Friction

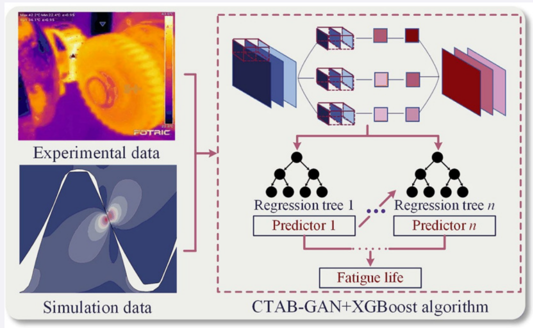
Research Article

A strategy for contact fatigue life prediction of polymer gears via an experimental-simulated hybrid data-driven model

Zehua Lu[1,2], Huaiju Liu[1,✉], Peitang Wei[1], Damijan Zorko[3]

Cite this article: Lu ZH, Liu HJ, Wei PT, et al. *Friction* 2026, **14**(2): 9441077. https://doi.org/10.26599/FRICT.2025.9441077

ABSTRACT: A timely trend in gear transmission involves the replacement of steel with polymers. Nevertheless, the absence of fundamental durability data for polymer gears impedes their reliable application during power transmission. The expensive and time-consuming gear fatigue experiments make it impossible to rely merely on experimental data. In this study, a strategy for contact fatigue life prediction of polymer gears via an experimental-simulated hybrid data-driven model is presented. The hybrid data are established with a certain mixture ratio of experimental and simulation data and are augmented by the conditional tabular generative adversarial network (CTAB-GAN) algorithm. This specific algorithm was combined with the extreme gradient boosting (XGBoost) algorithm to predict the contact fatigue life of gears made from different polymer materials, with the prediction accuracy controlled within a 3-fold scatter band. Moreover, an empirical predictive formula for contact fatigue life was developed. The hybrid data-driven model, which merges experimental and simulated data, allows for efficient estimation of fatigue life and material selection strategies, generating insight into the anti-fatigue design of polymer gears.



KEYWORDS: polymer gear; contact fatigue; life prediction; hybrid data driven

1 Introduction

Compared with metallic gears, polymer gears have several advantages, such as lightweight, low cost, corrosion resistance, and high damping. In the early stages, polymer gears were used primarily in motion transmissions for medical devices, office equipment, and instrumentation. However, with advancements in material science and molding technology, the load-carrying capacities of polymer gears have distinctly increased. Polymer gears have been used in aeroengines and renewable energy vehicles to reduce costs and increase power density in power transmissions [1]. There is an expanding trend in the power transmission of polymers to replace metallic materials.

On one hand, Reitschuster et al. [2] examined the use of polymer gears in a high-speed small electric vehicle, demonstrating their feasibility through substantial experimentation. On the other hand, Demšar et al. [3] studied the use of polymer gears in drive reduction gearboxes for electric bicycles, providing valuable insights to ensure their widespread application. Establishing the load-carrying capacity of polymer gears in power transmissions is a complex task because of the wide range of polymer materials used and the critical impact of external environmental factors such as temperature and load. State-of-the-art material selection for polymer gears significantly relies on engineering experience, while insufficient experimental data restrict their application in power transmissions.

The durability performance of polymer gears is affected by lubrication conditions, molding processes, ambient temperature, geometric structure, and so forth. A number of potential failure modes can occur, including wear, thermal melting, contact and bending fatigue, and tooth deformation. The intricate nature of these factors and failure modes renders the study of polymer gear durability a complex and challenging task. Several studies have investigated the load-carrying capacities and failure modes of polymer gears under dry contact conditions in light of their self-lubricating properties. For example, Mao et al. [4, 5] conducted load-carrying capacity tests of polyoxymethylene (POM) gears under dry contact conditions and discovered the vital load transition from wear failure to thermal melting. Concurrently, Senthilvelan and Gnanamoorthy [6, 7] investigated the load-carrying capacity and failure modes of polyamide (PA) gears, focusing on the effects of the fiber type, tooth root fillet radius, and torque–speed combination under dry contact conditions.

Nevertheless, the load-carrying capacity of polymer gears under

[1] State Key Laboratory of Mechanical Transmission for Advanced Equipment, Chongqing University, Chongqing 400044, China. [2] College of Aerospace Engineering, Chongqing University, Chongqing 400044, China. [3] Faculty of Mechanical Engineering, University of Ljubljana, Aškerčeva cesta 6, Ljubljana 1000, Slovenia.

✉ Corresponding author. E-mail: huaijuliu@cqu.edu.cn

Received: April 8, 2024; Revised: December 21, 2024; Accepted: February 6, 2025

Nomenclature

A	Yield stress at the reference temperature and strain rate (MPa)	T_{trans}	Transition temperature (°C)
A^*	Accuracy index	$T(X,\theta_i)$	Decision tree
B	Strain hardening coefficient	u	Gear ratio
b_{i}	Tooth width (mm)	x	Data of the original sample
b_{w}	Common face width of the gear pair (mm)	$\tilde{x}$	Joint probability distribution
C	Strain rate sensitivity parameter	$\breve{x}$	Data of the generator
C^*, C_1, C_2	Material constants	Z_{E}	Elasticity factor
c	Fatigue ductility exponent	Z_{H}	Zone factor
$D(x)$	Neural network of discriminator	Z_{ε}	Contact ratio factor
d_1	Reference circle diameter of pinion (mm)	Z_{R}	Surface roughness factor
E	Elastic modulus (MPa)	α	Modulus attenuation coefficient (MPa/°C)
E_0	Elastic modulus at 0 °C (MPa)	β	Material constant
$\mathop{\mathbb{E}}_{\tilde{x}\,\mathbb{P}_r}$	Expectation of the Wasserstein distance	$\dot{\varepsilon}$	Strain rate
F_{t}	Nominal tangential force (N)	$\dot{\varepsilon}_0$	Reference strain rate
k	Fatigue strength exponent	ε^{P}	True plastic strain
L	Wasserstein loss function	$\varepsilon_{\mathrm{f}}^{'}$	Fatigue ductility coefficient
n	Strain hardening exponent	σ	Flow stress (MPa)
N_{f}	Fatigue life of the material point	σ_{H}	Maximum contact stress (MPa)
T	Operating temperature (°C)	σ_{m}	Mean normal stress on the critical plane (MPa)
o	Tree number	$\sigma_{\mathrm{u}}^{'}$	Conventional monotonic ultimate tensile stress (MPa)
$\mathbb{P}_g$	Generated sample distribution	$\sigma_0^{'}$	Ultimate tensile stress at 0 °C (MPa)
$\mathbb{P}_r$	Original sample distribution	σ_{HlimN}	Allowable contact stress (MPa)
S_{Hmin}	Required minimum safety factor	$\Delta\gamma_{\mathrm{max}}/2$	Maximum shear strain amplitude on the critical plane
T_{room}	Ambient temperature (°C)	$\Delta\varepsilon_{\mathrm{n}}/2$	Normal strain amplitude on the critical plane
T_{melt}	Melting point temperature (°C)	λ	Lubricant coefficient

dry conditions is limited because of the significant increase in temperature. Hasl et al. [8] shed light on how oil-lubricated POM gears can support transmission power of nearly 30 kW based upon durability experiments, which presented an especially significant stage in the development of polymer gears in power transmissions. Under dry contact conditions, the load-carrying capacity of polymer gears becomes significantly influenced by the load conditions and operating temperature. The transition from dry contact to oil lubrication makes fatigue failure the dominant failure mode of polymer gears. Specifically, when polymer gears with high geometric precision running under oil jet lubrication are considered, contact fatigue failure becomes the primary failure mode [9, 10].

Currently, fatigue experiments are the leading way to investigate the contact fatigue performance of polymer gears, with several studies employing numerical simulations. Illenberger et al. [11, 12] utilized the group testing method to assess the contact fatigue of stress vs. number of load cycles (*S*–*N*) curves for polyether-ether-ketone (PEEK) gears under oil lubrication. This supported the strength design for PEEK gears intended for power transmissions. Additionally, Lu et al. [13, 14] performed numerical analysis on the contact fatigue of POM gears given the effect of temperature. These authors indicated that elevated temperatures resulted in a decrease in the contact fatigue performance of gears, despite lower contact stress.

