haddar[1]

[1] Mechanics, Modelling and Production Laboratory (LA2MP), National Engineering School of Sfax, University of Sfax, Sfax 3038, Tunisia
[2] Laboratory of Electromechanical Systems (LASEM), National School of Engineers of Sfax, University of Sfax, Sfax 3038, Tunisia
*** Corresponding author:** Nourhaine Yousfi, nourhainegem@gmail.com

CITATION

Yousfi N, Mejdoub F, Akrout A, et al. Nonlinear damping identification using extended continuous wavelet transform and long-short term memory: Application to a spur gear pair system. Sound & Vibration. 2026; 60(2): 3869. https://doi.org/10.59400/sv3869

ARTICLE INFO

Received: 30 December 2025
Revised: 11 March 2026
Accepted: 16 March 2026
Available online: 27 March 2026

Abstract: Damping significantly affects the dynamic analysis of spur gear pair systems. The identified damping ratios may suffer from instability owing to many reasons, such as time-varying conditions and nonlinear effects. Long-short-term memory (LSTM) has been developed for damping model identification in systems with nonlinear behavior. However, owing to the poor quality of the input data, the identification results may not be reliable. In this regard, this study proposes a hybrid technique based on the Continuous Wavelet Transform (CWT) and LSTM methods. The major novelty of this study is the utilization of the time-frequency information provided by the CWT as the input data. Therefore, the LSTM network was fed these extracted features to enhance noise attenuation and improve the robustness and stability of nonlinear damping identification. Thus, the CWT technique is used as a preprocessing tool for the observed signals, enhancing the data quality by reducing the influence of noise. To verify the effectiveness of the proposed CWT–LSTM approach for damping identification, CWT representations of the simulated spur gear pair system response were used in a series of analyses. The numerical results indicate that the combined CWT–LSTM approach provides more reliable and accurate nonlinear damping estimation than the conventional LSTM model. This methodology has a strong potential for the accurate identification of damping in gear transmission systems.

Keywords: nonlinear damping; continuous wavelet transform (CWT); long short-term memory (LSTM); system stability; time-frequency features

1. Introduction

Damping estimation has a strong effect on engineering system dynamics because it profoundly impacts vibration level predictions, system stability, and the accuracy of dynamical models [1]. Incorrectly detected damping may induce considerable errors in the predictions of important dynamical phenomena, as proven by Ghahari et al. [2], who showed that damping model accuracy has a strong impact on system dynamics. Despite its importance, this estimation remains one of the most difficult tasks to perform accurately, particularly with respect to nonlinear and time-varying changes. Traditional methods for damping identification, such as modal analysis, often face limitations, including computational complexity, sensitivity to noise, and dependence on idealized system models [3]. Extensive research has been devoted to the direct identification of linear damping matrices, and several refined techniques have been developed. Among these, the Stabilized Layers Method (SLM) proposed by Lisitano

et al. [4] has proven to be robust and reliable, providing physically meaningful damping coefficients. However, the SLM method presented by Bonisoli et al. [5] is limited to linear equivalent models.

In practice, many mechanical systems exhibit nonlinear damping behavior caused by material hysteresis, friction at joints, or geometric effects [6]. Nonlinear damping plays a crucial role when the system experiences large oscillation amplitudes, and neglecting such effects can lead to significant modeling inaccuracies [7–9]. Therefore, recent research efforts have focused on developing parametric identification methods capable of separating linear and nonlinear contributions to damping, allowing for a clearer understanding of energy dissipation mechanisms. Chatterjee and Chintha [10] developed an identification approach based on harmonic probing and Volterra series to estimate cubic nonlinear damping in mechanical systems. Recently, Eberle [11] introduced a nonparametric and computationally efficient technique for characterizing nonlinear damping in single-degree-of-freedom systems. Han and Kinoshita [12] proposed a stochastic inverse approach for identifying nonlinear damping in oscillatory systems with constant stiffness.

Recently, novel machine learning algorithms, such as Long Short-Term Memory (LSTM) networks, have attracted growing attention in dynamic system identification, which may bring certain benefits to avoid the disadvantages of traditional methods [13]. Unlike traditional approaches, LSTM networks directly address the issue of damping characteristics from vibration measurements. Several studies have established that better data preprocessing is essential for enhancing the performance of the proposed LSTM model [14]. For instance, Yun et al. [15] proposed the integration of the frequency domain decomposition (FDD) approach with an LSTM network to determine the modal damping ratio. The effect of the instabilities on the FDD method was eliminated. It has been established in the literature that robust extraction of damping characteristics from noisy measurements can be achieved using LSTM network-based techniques, ensuring their real-life applications [16, 17]. Afridi et al. [18] proposed that LSTM models can estimate damping under different operating conditions. The potential of an Extended Kalman Filter (EKF) with an LSTM network was proposed by Kulkarni and Paul [19], where an accurate estimation of states can be acquired using the proposed approach. To enhance the LSTM technique, Zheng et al. conducted research [20] by incorporating Enhanced Frequency Domain Decomposition (EFDD) into the identification of damping variability in the early stage. Yun and Park [21] developed a hybrid technique for modal identification by incorporating the Extended Kalman Filter (EKF) technique and LSTM networks. Previous studies have indicated the possible accuracy of data estimation using the LSTM technique, despite the noise involved [22–24]. In gear system applications, the LSTM neural network approach is important for the dynamic modeling and fault prediction of gearboxes. Various studies have proven the efficiency and potential of the LSTM model approach for gear health evaluation and fault identification based on vibrations [25, 26]. For example, He et al. [27] proposed a hybrid model by inserting a Convolutional Block Attention Module (CBAM) into LSTM neural networks for gear fault identification in planetary gearboxes under constant operating conditions. In addition, Xia et al. [28] presented

the theory of Variational Mode Decomposition (VMD) and the LSTM model approach for gear health state classification under various operating conditions, such as variable speeds and interfering noise. Again, enhanced dependability for the fault identification approach was offered by the strategy of Cheng et al. [29] through the application of the LSTM model approach.

Although LSTM neural networks have been heavily utilized in other areas such as fault diagnosis and condition monitoring, it has been found that their effectiveness heavily relies on the accuracy and credibility of the input data [30–32]. For this purpose, Continuous Wavelet Transform (CWT)-based preprocessing was adopted in this study. Previous studies [33, 34] have revealed that wavelet transformations can also be efficiently utilized to transform raw time-domain signals into time-frequency signals. This further leads to an increased amount of information regarding the gear system dynamics. These time-frequency features were then used to train an LSTM neural network. By using both CWT preprocessing steps and LSTM neural networks, this model further improves the effectiveness and accuracy of damping identification, including nonlinearities and defects. A simulation test problem has been analyzed to test and verify the accuracy and efficacy of this novel technique using the novel CWT–LSTM technique.

The key contributions of this study are briefly summarized as follows: In Section 1, a dynamic gear system with time-varying stiffness and nonlinear damping is presented. In Section 2, a novel hybrid CWT–LSTM approach is developed for damping ratio identification based on vibration responses. In Sections 3 and 4, a systematic robustness test is conducted by applying the proposed method to various constant damping ratios and under varying torque conditions.

2. Methodology

2.1. Time–frequency feature extraction using CWT

The CWT analysis technique was used to provide the simulation data in the form of $x(t)$ in the time-frequency domain. This technique allows for a wide array of frequencies to obtain a more informative and detailed dataset. The formula for the CWT is stated as follows

$$W[n,s] = \frac{1}{s^{\frac{1}{2}}} \int_{-\infty}^{\infty} x[k]\omega * \left(\frac{k-n}{s}\right) dt \tag{1}$$

In this formulation, $x[k]$ denotes the input signal, and $\omega * \left(\frac{k-n}{s}\right)$ represents the wavelet function. the Ricker wavelet (Mexican hat) is employed due to its good localization in both time and frequency domains.

Parameter s corresponds to the frequency, and n defines the time. The term $\omega * \left(\frac{k-n}{s}\right)$ is the complex mother wavelet, and S is the scale parameter, which controls the analyzed frequency. The wavelet $\omega * \left(\frac{k-n}{s}\right)$ is localized at time $n = [t_1.....t_a]$ and the frequency range is $s = [s_1.....s_a]$ and translated throughout the simulated data. As a result, the CWT generates a two-dimensional representation, as expressed by the

following equation

$$wavelet_{output} = \begin{bmatrix} \Psi(s_1,t_1) & \Psi(s_1,t_2) & \dots & \Psi(s_1,t_a) \\ \Psi(s_2,t_1) & \Psi(s_2,t_2) & \dots & \Psi(s_2,t_a) \\ \dots & \dots & \dots & \dots \\ \Psi(s_a,t_1) & \Psi(s_a,t_2) & \dots & \Psi(s_M,t_a) \\ \end{bmatrix} \tag{2}$$

In this representation, each row corresponds to a specific scale (or frequency band). Each coefficient $\Psi(s_j,t_j)$ represents the contribution of frequency s_j at time t_j. The obtained components are then provided as inputs to the Long Short-Term Memory (LSTM) network as follows

$$LSTM_{input} = \Psi(s,t_1), \Psi(s,t_2).....\Psi(s,t_a) \tag{3}$$

2.2. Neural network framework based on LSTM method

An LSTM is composed of three main gates, which can be characterized as follows:

- **Forget gate (Ft)**: This gate controls the information removed from the cell state. It is calculated as

$$F_t = \sigma\left(U_F\left[H_{t-1}, X_t\right] + B_F\right), \tag{4}$$

  where σ corresponds to the sigmoid activation function, H_{t-1} is the hidden state from the previous time step, X_t is the current input vector, and U_F and B_F are the weight matrix and bias associated with the forget gate, respectively.
- **Input gate (J_t):** This gate decides which information should be added to the cell state and is computed as

$$J_t = \sigma(U_J\left[H_{t-1}, X_t\right] + B_J), \tag{5}$$

  where U_J is the weight matrix associated with the input gate, B_J is the corresponding bias term, H_{t-1} represents the previous hidden state, and X_t is the current input vector.
- **Output gate (O_t):** This gate controls the information to be output as the next hidden state. It is computed as

$$O_t = \sigma(U_o\left[H_{t-1}, X_t\right] + B_o), \tag{6}$$

  where U_o denotes the weight matrix of the output gate along with its corresponding bias B_o.

2.3. Damping estimation using a hybrid CWT–LSTM approach

The CWT-extracted features were fed into a long short-term memory (LSTM) network, which is well-suited for modeling temporal dependencies in sequential data. The LSTM consists of a hidden recurrent layer followed by a fully connected output layer with a softplus activation function to enforce the non-negativity of the predicted damping coefficient. The LSTM network was built with 60 units, followed by a layer

that outputs the estimated damping. To ensure reproducibility of the experiments, the key training parameters have been explicitly specified. The model was trained using the Root Mean Square Error (RMSE) as the loss function, with the Adam optimizer and a learning rate of 0.001. A batch size of 32 was used, and training was performed for up to 200 epochs. To prevent overfitting, dropout with a rate of 0.2 was applied to the hidden layers, and early stopping with a patience of 20 epochs was implemented. Additionally, 20% of the training data was set aside for validation. **Figure 1** presents the hybrid CWT–LSTM approach for estimating the damping. First, the continuous wavelet transform (CWT) is used for time-frequency feature extraction of the input signal, offering a refined description in terms of various frequency bands using the Ricker wavelet. The Ricker wavelet was selected as the mother wavelet because of its suitability for vibration signal analysis and its ability to represent transient features. Wavelet-based approaches have been widely used for extracting transient information from gear vibration signals and for fault detection in rotating machinery, as presented by Li [35]. The impulse-like shape of the Ricker wavelet makes it particularly appropriate for representing transient vibration components. The extracted features were then fed into a long short-term memory (LSTM) network, which is particularly effective for capturing temporal dependencies in sequential data. The LSTM network architecture was designed with gated memory cells and a fully connected output layer with a softplus activation function to guarantee non-negative damping estimates. The hybrid model is capable of effective damping estimation owing to the improved feature description offered by time-frequency analysis. After passing through the LSTM layer, the calculated predicted output was used to compute the loss function with the target output (output data) as part of the weight optimization. The loss function employed in this study is root mean square error (RMSE), as presented in **Figure 2**.

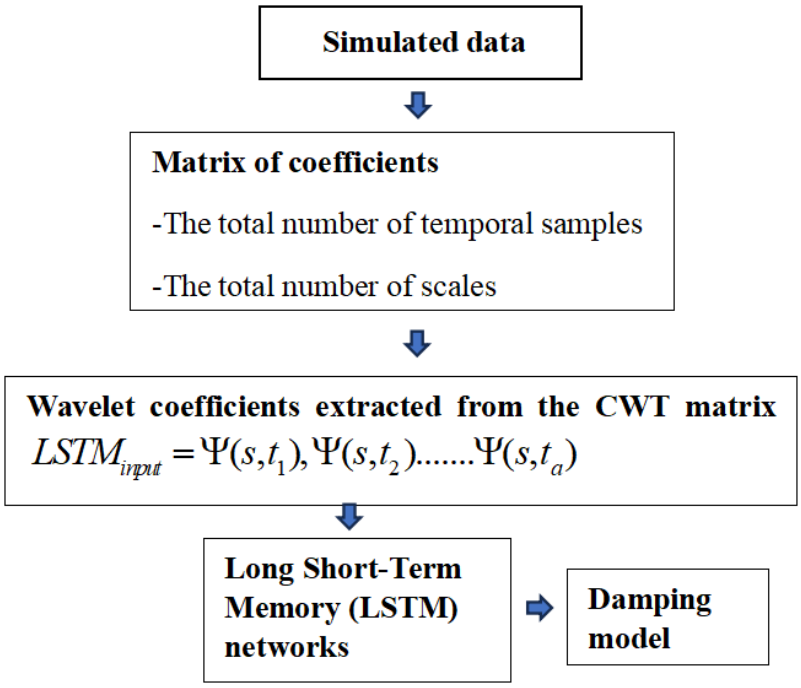

Figure 1. CWT–LSTM for Damping estimation algorithm.

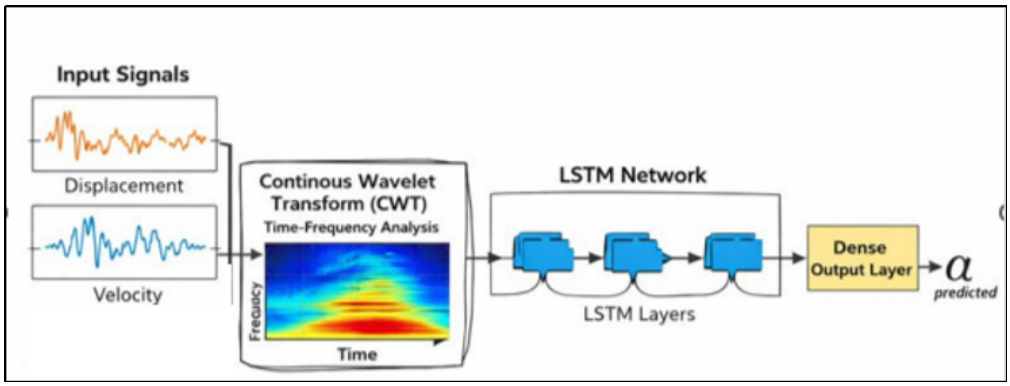

Figure 2. Damping estimation algorithm.

Root mean square error (RMSE) is calculated in the following equation,

$$RMSE = \sqrt{\frac{1}{2}\sum_{i=1}^{N}(\alpha_i - \widetilde{\alpha_i})^2} \tag{7}$$

α represents the target output and $\widetilde{\alpha}$ represents the LSTM network's predictions. The target output αcorresponds to the damping ratio predicted by CWT–LSTM. For comparison, a baseline LSTM model is also trained using raw time-domain signals as inputs, allowing quantitative evaluation of the benefit introduced by time–frequency preprocessing. For data generation, the dynamic responses of the system were integrated over the time interval of 0.05 s with 2,000 steps and transformed to input/output sequences by using a sliding window of 30 elements to produce approximately 1,900 elements for each run. The training and test data were divided according to their chronological order, with 80% of the data for training and 20% for testing to maintain the signal's time dependency. For validation, another set of data was randomly taken from the training data to monitor the training process and avoid overfitting. The weights of the neural networks are randomly initialized; however, to maintain the same random sequence for NumPy and TensorFlow libraries, random seeds are fixed for these libraries. The experiments are repeated multiple times to confirm the stability of the results. The results are reported for some of the runs with minimal variation between the repeated runs.