Notably, while experiments and numerical simulations provide effective methods for evaluating the contact fatigue performance of polymer gears, practical limitations in terms of techniques, time, and cost remain. For a specific metallic material, only one *S*–*N* curve is needed, whereas for a selected polymer material, owing to the polymer's temperature-dependent properties, several *S*–*N* curves tend to be needed, considering different temperature conditions. Fatigue experiments for polymer gears lead to a much longer testing time, incurring higher costs. The analysis of fatigue data relies primarily on an explicit regression algorithm, which demands a substantial quantity of fatigue data to construct an accurate model for gear life prediction. Numerical models promote the efficient prediction of the contact fatigue life of the polymer gear, following the complete outline of boundary conditions. However, specific inherent issues of numerical approaches remain, such as complex damage criteria, a lack of validated data, and the inability to load sequence effects.

In recent years, machine learning (ML) has been more prominently utilized to address the challenges posed by oversimplified numerical simulations as well as time-consuming fatigue experiments. ML has been popularly applied in gear fatigue life prediction [15], data processing [16], and failure mechanism analysis [17]. The present ML algorithms for fatigue life prediction overwhelmingly include artificial neural networks [18], support vector machines [19], and random forests [20].

Conventional ML algorithms demand a significant quantity of experimental data, which can be quite challenging to obtain. To increase the accuracy of life prediction, it is necessary to apply active learning or transfer learning algorithms. Zhang et al. [21] developed a transfer learning model to predict the stress–strain curves of diverse additively manufactured fiber-reinforced polymer composites, achieving satisfactory prediction accuracy. Li et al. [22] proposed a transfer learning neural network framework to predict the contact fatigue life of 18CrNiMo7-6 gears based on fatigue data of AISI 9310 rollers. The framework significantly minimized the required amount of test data. Nevertheless, owing to the complex performance of different polymer materials, transfer learning has not yet matured sufficiently to transfer fatigue lives between polymer gears. Nevertheless, the results suggest a promising way of improving the life prediction accuracy of ML models through data augmentation. Jia et al. [23] suggested a Wasserstein generative adversarial network (WGAN)-–extreme gradient boosting (XGBoost) data augmented model to predict the fatigue lives of POM gears, alongside a gear-life formula considering temperature, load, and several other factors. Currently, data augmentation is a prevailing technique in ML models used for predicting the fatigue performance of components made from the same material. However, it is critical to acknowledge that different polymer materials present substantial disparities. Employing a limited experimental dataset for data augmentation between different polymer materials may diminish the generalizability of the results.

This study offers a strategy for predicting polymer gear contact fatigue life based on experimental-simulated data via a conditional tabular generative adversarial network (CTAB-GAN)+XGBoost algorithm. Contact fatigue experiments were conducted for POM, PEEK, and polyketone (POK) gears, and a simulation model was developed for the temperature of the polymer gears. The experimental and simulated fatigue data were mixed, forming a hybrid dataset, which was further augmented by the CTAB-GAN algorithm. This specific algorithm was combined with the XGBoost algorithm to predict the contact fatigue life of gears made of different polymer materials. Additionally, an empirical formula for calculating the contact fatigue life of polymer gears according to material properties and loading conditions was developed. Leveraging experimental–simulated hybrid data-driven models to study the contact fatigue performance of polymer gears will further benefit the design of high-performance polymer gears for power transmissions.

2 Methodology

2.1 Contact fatigue experimental data

To generate experimental data, contact fatigue experiments were conducted for polymer gears using a power-open gear test rig, as demonstrated in Fig. 1. The test rig includes two primary spindle boxes, a drive motor, guide rails, and a monitoring system. The drive spindle box moves along the guide rail to adjust the center distance. Accelerometers are installed on the bearing supports for real-time vibration monitoring. Once the vibration signal value exceeds 10%–20% over normal operation, the durability experiment is automatically terminated. Gear tests were conducted under oil/grease-lubricated or dry contact conditions. To promote steady lubrication conditions, a jet lubrication system with active control of the lubricant's temperature in the range of 30–200 °C was applied. Additionally, the ambient temperature was controlled at 23±2 °C.

In all the tests, the pinion was composed of 18CrNiMo7-6 carburizing steel. Following carburization and quenching, the surface hardness reached 59–62 HRC. A digital automatic turret microhardness tester (MHVS-1000AT, Aolongxingdi) was utilized to perform hardness measurements, which permitted a depth range of 1,000 μm. The intervals for measuring were 50 and 100 μm for depths of 0–100 μm and 100–1,000 μm, respectively. At every point, three measurements were taken, and the average value was calculated. Moreover, a load of 1,000 g was applied with a load retention time of 12 s.

The polymer wheels were made of PEEK (Victrex, 450G), POM (Duracon, M90-44), and POK (Hyosung, M630A), as shown in Fig. 1(b). The main mechanical properties of the tested polymer wheels and polyethylene-terephthalate (PET, from Verein Deutscher Ingenieure (VDI) 2736 [27]) at 23 °C are shown in Table 1.

Table 2 shows the geometrical parameters for the tested gear pairs. The maximum specific sliding was maintained below 1, ensuring that contact fatigue failure occurred before the other failure modes. A white light interferometer (RTEC MFT-5000) with a magnification of 10× and a spot size of 1.11 mm × 0.89 mm was used for the measurement of surface morphology. The geometric accuracy was measured on a computer numerical control (CNC) gear measuring machine (Klingelnberg, P26). The metallic gears were ground to R_a = 1.6 μm with a geometric accuracy of Q7, whereas the polymer gears were hobbed to R_a = 1.2–1.8 μm with a geometric accuracy of Q10, per the Chinese

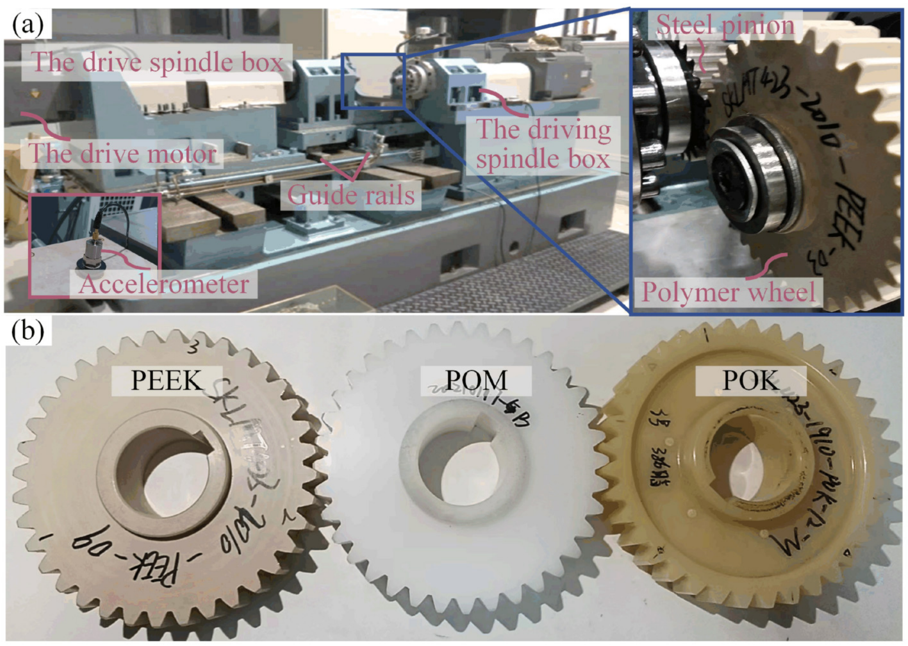


Fig. 1 Gear test rig and tested gears: (a) power-open gear test rig and (b) tested polymer gears.

Standard GB/T 10095.1-2022 [24]. None of the tested gears had tooth tip modifications.

The contact fatigue experiments of the polymer gears were conducted following the Chinese Standard GB/T 14229-2021, “Test method for gear contact fatigue strength” [25]. The input speed was set as 1,000 r/min, and the output torques were 40, 60, 80, and 100 N·m. Following the German Guideline VDI 2736.2-2014 [27], the maximum contact stress of the polymer gears σ_H can be calculated as Eq. (1):

$$\sigma_{\mathrm{H}} = Z_{\mathrm{E}} \cdot Z_{\mathrm{H}} \cdot Z_{\varepsilon} \cdot Z_{\beta} \cdot \sqrt{\frac{F_{\mathrm{t}} \cdot K_{\mathrm{H}}}{b_{\mathrm{w}} \cdot d_1} \cdot \frac{u+1}{u}} \leq \sigma_{\mathrm{HP}} = \frac{\sigma_{\mathrm{HlimN}} \cdot Z_{\mathrm{R}}}{S_{\mathrm{Hmin}}} \quad (1)$$

where σ_H represents the maximum contact stress with units of MPa; Z_E represents the elasticity factor; Z_H corresponds to the zone factor; Z_ε represents the contact ratio factor; Z_R is the surface roughness factor; F_t is the nominal tangential force with units of N; u is the gear ratio; b_w is the common face width of the gear pair with units of mm; d_1 is the reference circle diameter of the pinion with units of mm; σ_{HP} represents the permissible contact stress; σ_{HlimN} represents the allowable contact stress; S_{Hmin} represents the required minimum safety factor.

The oil lubricant was MOBILGEAR SHC 627, classified as ISO-VG100. The grease lubricant was a lithium-based grease whose base oil was the 1-dodecene homopolymer 1-dodecane. The fundamental properties of the lubricants are listed in Table 3.

The oil inlet temperature was controlled at 30, 40, or 80 °C, while the oil feed rate was fixed at 1 L/min. Grease-lubricated gears were operated in an open environment at an ambient temperature of 23±2 °C. The contact fatigue failure criterion was established according to the degree of pitting damage on the tooth surface by intermittent stops and visual checking of the pitted area. When the pitting area exceeded 10% of the total tooth surface area or when 10^7 loading cycles were performed, the experiment was halted.

As shown in Fig. 2(a), an infrared thermal imager (Fotric, 238) with an emissivity of 0.95 was used to detect the gear bulk temperature [26]. Furthermore, a thermal sensor was used to measure the oil temperature at the outlet zone. The oil temperature of the polymer gears remained stable during their service life, as shown in Fig. 2(b). The tooth temperature of oil-lubricated polymer gears remains independent of the output torque, whereas that of grease-lubricated polymer gears increases with the output torque, as shown in Fig. 2(c).

Figure 3(a) highlights how the PEEK gear fatigue experiments result in contact fatigue failure, which manifests primarily as surface pitting close to the pitch line. Surface pitting is the dominant failure mode of the tested polymer gears. The relationships between the contact fatigue *S*–*N* curves are outlined in Fig. 3(b). The contact fatigue life of oil-lubricated PEEK gears is 1–2 times longer than that of POM gears at a given contact stress and temperature. The contact fatigue life of oil-lubricated POK gears is slightly lower than that of POM gears.

Table 4 details the test status of the polymer gears, including information on the output torque, lubrication mode, tooth

Table 1 Main mechanical properties of tested polymer wheels at 23 °C

Parameter	PEEK	POM	POK	PET
Grade ID	450G	M90-44	M630A	From VDI 2736
Density (g/cm³)	1.32	1.41	1.24	1.39
Elastic modulus (MPa)	3,600	2,700	1,350	2,800
Tensile strength (MPa)	95	62	58	75
Poisson’s ratio	0.38	0.3	0.47	0.33
Thermal conductivity (W/(m·°C))	0.25	0.3	0.2	0.24
Melting temperature (°C)	334	165	222	250

Table 2 Geometric parameters of tested gear pairs

Parameter	Pinion	Wheel
Normal module (mm)	3	
Pressure angle	20°	
Gear ratio	1.5	
Centre distance (mm)	94.5	
Root fillet factor	0.38	
Teeth number	24	36
Tooth width (mm)	22	20
Tip diameter (mm)	81.26	118.28
Base diameter (mm)	69.22	106.24
Profile shifting	0.7873	0.9567
Gear accuracy	GB/T 10095 Q7	GB/T 10095 Q10

Table 3 Basic properties of the lubricants

Parameter	Value of oil	Parameter	Value of grease
Density (kg/L)	0.89	NLGI	1
Viscosity index	98	Drop point (°C)	200
Viscosity at 40 °C (cSt)	100	Flash point (°C)	260
Viscosity at 100 °C (cSt)	12	Penetration	310–340

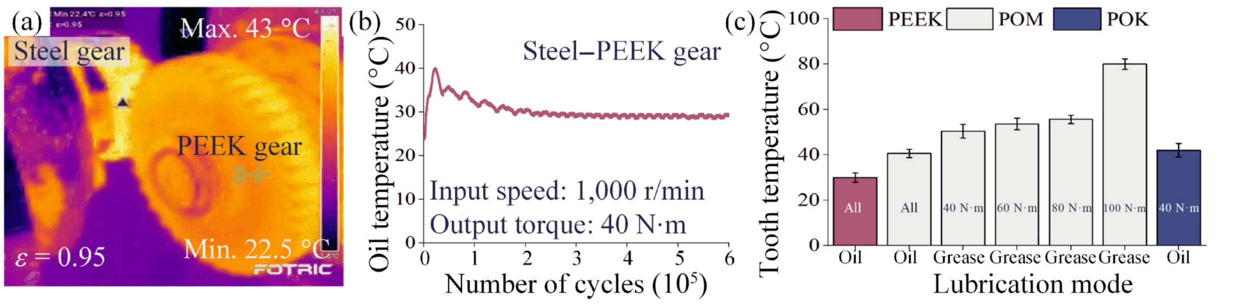


Fig. 2 Temperatures of the polymer tested gears: (a) thermal image showing the bulk temperature; (b) oil temperature; (c) tooth temperature.

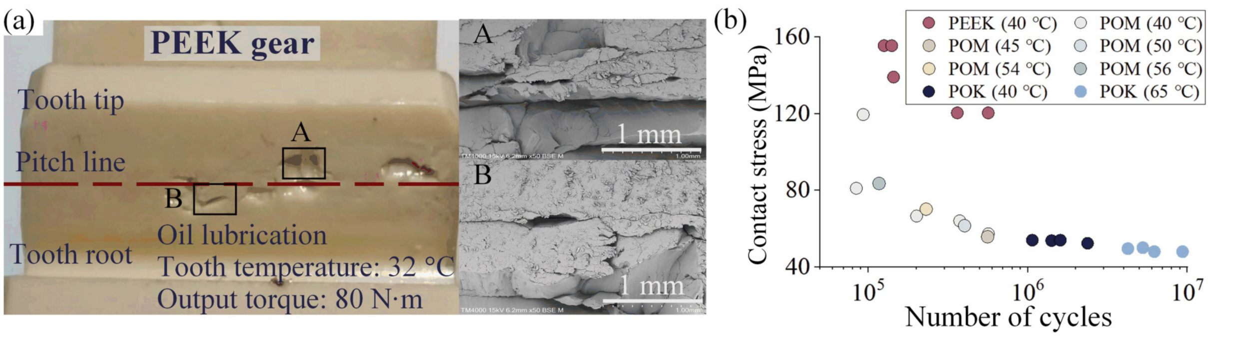


Fig. 3 Contact fatigue failure of polymer gears: (a) tooth pitting failure and (b) contact fatigue *S*–*N* curves.