3. Numerical validation

3.1. Model of spur gear pair system

To assess the effectiveness of the proposed CWT–LSTM approach for damping identification, a series of analyses was conducted using the CWT representations of the simulated gear system response. The CWT–LSTM procedure outlined in Section 2 was applied to the gear system. Although the effectiveness of the proposed method has been demonstrated through numerical simulations, its applicability to real-world gear systems has not yet been experimentally validated. The dynamic model of the gear system is illustrated in **Figure 3**. The pinion and gear are modeled as rigid rotating bodies with rotational inertias I1 and I2, and angular displacements θ1(t) and θ2(t), respectively. The gear mesh interaction is represented by a time-varying mesh stiffness K(t) and damping coefficient C(t). The shafts are supported by rigid bearings,

and no bearing flexibility is considered. The system is excited by an external torque T1(t) applied to the pinion. This simplified torsional model describes the dynamic transmission error between the pinion and the gear.

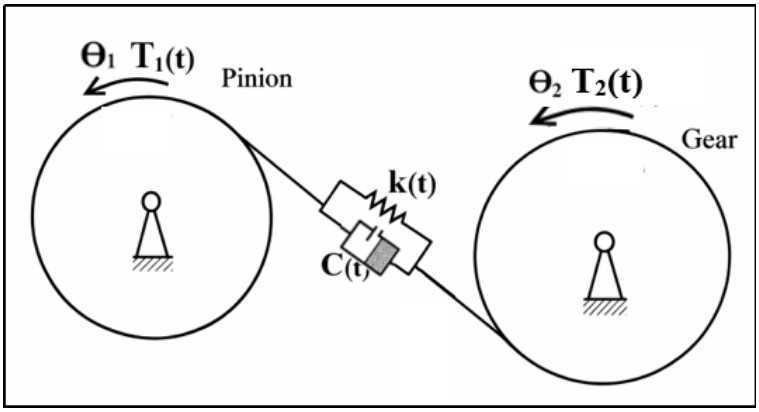

Figure 3. Dynamic gear model.

In this study, a single-stage spur gear system was used as the test case to apply the LSTM and CWT–LSTM techniques for damping analysis. All relevant mechanical parameters, including gear geometry, material properties, boundary conditions, and loading, are presented in **Table 1**. Although the present study focuses on a single-stage system, the methodology is general and can be extended to more complex multi-stage gear systems or other mechanical structures, as it only requires simulated data and their wavelet transforms as input. This approach allows for a rigorous demonstration of the effectiveness of the CWT and CWT–LSTM techniques in identifying linear and nonlinear damping behaviors.

Table 1. Parameters of spur pair-gear system.

Parameters	Pinion	Gear
Teeth numbers	20	40
Inertia moments (kg.m^2)	0.00026	0.0045
Base circle (m)	0.05	0.11
Stiffness (N/m)	Kmax = 4.10^8; Kmin = 2.10^8	
Module (m)	0.003	
Pressure angle (°)	20	
Contact ratio	1.6	
Teeth width (m)	0.023	
Torque T1 (N.m)	150	

The system's equation of motion can be expressed as follows

$$\ddot{\mathrm{x}} + \mathrm{C}\dot{\mathrm{x}} + \mathrm{K}(t)\mathrm{x} = \mathrm{f}. \tag{8}$$

The transmission error x(t) was obtained using the Runge–Kutta method. The mesh damping model is presented by C. In this study, we introduced a nonlinear damping model into the dynamic equation of motion. Unlike classical linear damping, which is proportional to velocity, the proposed model incorporates a nonlinear relationship between the dissipative force and the system's velocity. The nonlinear

damping is presented by the following equation

$$C = \alpha \text{sign}(\dot{x})|\dot{x}|^n, \tag{9}$$

where $\dot{x}$ is the velocity, α is the damping coefficient and *n* is the non-linearity exponent. Nonlinear damping is incorporated to the differential equation. This formulation improves the model performance and leads to a dynamic response that is more consistent with the time varying system conditions.

The time-varying mesh stiffness is illustrated in **Figure 4**. The horizontal axis represents time (s), while the vertical axis represents the mesh stiffness K(*t*) expressed in N/m. The stiffness varies periodically due to the continuous change in the number of teeth in contact during the meshing process. Higher stiffness values correspond to double-tooth contact, while lower values correspond to single-tooth contact.

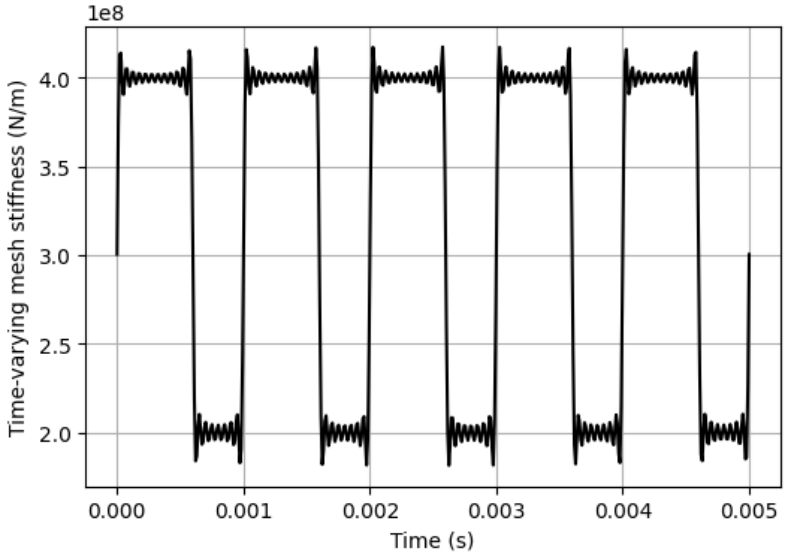

Figure 4. Time varying stiffness.

The mesh stiffness is modeled using a truncated Fourier series, as given [36]

$$K(t) = K_0 + \sum_{h=1}^{Nh} \left(K_h^c \cos(hw_m t) + K_h^s \sin(hw_m t)\right), \tag{10}$$

where K_h^c and K_h^s are the harmonic coefficients of the time varying stiffness. Nh is the number of harmonic sideband and w_m is the mesh frequency. The mesh stiffness of a spur gear pair varies cyclically with the mesh period, and Fourier series are well-suited for efficiently representing such periodic stiffness fluctuations. This approach provides closed-form analytical expressions that can be directly integrated into the system differential equations, enabling fast and stable simulations in both the time and frequency domains. The simulated data x is obtained by solving Equation (8) using the Runge-Kutta method. The CWT method was then applied to the simulated data to obtain the wavelet characteristics, which were used as the input of the CWT–LSTM technique. The steps of the LSTM and CWT–LSTM techniques were then applied to identify the damping coefficient α. The proposed model was developed and tested using the Python programming language. To ensure the validity of the

proposed method, additive zero-mean Gaussian white noise effects, gear eccentricity, and backlash effects were considered in the dynamic model. This helps verify the robustness and reliability of the proposed approach.

3.2. Damping ratio identified from the CWT–LSTM method

The simulated data were first processed using the Continuous Wavelet Transform to extract the time–frequency features. The simulated vibration signals were processed using the Continuous Wavelet Transform (CWT) to extract time–frequency features, as shown in **Figure 5**. The horizontal axis represents frequency (Hz), and the vertical axis represents time (s), while the color scale indicates the amplitude of the wavelet coefficients. The black color in **Figure 5** indicates the cone of influence of the Continuous Wavelet Transform. The cone of influence indicates the region where edge effects due to the finite length of the signal are significant. The CWT representation highlights the dominant frequency components of the vibration signal. Several high-energy regions can be observed, corresponding to the principal dynamic components of the gear system. These components are associated with the mesh frequency and its harmonics. The CWT provides a joint time–frequency representation that allows transient vibration features to be clearly identified. The wavelet coefficient matrix obtained from the CWT was used as input to the LSTM model for damping identification. The advantage of the CWT lies in its ability to clearly represent the dominant frequencies while providing rich information in the time and frequency domains. The CWT plot reveals five dominant frequencies, characterized by their respective maximal wavelet coefficients in the CWT representation. The wavelet coefficients were used as inputs to the LSTM model.