Table 5 Test conditions and results

Material	Output torque (N·m)	Lubrication	Tooth temperature (°C)	Contact stress (MPa)	$\log(N_f)$
PEEK	40	Oil (oil temperature: 30 °C)	30	87.93	7.10
	60			120.40	6.66
	80			139.03	6.27
	100			155.44	6.12
	60	Oil (oil temperature: 80 °C)	80	108.21	6.01
	80			116.88	5.54
POM	40	Oil (oil temperature: 40 °C)	40	57.32	6.71
	60			70.21	6.36
	80			81.07	5.93
	100			90.63	5.32
	40	Grease (ambient temperature: 23 °C)	45	55.76	6.75
	60		50	66.61	6.30
	80		54	75.63	5.61
	100		56	83.57	5.15
POK	40	Oil (oil temperature: 40 °C)	40	53.75	6.79
	40	Grease (ambient temperature: 23 °C)	65	49.57	6.34

temperature, and median fatigue life, etc. The tooth temperature of the oil-lubricated polymer gears is independent of the output torque and is in good agreement with the oil temperature. However, grease lubrication has a weak heat dissipation capacity; thus, the tooth temperature increases gradually with increasing output torque. When the output torque increased from 40 to 100 N·m, the tooth temperature of the POM gears increased from 45 to 56 °C, whereas the median fatigue life decreased from 5.62×10^6 to 1.41×10^5 cycles.

2.2 Numerical simulation data

The thermoviscoelastic–plastic constitutive model was utilized to outline the stress–strain responses of semicrystalline polymer gears. The elastic part of the equation (Eq. (2)) assumes a linear relationship between the elastic modulus and temperature, expressed as

$$E = \alpha T + E_0 \tag{2}$$

where E_0 stands for the elastic modulus at an ambient temperature of 0 °C, MPa; T corresponds to the operating temperature, °C; α represents the modulus attenuation coefficient, MPa/°C.

The Johnson–Cook (J–C) model, a phenomenological model, has been widely applied for predicting the flow stress of metallic and polymer materials. Because strain failure was dependent on the strain rate and temperature of semi-crystalline polymers, the J–C model was utilized as follows, given by Eq. (3):

$$\sigma = \left[A + B(\varepsilon^{p})^{n}\right]\left(1 + C\ln\left(1 + \frac{\dot{\varepsilon}}{\dot{\varepsilon}_0}\right)\right)(1 - T^{*m}) \tag{3}$$

where σ represents the flow stress, MPa; A is the yield stress at the reference temperature and strain rate; B stands for the strain hardening coefficient; n represents the strain hardening exponent;

ε^{P} is the true plastic strain; $\dot{\varepsilon}$ is the strain rate; $\dot{\varepsilon}_0$ is the reference strain rate; C represents the strain rate sensitivity parameter.

The temperature coefficient $T^{*\mathrm{m}}$ is calculated by Eq. (4):

$$T^{*\mathrm{m}} = \begin{cases} (T - T_{\mathrm{room}})/(T_{\mathrm{melt}} - T_{\mathrm{room}}), T > T_{\mathrm{trans}} \\ 1, \qquad\qquad\qquad\qquad\quad T < T_{\mathrm{trans}} \end{cases} \tag{4}$$

where T_{room} stands for the ambient temperature; T_{melt} represents the melting point temperature; T_{trans} corresponds to the transition temperature, which is defined as the critical temperature at which the yield stress does not change. The mechanical properties of semi-crystalline polymers, such as POM and PEEK, can be determined via this constitutive relationship, as reported by Chen et al. [28] and Lu et al. [13].

A Brown–Miller multiaxial fatigue criterion alongside the critical plane method was employed to describe the contact fatigue behavior of polymer gears [13]. It is assumed that a fatigue crack initiates along the plane with the maximum shear strain and then propagates along the normal strain direction on the critical plane. Based on the previous work, the Brown–Miller fatigue model, modified with the mean stress and lubrication effect, is expressed as Eq. (5) [29]:

$$\frac{\Delta\gamma_{\max}}{2} + \frac{\Delta\varepsilon_{\mathrm{n}}}{2} = C_1\frac{\lambda\sigma_{\mathrm{f}}^{'} - \sigma_{\mathrm{m}}}{E}(2N_{\mathrm{f}})^{k} + C_2\varepsilon_{\mathrm{f}}^{'}(2N_{\mathrm{f}})^{c} \tag{5}$$

where $\Delta\gamma_{\max}/2$ represents the maximum shear strain amplitude; $\Delta\varepsilon_{\mathrm{n}}/2$ refers to the normal strain amplitude on the critical plane; σ_{m} is the mean normal stress on the critical plane; $\sigma_{\mathrm{f}}^{'}$ and $\varepsilon_{\mathrm{f}}^{'}$ represent the fatigue strength coefficient and fatigue ductility coefficient, respectively; λ represents the lubricant coefficient, which indicates the effect of the lubrication mode, extreme-pressure additives and other lubrication-related elements on the contact fatigue life and is empirically taken as 1.2 in this work [13, 14]; k is the fatigue strength exponent; c is the fatigue ductility exponent; N_{f} is the fatigue life of the material point; C_1 and C_2 are material constants.

Several approximate relationships between the fatigue parameters and tensile strength have been suggested for metals and polymers. Among those relationships, Bäumel and Seeger's relation [30] was used, expressed as Eq. (6):

$$\sigma_{\mathrm{f}}^{'} = 1.67\sigma_{u}^{'}, \varepsilon_{\mathrm{f}}^{'} = 0.35, k = -0.095, c = -0.69 \tag{6}$$

where $\sigma_{\mathrm{u}}^{'}$ is the conventional monotonic ultimate tensile stress, MPa.

The polymer materials, including POM, PEEK, PET, and POK, exhibit an approximately linear relationship with temperature, expressed as Eq. (7):

$$\sigma_{\mathrm{u}}^{'} = \beta T + \sigma_{0}^{'} \tag{7}$$

where β corresponds to the material constant; $\sigma_{0}^{'}$ reflects the ultimate tensile stress at the reference temperature of 0 °C, MPa.

The numerical model was developed in Abaqus software (Dassault Systems Simulia Co.). Five teeth of each gear were generated in the numerical model, with the gear pair assembled to mesh with each other, as shown in Fig. 4(a). To meet the requirements of power transmissions, the output torque was set to 20, 40, 60, 80, and 100 N·m, whereas the ambient temperature was set to 20, 40, 60, 80, and 100 °C. As full stress cycles must be guaranteed for critical material points, transient temperature–displacement coupled analysis was conducted in this study with the element type specified as CPE4RT. Mesh refinement was conducted on the meshing flanks, using an element size of 0.2 mm for the flanks of the steel gear and an element size of 0.1 mm for the polymer gear, as demonstrated in Fig. 4(b). The maximum von Mises stress in the subsurface near the tooth pitch line is shown in Fig. 4(c), and the predicted contact fatigue life in Fig. 4(d), implying that tooth contact fatigue cracks sprout from the subsurface and progress to tooth pitting.

The element size was chosen based on the conducted mesh convergence test, following the h-refinement method in Fig. 5(a). Subsequently, the friction formulation of contact interfaces was solved by the penalty function method, and the friction coefficient was set to 0.04. The output torque was applied to the driven polymer gear while the rotational movement of the steel pinion was adjusted with a rotational speed of 1,000 r/min. The contact stress results from VDI 2736.2-2014 [27] and the numerical model, both of which are based on Hertzian pressure theory and assume dry contact at the interface, show little difference under different loading conditions, as shown in Fig. 5(b).

Several parameters influence the contact fatigue performance of polymer gears, such as elastic modulus, contact stress, coefficient of friction, tensile strength, and Poisson's ratio. Thirty samples were obtained through numerical simulation to establish the simulated life data. In contrast to the experimental results for POM and PEEK gears, the errors in the simulated lives were within a 3-fold scatter band, demonstrating the validity of the simulation, as shown in Figs. 6(a) and 6(b).