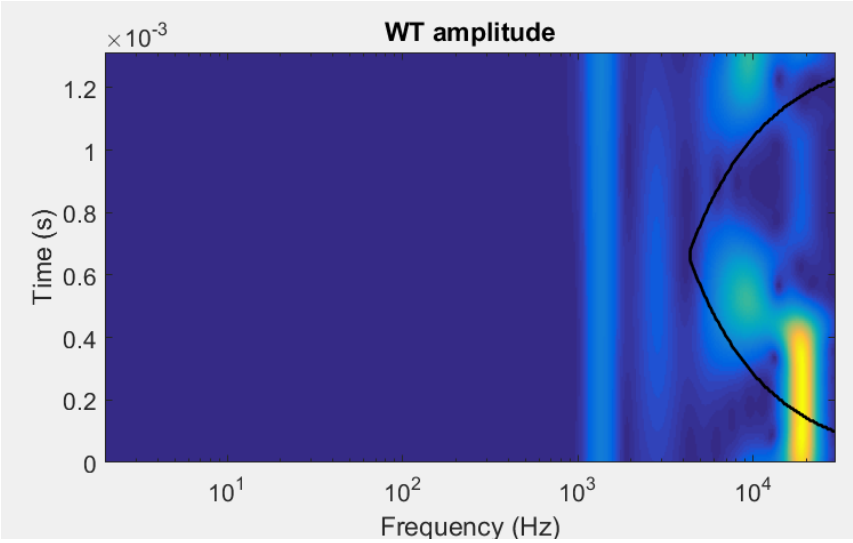

Figure 5. CWT-based time–frequency representations of the transmission error.

To further highlight the limitations of the proposed LSTM method compared with those of the proposed CWT–LSTM method, convergence analyses of the techniques were considered. In **Figure 6**, it can be noticed that the convergence of the training loss values calculated using the proposed CWT–LSTM method, as well as the proposed LSTM method, is depicted. The proposed models were trained for a maximum of

20 epochs. The proposed models converged at the 18th iteration. The results show that the training and validation errors of the proposed methods decreased with each subsequent iteration, justifying the effectiveness of the proposed models. The proposed CWT–LSTM model achieved convergence after the second iteration. However, the proposed LSTM model converged after the fifth iteration. The above results clearly demonstrate the effectiveness of the proposed CWT–LSTM method, which improves the stability of model convergence based on wavelet properties.

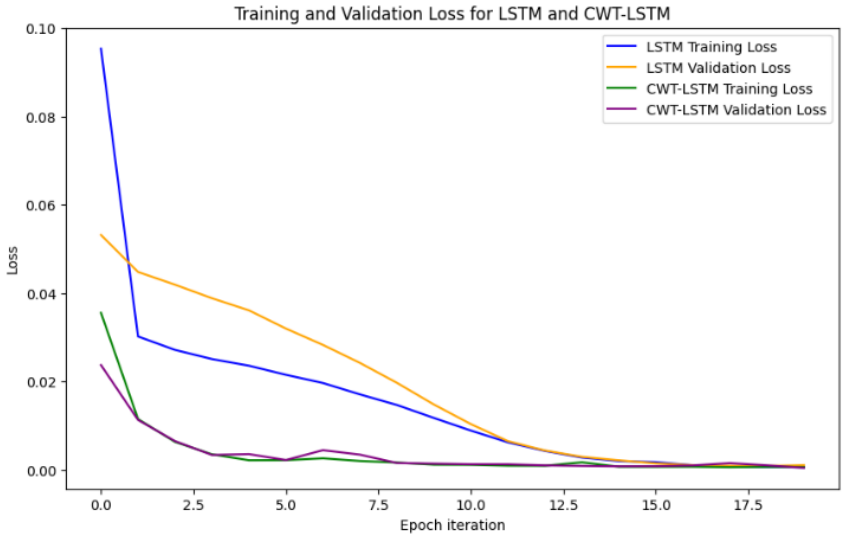

Figure 6. Performance loss during training and validation for non-linear damping.

Figure 7 compares the identified damping ratios obtained using the LSTM and CWT–LSTM frameworks trained with 400 training datasets. The RMSE obtained using the CWT–LSTM method was 3.19×10^{-4}. The RMSE obtained by the LSTM method was 6.37×10^{-4}. As presented by the identified results, the proposed CWT–LSTM method improved the identification compared with the results identified by the LSTM method. Therefore, the option of performing damping identification using the CWT technique should be used to improve the quality of the input data. This indicates that the continuous wavelet transform enhances the ability of the LSTM to capture fine temporal dynamics, providing more precise predictions even under ideal conditions. Therefore, incorporating the CWT improves the model's sensitivity to subtle variations in the damping ratio. In the next section, noise is added to the dataset to analyze the effect of noise on the performance of the proposed CWT–LSTM model in identifying damping.

3.3. Damping identification of the CWT–LSTM method in the case of noise

This section highlights the efficiency of the proposed CWT–LSTM approach in noisy inputs. A noisy input load is added to the gear system model. The Ricker wavelet is selected based on the stability of the wavelet in the presence of noisy conditions. **Figure 8** highlights the damping ratio estimates of the proposed LSTM and CWT–LSTM approaches. The value of RMSE of the proposed CWT–LSTM approach is 1.034×10^{-4}. The value of RMSE of the proposed LSTM approach is 3.99×10^{-4}. Unlike the LSTM technique, which uses a signal with noise, the CWT–LSTM

decomposes the signal into the scales of a mother wavelet. This property allows transient features and non-stationary components to be distinguished from noise, which is typically spread across scales.

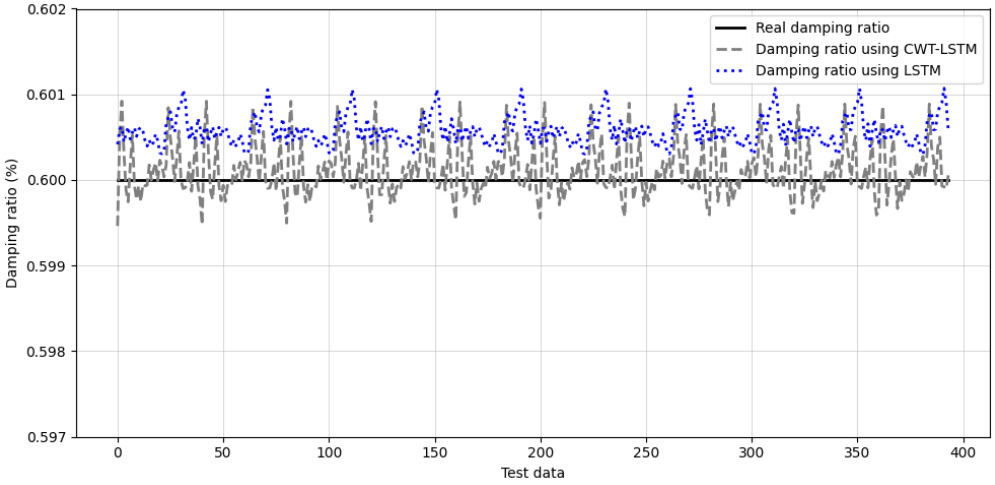

Figure 7. Damping ratio prediction using the CWT–LSTM and LSTM model.

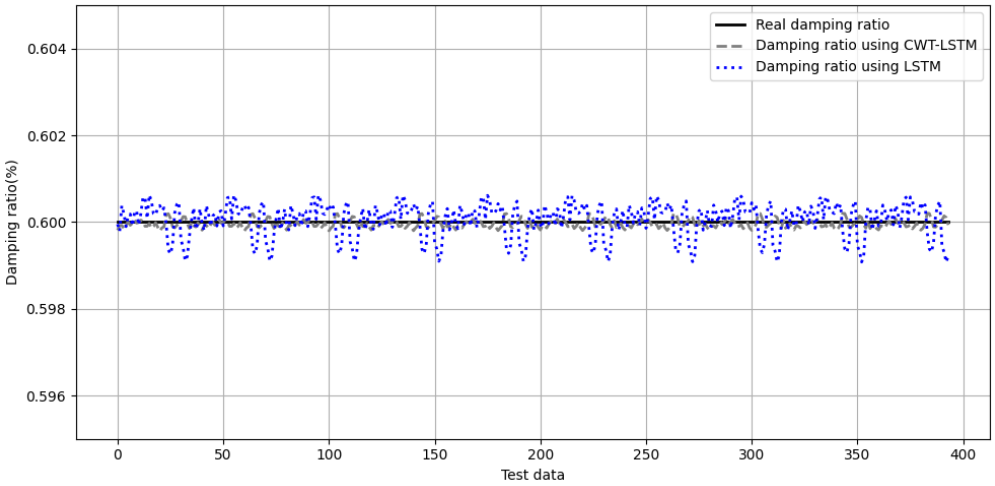

Figure 8. Damping ratio prediction using the CWT–LSTM and LSTM model.

Continuous Wavelet Transform is used as a preprocessing technique to improve stability in noisy inputs of the gear system. Importantly, although the noise modifies the excitation characteristics of the system, the proposed damping identification CWT–LSTM method remains robust and effective in their presence. **Figure 9** examines whether the CWT–LSTM method demonstrates satisfactory stability, even when noise is present in the input data. The noise input load applied to the gear system model, increasing the SNR from 1 to 50 dB. In the natural frequency results, the CWT–LSTM method shows significantly higher robustness under noisy conditions. As the value of SNR drops to less than 30 dB, the standard LSTM shows considerable degradation, whereas the results for the CWT–LSTM are still acceptable. Hence, it can be said that the advantage of the CWT–LSTM does not exist in the improvement of the final accuracy value, but it provides increased stability and robustness during the training phase. The CWT helps in stabilizing the training phase and increases the robustness to noise during training. In addition, the multi-scale characteristics of the CWT allow

for the effective reduction of noise-dominated features while maintaining physically meaningful dynamic information. This pre-processing stage ensures the stability of feature learning and avoids degradation in performance for low SNR conditions. Therefore, the proposed framework of CWT–LSTM shows better robustness during the training and testing phases, making it more reliable compared to the traditional LSTM model for noisy excitation load signals.

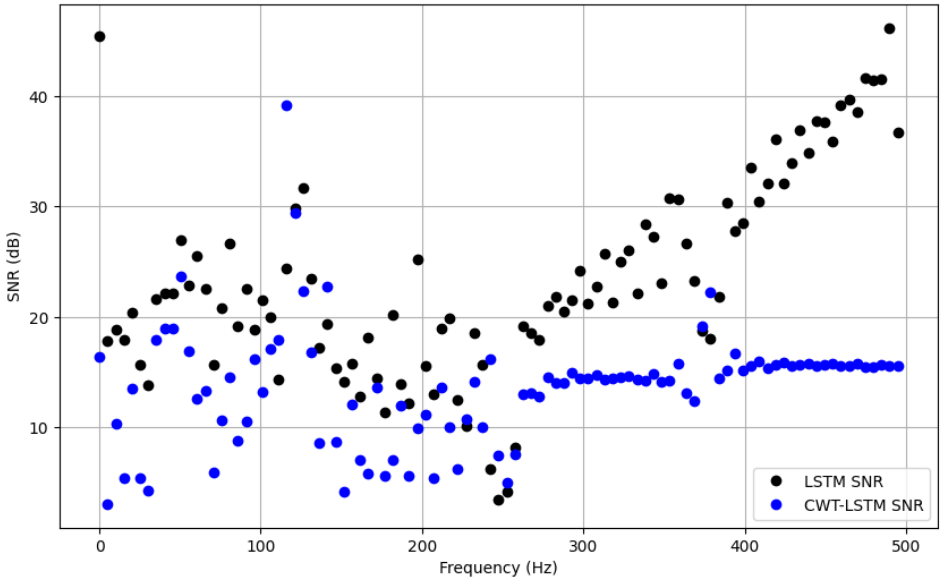

Figure 9. CWT–LSTM-based estimation of modal damping ratios under noisy conditions.

3.4. Modal identification of the CWT–LSTM method in the case of eccentricity

Eccentricity is defined as a shift between the gear's center of mass and its axis of rotation. This eccentricity provides additional dynamic excitations that can increase vibrations and noise. The eccentricity defect, denoted as $E = 0.001$, is represented as follows

$$\delta_e(t) = E\cos(\Omega t), \tag{11}$$

where Ω is the gear rotation speed. This eccentricity is incorporated into the time-dependent mesh stiffness $K(t)$ as follows

$$K_e(t) = K_0 + \Delta K f(t) + \delta_e(t), \tag{12}$$

where K_0 represents the nominal mesh stiffness, ΔK accounts for the periodic variation caused by tooth contact, and $f(t)$ is a periodic function representing the stiffness modulation over time.

Eccentricity was introduced into the spur gear system. The dynamic response was presented in the time–frequency domain using the CWT method, with the resulting wavelet plot shown in **Figure 10**, highlighting five dominant frequency components. The horizontal axis represents frequency (Hz), and the vertical axis represents time (s), while the color scale indicates the amplitude of the wavelet coefficients. The CWT representation highlights the dominant frequency components of the vibration signal. As shown in **Figure 5**, without eccentricity, the signal has a more focused distribution

in the time-frequency plane, indicating a relatively stable dynamic process. However, when eccentricity was added, new high-frequency components and energy spreading were observed, particularly in the higher frequency domain. This is because eccentricity affects the dynamic process of the system, and the CWT can effectively identify and distinguish these nonlinear effects in the time-frequency plane.

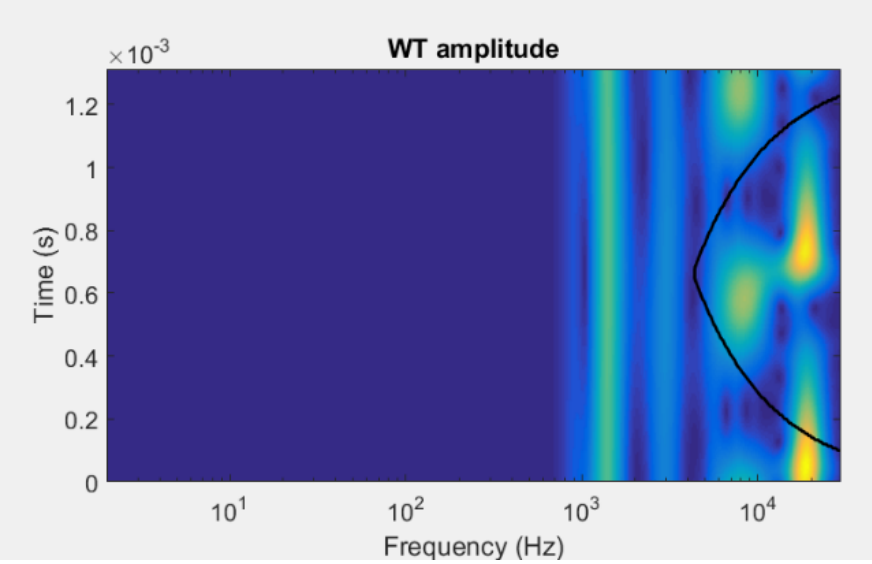

Figure 10. CWT-based time-frequency representations of the transmission error under eccentricity conditions.

Figure 11 presents the training loss evolution for both methods over a maximum of 20 epochs, with convergence achieved after 18 iterations. Both the training and validation losses steadily decreased, confirming successful network training. The CWT–LSTM model converged after the second epoch. However, the LSTM training convergence was reached at epoch 10. These results indicate that the proposed CWT–LSTM framework enhances the convergence stability in the presence of eccentricity.