2.3 Life prediction via CTAB-GAN+XGBoost algorithm

The development process of the contact fatigue life prediction model according to the hybrid data was derived, as shown in Fig. 3. Experimental and simulated data were blended at a predetermined ratio to establish a hybrid dataset of the contact fatigue lives of polymer gears, and the simulated data were used to extend the parameter range of the experimental data. The

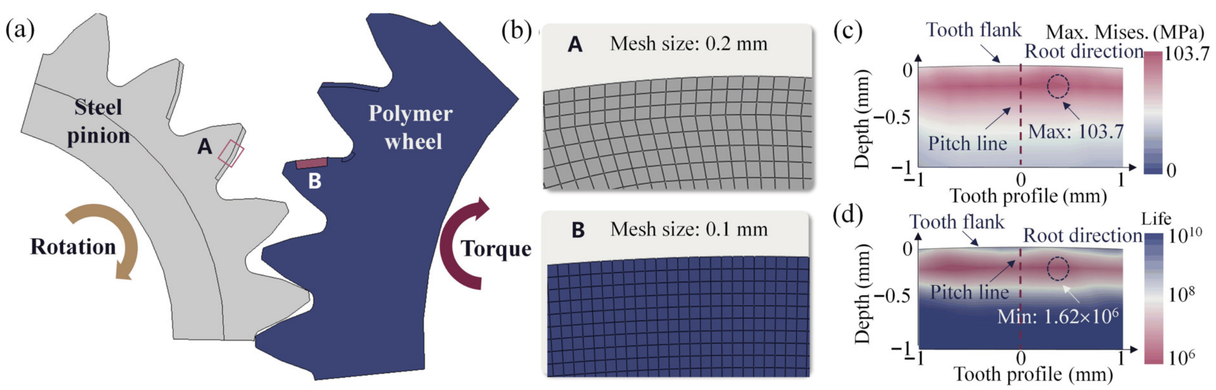


Fig. 4 Numerical model for polymer gear contact fatigue life: (a) engagement model; (b) mesh refinement; (c) von Mises stress distribution; (d) life distribution.

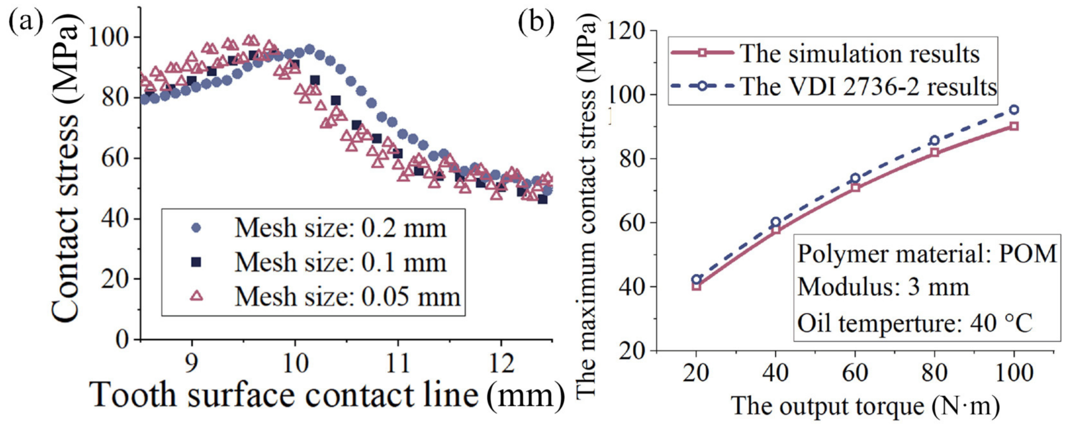


Fig. 5 Numerical model check: (a) convergence check; (b) comparison of simulated contact stress to standard calculations.

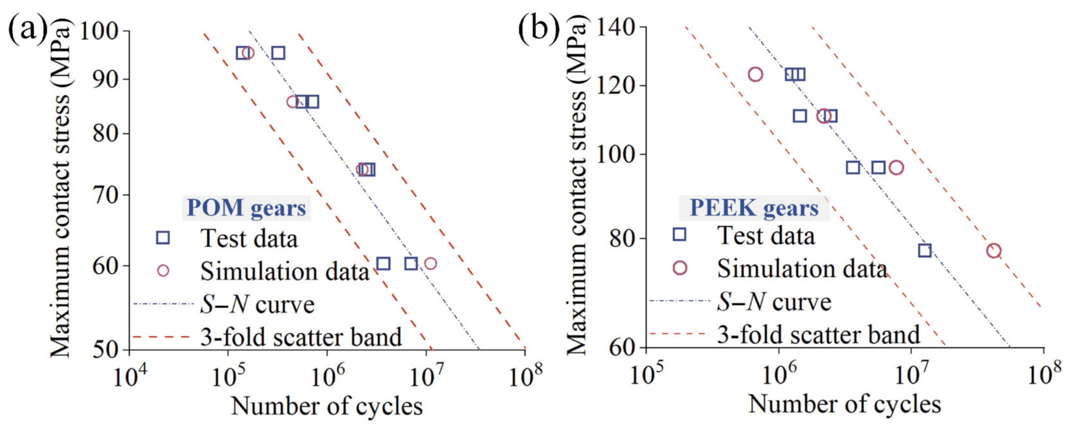


Fig. 6 Numerical results of fatigue life: (a) comparison of simulated life values with experimental results for POM gears; (b) comparison of simulated life values with experimental results for PEEK gears.

augmented data of the contact fatigue lives were effectively generated from the hybrid dataset via the CTAB-GAN algorithm. Consecutively, an XGBoost algorithm was trained on the augmented data to predict the contact fatigue life of polymer gears in terms of the elastic modulus, contact stress, coefficient of friction, tensile strength, and Poisson's ratio.

Gear fatigue data are typical tabular data. It is challenging for conventional data augmentation algorithms, such as GANs, variational autoencoders (VAEs), and adaptive synthetic sampling, to capture the long-tail distribution of tabular data, which may result in data imbalance. The CTAB-GAN algorithm, which is designed for tabular data, was utilized to augment the hybrid data of fatigue lives. Relying on the WGAN algorithm, the CTAB-GAN algorithm uses the Wasserstein loss function with a gradient penalty for more effective training convergence, expressed as Eq. (8) [31]:

$$L=\underset{\tilde{x}\sim\mathbb{P}_g}{\mathbb{E}}\left[D(\tilde{x})\right]-\underset{\tilde{x}\sim\mathbb{P}_r}{\mathbb{E}}\left[D(x)\right]+\xi\underset{\breve{x}\sim\mathbb{P}_g}{\mathbb{E}}\left[\left(\|\nabla_{\breve{x}}D\left(\breve{x}\right)\|_2-1\right)^2\right] \quad (8)$$

where L represents the Wasserstein loss function; $\mathbb{P}_g$ and $\mathbb{P}_r$ represent the generated sample distribution and the original sample distribution, respectively; $\tilde{x}$ corresponds to the joint probability distribution; x refers to the data of the original sample; $\breve{x}$ represents the data of the generator; $D(x)$ represents the neural network of the discriminator; and $\underset{\tilde{x}\sim\mathbb{P}_r}{\mathbb{E}}$ represents the expectation of the Wasserstein distance under the joint probability distribution.

The CTAB-GAN algorithm models continuous, categorical, and intermixed variables through innovative data encodings. A logarithm transformation was employed to decrease the difficulty associated with generating continuous long-tail distributions of tabular data. Detailed descriptions can be found in Ref. [32].

Convolutional neural networks with 64 feature channels were chosen for the generator and discriminator in the CTAB-GAN algorithm. The generator achieved an expansion of the space dimension for the first batch of virtual data through several upsampling and convolution operations. This process concurrently minimized the number of feature channels, ultimately generating virtual data that matched the dimensions of the hybrid data. Henceforth, the dimension of the random noise data was set as 100, which was produced by the generator. The rectified linear unit (ReLU) and batch normalization (BatchNorm) functions were applied for information transformation between layers. The output layer employed the hyperbolic tangent (Tanh) activation function, whereas the discriminator performed the convolution and down-sampling operations from data generation, and gradually increased the number of feature channels. A completely connected layer output of the discriminator was utilized to showcase the evaluated result of the virtual data. In each training iteration, the generator produced a batch of virtual data. Afterwards, the discriminator evaluated the consistency between this batch of virtual data and the hybrid data to update the parameters of the generator. Notably, with a learning rate of 2×10^{-4}, adaptive moment estimation (Adam) was used to optimize the parameter updates of the discriminator and generator. To avoid overfitting, weight decay, known as the L2 regularization factor, was employed. A decrease in this parameter would, in turn, reduce the likelihood of overfitting. In this work, the weight decay was set as 10^{-5}.