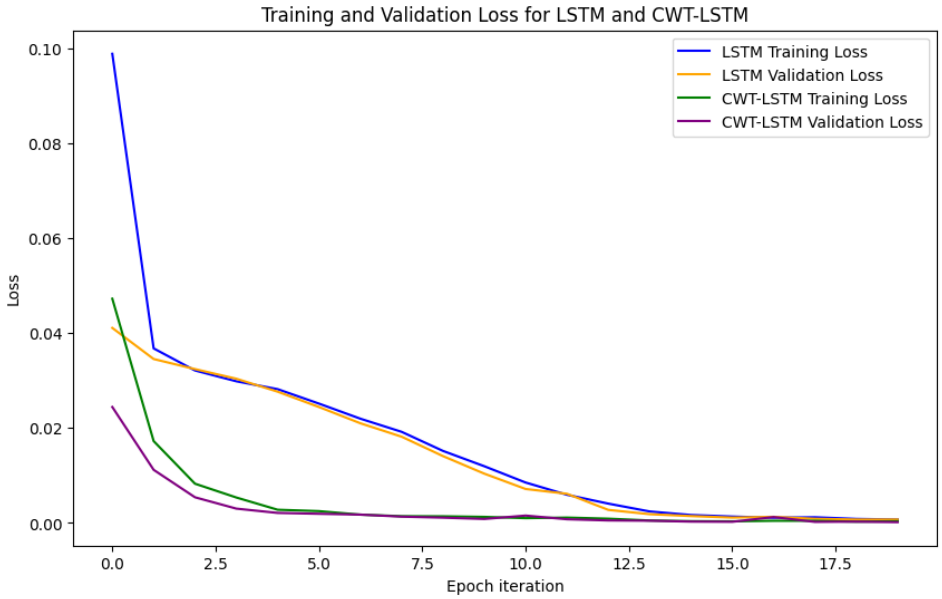

Figure 11. Performance loss during training and validation in the case of eccentricity.

Figure 12 compares the identified damping ratios obtained using the LSTM and CWT–LSTM frameworks trained with 400 training datasets. The RMSE obtained using the CWT–LSTM method was 8.33×10^{-4}. The RMSE obtained using the LSTM method was 2.38×10^{-4}. As presented in the identified results, the proposed CWT–LSTM method in this study improved the identification compared with the results identified by the LSTM method. Importantly, although geometric defects such as eccentricity modify the excitation characteristics of the system, the proposed damping identification CWT–LSTM method remains robust and effective in their presence.

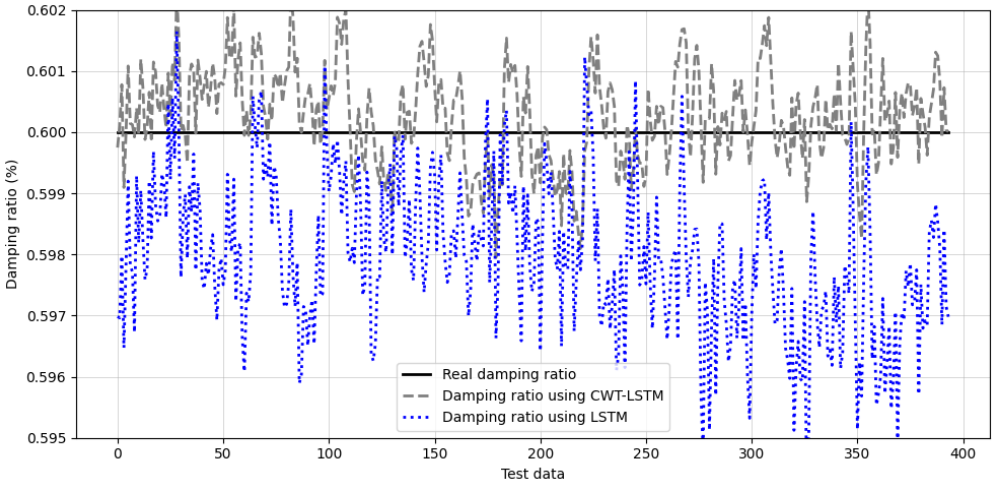

Figure 12. Damping ratio prediction using the CWT–LSTM and LSTM model under eccentricity.

3.5. Modal identification of the CWT–LSTM method in the case of backlash

The presented model incorporates a mechanical backlash to represent the nonlinear behavior of a rotational dynamic system. Backlash is a zone around zero where relative displacement between two components produces no effective motion. Mathematically, the backlash can be expressed as:

$$f_b = \begin{cases} 0 & |y| \le b \\ y - b.\text{sign}(y) & |y| \ge b \end{cases}. \tag{12}$$

Equation (10) can be expressed in the case of backlash as follows

$$\ddot{y} + C\dot{y} + \kappa(t)f_b = f \tag{14}$$

Backlash in mechanical transmission systems results in strongly nonlinear and discontinuous system behaviors. The aforementioned nonlinearities substantially affect the performance of Long Short-Term Memory (LSTM)-based identification algorithms. The presence of a nonlinear phenomenon, such as backlash, results in a nonstationary and impulsive vibration pattern, which makes learning by the LSTM algorithm more complex. Wavelet-based features are informative in the time scale, enabling the LSTM to effectively capture the temporal evolution of the backlash-related pattern across multiple scales. The CWT–LSTM framework results in significantly enhanced noise

robustness, feature separability, and generalization capability, yielding more accurate identification of system parameters such as the damping ratio. **Figure 13** compares the real damping ratio with the damping ratios identified using the LSTM and CWT–LSTM techniques when nonlinearity is introduced by the backlash. The RMSE obtained using the CWT–LSTM method was 1.93×10^{-4}. However, the RMSE obtained using the LSTM method was 3.92×10^{-4}. It can be observed that a good correspondence is obtained using the CWT–LSTM method. By contrast, the conventional LSTM method presents larger fluctuations around the true damping ratio. These results prove that using Continuous Wavelet Transform (CWT) enhances the feature representation of the signal, allowing the LSTM model to better capture system dynamics in the presence of backlash nonlinearity. Consequently, the proposed CWT–LSTM method demonstrated higher accuracy and robustness for damping ratio identification than the standard LSTM approach. Overall, the results confirm the effectiveness of the proposed method for reliable parameter identification in nonlinear systems affected by a backlash. The results confirm that even in the case of noise and nonlinear effects, such as gear eccentricity and backlash effects, the proposed method can correctly identify the damping coefficients.

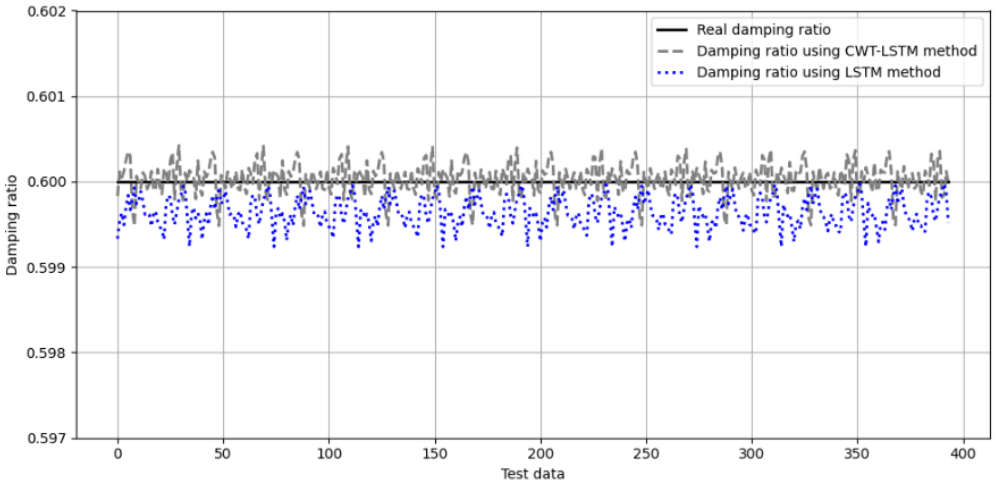

Figure 13. Damping ratio prediction using the CWT–LSTM and LSTM models.

4. Robustness evaluation of the CWT–LSTM method

4.1. Robustness evaluation of the CWT–LSTM method for different damping ratios

To evaluate the robustness of the CWT–LSTM method, the suggested identification method was examined for a range of constant damping ratios from 0.1 to 0.6%, as opposed to a single chosen value, to evaluate its robustness in terms of parameter variation. The damping ratio of 0.6% was first chosen as a representative case, which corresponds to relatively high dissipation levels, as can often be found in practice for mechanical systems. As shown in **Figure 14** and **Table 2**, the suggested CWT–LSTM model exhibits a high level of accuracy for the full range of damping ratios, with a close match to the real damping ratios. There was a significant difference

in the prediction errors between the LSTM and CWT–LSTM methods. Although the LSTM model can learn the temporal characteristics from the time-domain signals, its performance is still restricted in the case of nonlinear conditions. In contrast, the CWT–LSTM model always has a smaller prediction error over the entire range of damping ratios. This is because the CWT–LSTM model benefits from the time-frequency characteristics extracted by the continuous wavelet transform, which can provide a more comprehensive description of the system dynamics. The results indicate that the CWT–LSTM method is effective for a wide range of damping coefficients, which makes it possible to generalize and obtain accurate results for the gear system under different damping conditions.