The life prediction method utilized was based on the XGBoost algorithm demonstrated by Chen and Guestrin [33]. XGBoost, a multiparameter algorithm for supervised learning, effectively utilizes an integration and superposition strategy to connect multiple decision trees into a series. The gradient boosting algorithm was employed to minimize the loss incurred by previous decision trees and promote new trees to update the predictive model iteratively until the loss function reached a threshold lower than the expected value [34]. Figure 7 outlines the framework of the XGBoost algorithm. Initially, the first tree fitted the polymer gear fatigue life data and provided a foreseen residual W_1. These predicted residuals were subsequently utilized as the target for the second tree, generating predicted residuals W_2, and so forth until the final tree produced updated predicted residuals W_k. The objective function of the XGBoost algorithm is exemplified in Eq. (9):

$$f_M(x) = \sum_{i=1}^{o} T(X, \theta_i) \tag{9}$$

where $T(X, \theta_i)$ is the decision tree, θ_i represents a parameter, and o is the tree number. The XGBoost algorithm was iterated 100 epochs, with the final prediction life obtained by aggregating the regression values from weak learners with appropriate weighting.

To compare the effects of diverse types of augmentation algorithms on the contact fatigue prediction of polymer gears, the triplet-based variational autoencoder (TVAE) algorithm was used to augment the hybrid data. The TVAE algorithm combines VAE and triplet-based metric learning alongside feature perceptual loss to learn latent representations enriched with more fine-grained data. Detailed information about the TVAE algorithm can be found in Ref. [35].

The root mean square error (RMSE) and life scatter band were applied as evaluation indicators to assess the prediction accuracy of the ML models. Moreover, the RMSE is defined in Eq. (10):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2} \tag{10}$$

where Y_i stands for the predicted life sample i, $\hat{Y}_i$ represents the predicted life sample, and n is the number of samples. Overall, a smaller RMSE indicates a smaller difference between the predicted and actual lives; hence, a higher prediction accuracy.

The life scatter band determines the deviation between the predicted life and the experimental life. In engineering practice, it proves imperative to guarantee that the fatigue life error is within a 3-fold scatter band [36]. Thus, the accuracy index A^* was specified as the probability of the predicted life within the 3-fold scatter band, expressed as Eq. (11):

$$A^* = \frac{n_A}{n} \times 100\% \tag{11}$$

where n_A is the number of predicted life samples within the 3-fold scatter band.

3 Results and discussion

3.1 Comparison of different ML models

To demonstrate the advantages of the present prediction model via the CTAB-GAN+XGBoost algorithm, this section offers a detailed comparison of the life results of numerous ML algorithms on hybrid data. There are 10 experimental data samples of PEEK and POM gears and 30 simulated data samples under different loads and temperatures, with the validation data consisting of 19 experimental data samples of PEEK, POM, PET, and POK gears.

The TVAE and CTAB-GAN algorithms were chosen to augment the hybrid data. Figure 8(a) shows the distributions of hybrid and augmented data using the TVAE and CTAB-GAN

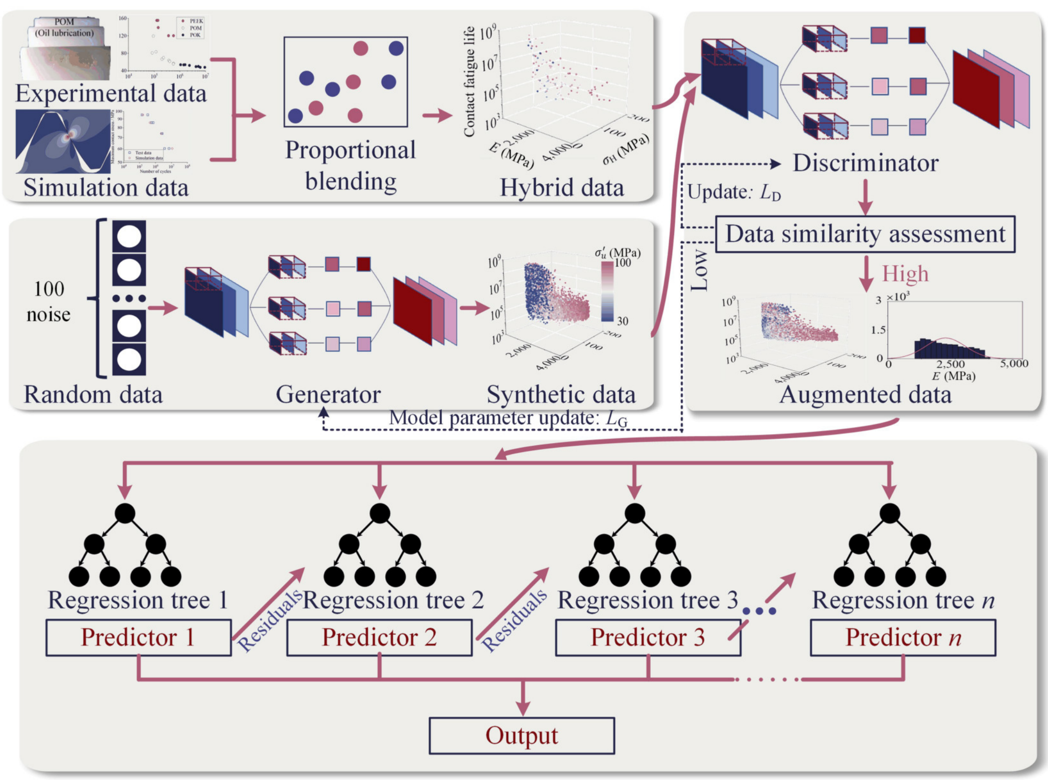


Fig. 7 Contact life prediction model with CTAB-GAN+XGBoost algorithm.

algorithms. The distribution of augmented data samples equals 10,000, and the mixture ratio of the hybrid data between the experimental and simulated data is set as 1:3. The distribution of the augmented data via the CTAB-GAN algorithm is comparable to that of the hybrid data. This similarity arises since the contact fatigue life decreases as the elastic modulus increases. The correlation between the elastic modulus and contact fatigue life in the augmented data generated by the TVAE algorithm stands apart from that observed in the hybrid data. This discrepancy may negatively influence the prediction accuracy of ML algorithms.

As shown in Figs. 8(b)–8(d), the mean value and standard deviation (STD) of the hybrid and augmented data were obtained using a normal distribution function for elastic modulus, maximum contact stress, and tensile strength. The comparison indicates that the similarity between the augmented data via the CTAB-GAN algorithm and the hybrid data is greater than that of the augmented data via the TVAE algorithm. With respect to the statistical characterization of the elastic modulus, the mean value of the augmented data via the CTAB-GAN algorithm exhibited a deviation of 11.8% from the hybrid data, whereas the STD of the augmented data via the CTAB-GAN algorithm differed by 0.6% from the hybrid data. Similarly, the mean value of the augmented data via the TVAE algorithm deviated by 17.3% from the hybrid data, exhibiting an STD difference of 13.4% between the augmented data and the hybrid data.

Additionally, the STD of the tensile strength of the augmented data via the CTAB-GAN algorithm differed from that of the hybrid data by only 1.2%. In contrast, the STD of the tensile strength of the augmented data via the TVAE algorithm differed from that of the hybrid data by a significant 30%. Thus, the CTAB-GAN algorithm is better suited than the TVAE algorithm for addressing the statistical characteristics of hybrid data.

For the construction of life prediction models, the XGBoost algorithm, random forest (RF) algorithm, and support vector machine (SVM) algorithm were utilized. The prediction accuracy of those ML models was evaluated considering various data, as shown in Fig. 9. The CTAB-GAN+XGBoost model has the highest prediction accuracy for the contact fatigue life of polymer

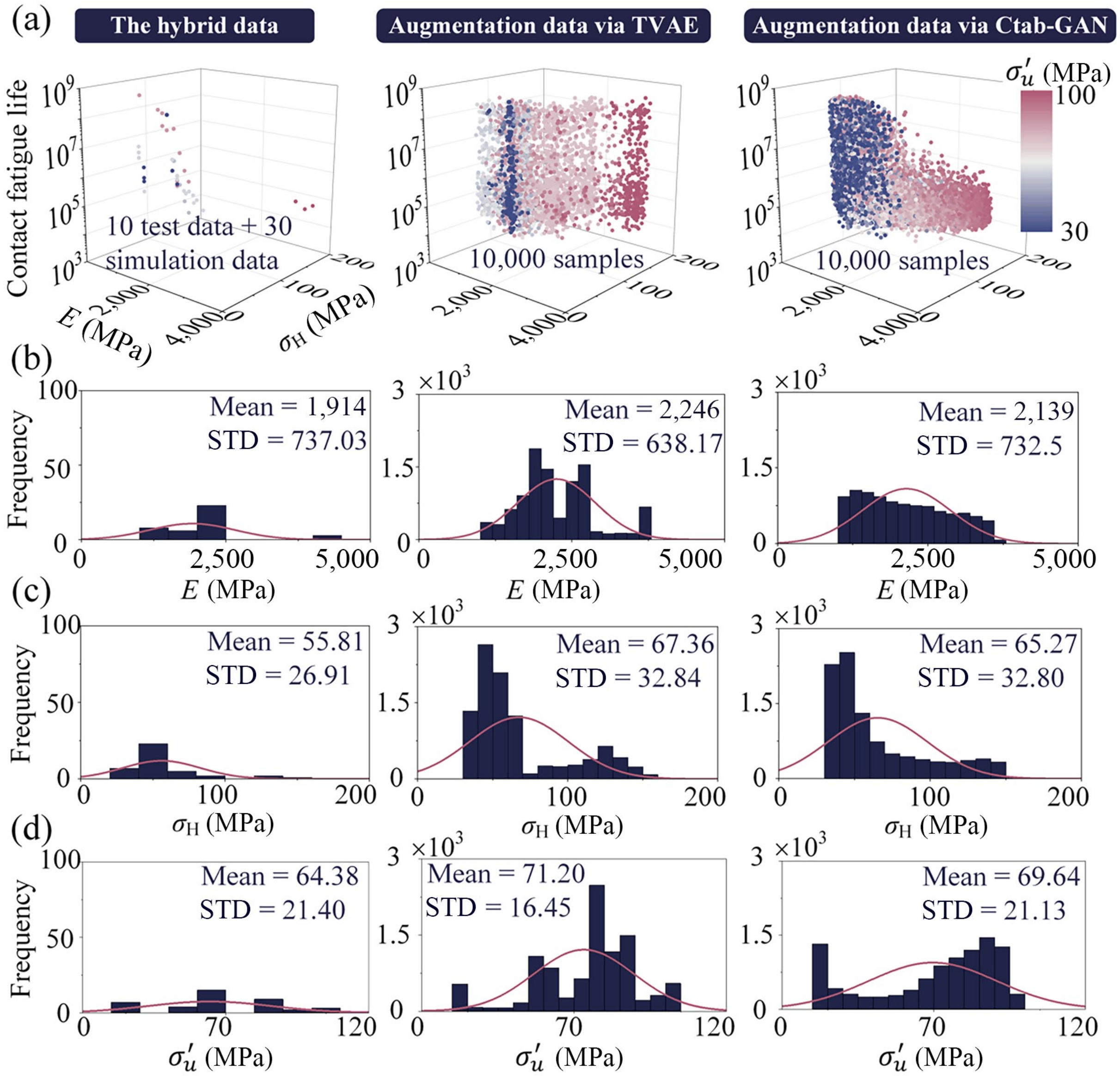


Fig. 8 Comparison of hybrid data with a 1:3 mixture ratio and augmented data via the TVAE and CTAB-GAN algorithms: (a) fatigue life distribution; (b) elastic modulus distribution; (c) contact stress distribution; (d) tensile strength distribution.

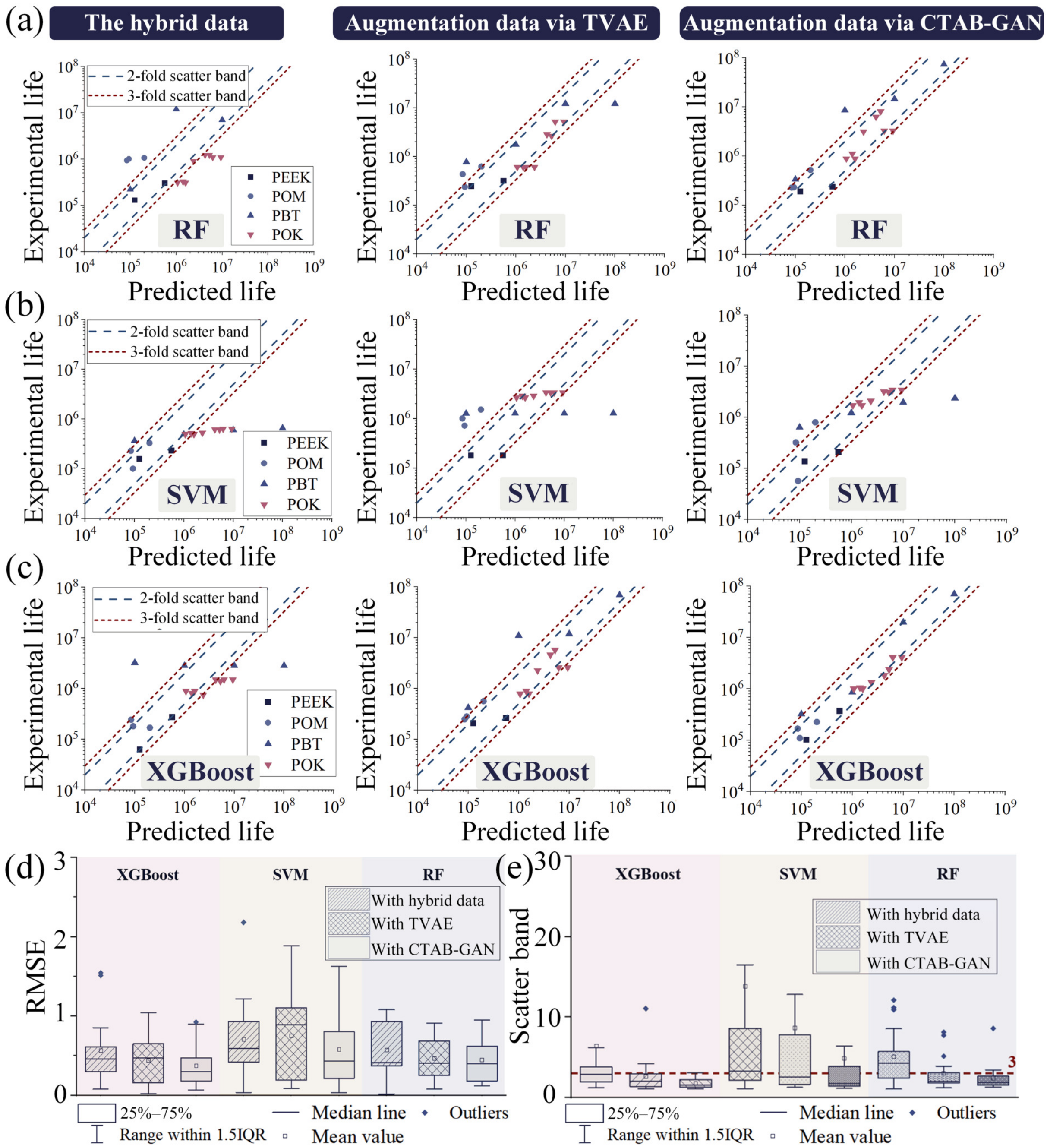


Fig. 9 Comparison of prediction accuracy of different models: (a) RF; (b) SVM; (c) XGBoost; (d) RMSE; (e) comparison of the lifetime scatter band.

gears in the context of this research. The predicted data for the POM, PEEK, PET, and POK gears deviate from the experimental data by no more than a 3-fold scatter band, as shown in Figs. 9(a)–9(c). Figures 9(d) and 9(e) present a comparison of the life prediction models following their RMSE and life scatter band, respectively. The mean RMSE of the CTAB-GAN+XGBoost model is 0.372, which is 52.6% lower than that of the XGBoost model and 18.9% lower than that of the TVAE+XGBoost model. This finding highlights the feasibility of utilizing the CTAB-GAN algorithm for data augmentation.