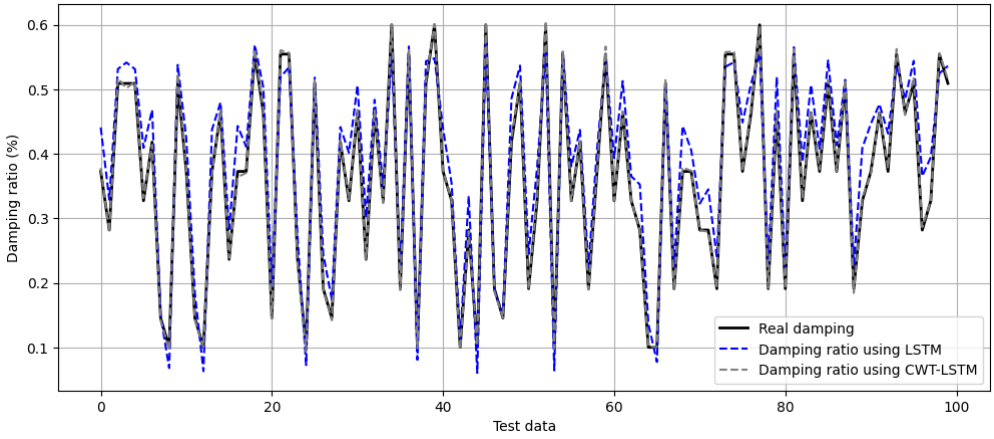

Figure 14. Comparison of LSTM and CWT–LSTM Methods for Different damping ratios.

Table 2. Root mean square error for different damping ratios.

Different damping ratios		**Time varying torque**	
CWT–LSTM method—RMSE	LSTM method—RMSE	CWT–LSTM method—RMSE	LSTM method—RMSE
0.0083	0.0534	0.003	0.062

4.2. Robustness evaluation of the CWT–LSTM method in the case of varying torque

To assess the robustness of the proposed CWT–LSTM methodology, the performance of the CWT–LSTM approach was analyzed under varying torque operating conditions. Varying torque operating conditions are considered realistic for assessing the dynamic response of gear systems. The varying torque conditions impose significant nonstationary and nonlinear influences on the dynamic response of the gear system. These effects considerably complicate the damping estimation tasks. **Figure 15** presents a comparison of the actual damping ratio with the predicted values based on the conventional LSTM approach and the proposed CWT–LSTM model for the test data associated with different torque values, as presented in **Table 2**. It can be noted that the predicted values based on the CWT–LSTM model follow the actual damping ratio curve closely, even in regions with abrupt changes in torque values.

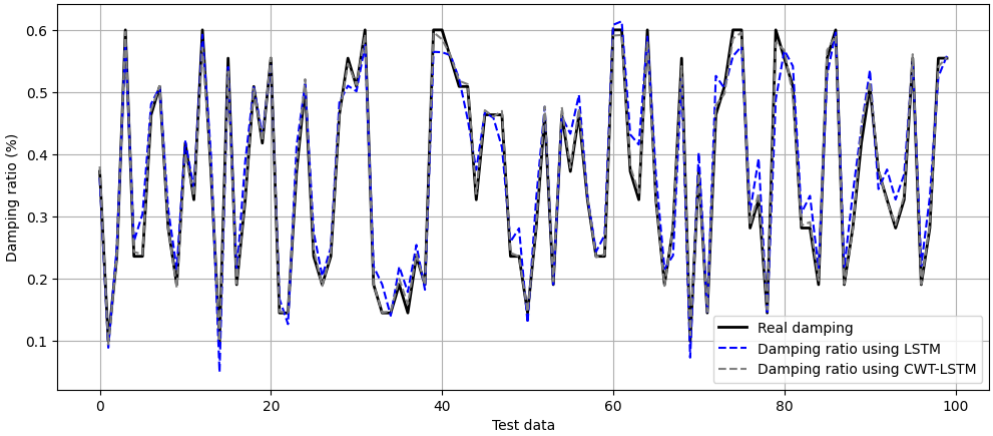

Figure 15. Comparison of LSTM and CWT–LSTM methods in the case of varying torque.

Overall, the results indicate that the proposed CWT–LSTM framework possesses good generalization capability and robustness in the case of varying torque. This is an important property, especially in real-life scenarios involving gear monitoring, where the operating conditions may not always be constant. The capability of the CWT–LSTM approach to ensure a high level of estimation accuracy under variable torque conditions suggests its potential for reliable damping identification.

5. Conclusion

In this study, the nonlinear damping parameter identification problem is addressed by integrating the Continuous Wavelet Transform technique with Long Short-Term Memory networks. The aim of this study is to improve the accuracy of nonlinear damping parameter identification models for spur gear pair systems under a more realistic operating environment. The Continuous Wavelet Transform technique is the most critical component of this study because it provides a more accurate method for analyzing vibration signals with a high resolution in the time-frequency domain. In traditional vibration signal analysis, this is not feasible because the method lacks accuracy in predicting the behavior of the system owing to its complexity.

The compact and informative time–frequency features derived from the CWT were used as inputs to the LSTM model. Feature transformation significantly enhances the learning capability of the neural network by emphasizing the transient events, frequency modulation, and multi-scale characteristics induced by the nonlinear effects of backlash, eccentricity, and noise. Consequently, the CWT–LSTM approach outperforms the standard LSTM method in estimating the nonlinear damping properties of spur gear pair systems. These results confirm that the proposed CWT–LSTM framework improves both the estimation accuracy and convergence behavior of the learning process. In particular, the faster convergence and higher robustness of the proposed method are evident for input signals polluted by noise or defects. The reliability and stability of nonlinear damping identification are increased by including CWT features in this approach. Therefore, the CWT–LSTM technique is more applicable and appropriate for practical engineering applications than conventional LSTM-based methods. The CWT–LSTM method was evaluated for different damping values and demonstrated good effectiveness.

In future work, we intend to enhance the CWT–LSTM model by considering the frictional force between the teeth of the gears. The friction between the teeth increases the energy damping terms, which may affect the damping ratio identification procedure. Hence, it is possible to determine the influence of the frictional force terms on the damping ratio identification procedure and examine how the influence of the energy damping terms affects the time-frequency characteristics of the response signal. Thus, a better understanding of the damping characteristics can be provided by the enhanced CWT–LSTM model. The simulated responses used in this study are generated from a physically consistent dynamic gear model, which provides realistic mechanical conditions; however, experimental measurements are still required to fully assess the generalization capability of the proposed identification method. Therefore, experimental validation using measured vibration data will be considered as an important direction for future work in order to further confirm the practical applicability of the proposed approach.

Author contributions: NY, FM and AA: original draft preparation, visualization, validation and investigation; LW: methodology and software, contributed to the discussions and first reviewing; MH: supervision, reviewing and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional review board statement: Not applicable.

Informed consent statement: Not applicable.

Data availability statement: The data used to support the findings of this study are available from the corresponding author upon request

Conflict of interest: The authors declare no conflict of interest.

AI use statement: The authors confirm that the manuscript was written by themselves. AI-based tools were used only for minor language polishing and grammar suggestions, and did not contribute to the scientific content, analysis, or interpretation of the work.

References

1. Al-hababi T, Cao M, Saleh B, et al. A Critical Review of Nonlinear Damping Identification in Structural Dynamics: Methods, Applications, and Challenges. Sensors. 2020; 20(24): 7303. doi: 10.3390/s20247303
2. Ghahari F, Sargsyan K, Taciroglu E. Quantification of modeling uncertainty in the Rayleigh damping model. Earthquake Engineering & Structural Dynamics. 2024; 53(9): 2950–2956. doi: 10.1002/eqe.4143
3. Zahid FB, Ong ZC, Khoo SY. A review of operational modal analysis techniques for in-service modal identification. Journal of the Brazilian Society of Mechanical Sciences and Engineering. 2020; 42(8): 398. doi: 10.1007/s40430-020-02470-8
4. Lisitano D, Bonisoli E, Mottershead JE. Experimental direct spatial damping identification by the Stabilised Layers Method. Journal of Sound and Vibration. 2018; 437: 325–339. doi: 10.1016/j.jsv.2018.08.055
5. Bonisoli E, Lisitano D, Vigliani A. Damping identification and localisation via Layer Method: Experimental application to a vehicle chassis focused on shock absorbers effects. Mechanical Systems and Signal Processing. 2019; 116: 194–216. doi: 10.1016/j.ymssp.2018.06.013
6. Amabili M, Alijani F, Delannoy J. Damping for large-amplitude vibrations of plates and curved panels, part

2: Identification and comparisons. International Journal of Non-Linear Mechanics. 2016; 85: 226–240. doi: 10.1016/j.ijnonlinmec.2016.05.004
7. Balasubramanian P, Ferrari G, Amabili M. Identification of the viscoelastic response and nonlinear damping of a rubber plate in nonlinear vibration regime. Mechanical Systems and Signal Processing. 2018; 111: 376–398. doi: 10.1016/j.ymssp.2018.03.061
8. Haghdoust P, Lo Conte A, Cinquemani S, et al. A Numerical Method to Model Non-linear Damping Behaviour of Martensitic Shape Memory Alloys. Materials. 2018; 11(11): 2178. doi: 10.3390/ma11112178
9. Müller F, Woiwode L, Gross J, et al. Nonlinear damping quantification from phase-resonant tests under base excitation. Mechanical Systems and Signal Processing. 2022; 177: 109170. doi: 10.1016/j.ymssp.2022.109170
10. Chatterjee A, Chintha HP. Identification and parameter estimation of cubic nonlinear damping using harmonic probing and volterra series. International Journal of Non-Linear Mechanics. 2020; 125: 103518. doi: 10.1016/j.ijnonlinmec.2020.103518
11. Eberle R. Estimating Nonlinear Damping for Mechanical Single-Degree-of-Freedom Systems: A Robust and Effective Approach. PAMM. 2025; 25(1): e202400158. doi: 10.1002/pamm.202400158
12. Han SL, Kinoshita T. Nonlinear Damping Identification in Nonlinear Dynamic System Based on Stochastic Inverse Approach. Mathematical Problems in Engineering. 2012; 2012(1): 574291. doi: 10.1155/2012/574291
13. Zaparoli Cunha B, Droz C, Zine AM, et al. A review of machine learning methods applied to structural dynamics and vibroacoustic. Mechanical Systems and Signal Processing. 2023; 200: 110535. doi: 10.1016/j.ymssp.2023.110535
14. Yemets K, Izonin I, Dronyuk I. Enhancing the FFT-LSTM Time-Series Forecasting Model via a Novel FFT-Based Feature Extraction–Extension Scheme. Big Data and Cognitive Computing. 2025; 9(2): 35. doi: 10.3390/bdcc9020035
15. Yun DY, Oh BK, Park K, et al. LSTM-Based Approach for Stable Identification of Modal Damping Ratio in Building Structures. Structural Control and Health Monitoring. 2024; 2024(1): 6645626. doi: 10.1155/2024/6645626
16. Liu D, Tang Z, Bao Y, et al. Machine-learning-based methods for output-only structural modal identification. Structural Control and Health Monitoring. 2021; 28(12). doi: 10.1002/stc.2843
17. Chen Y, Rao M, Feng K, et al. Physics-Informed LSTM hyperparameters selection for gearbox fault detection. Mechanical Systems and Signal Processing. 2022; 171: 108907. doi: 10.1016/j.ymssp.2022.108907
18. Afridi YS, Hasan L, Ullah R, et al. LSTM-Based Condition Monitoring and Fault Prognostics of Rolling Element Bearings Using Raw Vibrational Data. Machines. 2023; 11(5): 531. doi: 10.3390/machines11050531
19. Kulkarni A, Paul ASB. A hybrid EKF-LSTM framework with GA-tuned noise covariance for enhanced time series forecasting. Engineering Research Express. 2025; 7(3): 035266. doi: 10.1088/2631-8695/adf9be
20. Zheng Y, Lee CL, Guo J, et al. Improving EFDD with Neural Networks in Damping Identification for Structural Health Monitoring. Sensors. 2025; 25(22): 6929. doi: 10.3390/s25226929
21. Yun DY, Park HS. Modal identification of building structures under unknown input conditions using extended Kalman filter and long-short term memory. Integrated Computer-Aided Engineering. 2023; 30(2): 185–201. doi: 10.3233/ICA-220696
22. Yun DY, Shim HB, Park HS. SSI-LSTM network for adaptive operational modal analysis of building structures. Mechanical Systems and Signal Processing. 2023; 195: 110306. doi: 10.1016/j.ymssp.2023.110306
23. Han Z, She D, Liu J. LSTM-Based Fast Prediction of Seismic Response and Fragility for Bridge Pile-Group Foundations: A Data-Driven Design Approach. Designs. 2026; 10(2): 37. doi: 10.3390/designs10020037
24. Chen Z, Li C, Sanchez RV. Gearbox Fault Identification and Classification with Convolutional Neural Networks. Shock and Vibration. 2015; 2015: 1–10. doi: 10.1155/2015/390134
25. Yin A, Yan Y, Zhang Z, et al. Fault Diagnosis of Wind Turbine Gearbox Based on the Optimized LSTM Neural Network with Cosine Loss. Sensors. 2020; 20(8): 2339. doi: 10.3390/s20082339
26. Zhang BC, Sun SQ, Yin XJ, et al. Research on Gearbox Fault Diagnosis Method Based on VMD and Optimized LSTM. Applied Sciences. 2023; 13(21): 11637. doi: 10.3390/app132111637
27. He C, Yasenjiang J, Lv L, et al. Gearbox Fault Diagnosis Based on MSCNN-LSTM-CBAM-SE. Sensors. 2024; 24(14): 4682. doi: 10.3390/s24144682
28. Xia X, Sun H, Wang A. Fault Diagnosis of Planetary Gearboxes Based on LSTM Improved via Feature Extraction Using VMD, Fusion Entropy, and Random Forest. Entropy. 2025; 27(9): 956. doi: 10.3390/e27090956
29. Cheng X, Dou S, Du Y, et al. Gearbox fault diagnosis method based on lightweight channel attention mechanism and transfer learning. Scientific Reports. 2024; 14(1): 743. doi: 10.1038/s41598-023-50826-6
30. Fang X, Luo H, Tang J. Structural damage detection using neural network with learning rate improvement.

Computers & Structures. 2005; 83(25–26): 2150–2161. doi: 10.1016/j.compstruc.2005.02.029
31. Chang CC, Chang TYP, Xu YG, et al. Structural Damage Detection Using an Iterative Neural Network. Journal of Intelligent Materials Systems and Structures. 2000; 11(1): 32–42. doi: 10.1177/104538900772664387
32. Lei Y, Liu C, Liu LJ. Identification of multistory shear buildings under unknown earthquake excitation using partial output measurements: Numerical and experimental studies. Structural Control and Health Monitoring. 2014; 21: 774–783. doi: 10.1002/stc.1600
33. Yousfi N, Zghal B, Akrout A, et al. Damping models identification of a spur gear pair. Mechanism and Machine Theory. 2018; 122: 371–388. doi: 10.1016/j.mechmachtheory.2018.01.002
34. Dien NP. Damping identification in multi-degree-of-freedom systems using the continuous wavelet transform. Vietnam Journal of Mechanics. 2005; 27(1): 41–50. doi: 10.15625/0866-7136/27/1/5713
35. Li H. Gear Damage Detection Using Impulse Response Wavelet Transform. Advanced Materials Research. 2011; 204–210: 1378–1381. doi: 10.4028/www.scientific.net/AMR.204-210.1378
36. Nguyen TD, Nguyen PD. Improvements in the Wavelet Transform and Its Variations: Concepts and Applications in Diagnosing Gearbox in Non-Stationary Conditions. Applied Sciences. 2024; 14(11): 4642. doi: 10.3390/app14114642