Moreover, the prediction of contact fatigue life for POK and PET gears via hybrid data continuously poses a substantial challenge for the XGBoost, RF, and SVM algorithms. This is evident from the predicted contact fatigue life exceeding the 3-fold scatter band, indicating a significant deviation. This implies that the incorporation of simulated data alone proves insufficient to improve the generalization of the life prediction algorithms, whereas the application of data augmentation algorithms could effectively control the prediction errors of contact fatigue life. The A^* within the TVAE+RF model is 70.6%, whereas the A^* of the TVAE+SVM model achieves 58.8%.

In contrast, the A^* of the TVAE+XGBoost model is 82.3%. The XGBoost algorithm achieves the highest prediction accuracy among the three prediction algorithms because of its ability to capture the nonlinear relationship characteristics of the fatigue data. The XGBoost algorithm is an integrated approach that incorporates the regularization and a second-order Taylor expansion of the loss function, benefitting the robustness of the algorithm to noisy samples and outliers in augmented data and ensuring enhanced prediction accuracy.

3.2 Effects of model parameters on prediction accuracy

The mixture ratio of hybrid data plays a critical role in the prediction accuracy of the CTAB-GAN+XGBoost model, as shown in Fig. 10. Notable differences can be seen in the space of fatigue life data and elastic modulus distributions of the augmented data with diverse mixture ratios, as shown in Figs. 10(a) and 10(b). The mean elastic modulus of the augmented data with a 1:1 mixture ratio equals 2,152 MPa, which closely aligns with the mean elastic modulus of 1,914 MPa in the hybrid data.

Nevertheless, there are pronounced imbalances in the augmented data with a 1:1 mixture ratio, leading to significant aggregation around the mean elastic modulus. This imbalance stems from the limited amount of simulated data available, which constrains the coverage of the data augmentation process when a 1:1 mixture ratio is used. Alternatively, the augmented data with a 1:2 mixture ratio include numerous discrete samples in the space of fatigue life data because of the insufficient quantity of simulated data, degrading the augmented data quality. Figure 10(c) shows the prediction accuracy of the contact fatigue life for various mixture ratios. As such, the prediction accuracy is 47% for the 1:1 mixture ratio, increasing to 58.8% for the 1:2 mixture ratio and becoming completely controllable within the 3-fold scatter band for the 1:3 mixture ratio.

A comparison was performed to evaluate the prediction accuracy of the CTAB-GAN+XGBoost model under various mixture ratios, following five repetitions of the augmentation process, as shown in Figs. 10(d)–10(f). The model prediction performance significantly improves as the quantity of simulated data increases. The prediction accuracy of the model constructed

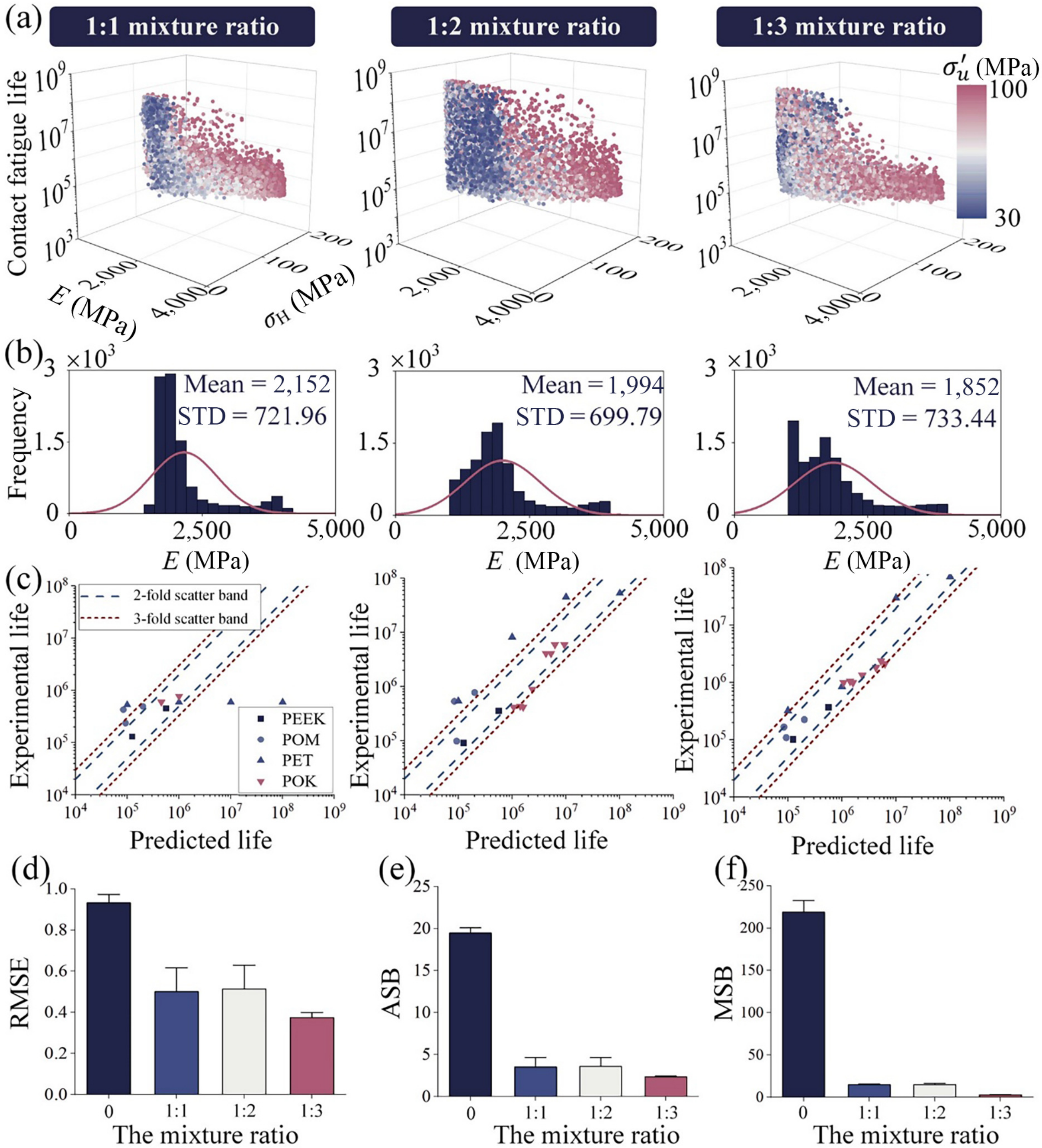


Fig. 10 Effects of mixture ratio on model performance: (a) fatigue life data spaces; (b) elastic modulus distributions; (c) comparison of model prediction accuracies; (d) comparison of RMSEs; (e) comparison of ASBs; (f) comparison of MSBs.

with augmented data without any simulated data is the lowest. In this case, the RMSE, average scatter band (ASB), and maximum scatter band (MSB) are calculated as 0.93, 19.45, and 218.76, respectively. Nevertheless, when the simulated data are increased and the mixture ratio is decreased from 1:1 to 1:3, the RMSE decreases from 0.5 to 0.37. Interestingly, with a mixture ratio of 1:3, the MSB decreases significantly to only 2.96, representing a mere 1.4% of the ASB observed in the augmented data without simulated data.

The number of augmented data is a vital factor that affects the prediction accuracy of the CTAB-GAN+XGBoost model. The distribution of diverse quantities of augmented data with a mixture ratio of 1:3 in the space of fatigue life data is shown, along with the distributions of the elastic modulus, in Figs. 11(a) and 11(b). Considering the mean value of the elastic modulus, the error between 100 augmented data and the hybrid data is 14.9%, and between 10,000 augmented data and the hybrid data is 0.8%. As such, the amount of augmented data increases with increasing data quantity. The prediction accuracy of the contact fatigue life with different quantities of augmented data is displayed in Fig. 11(c). With the 100 augmented data, the A^* of the CTAB-GAN+XGBoost model represents the contact fatigue life for PET gears and POK gears as 33.3%, whereas that of the POM gears and PEEK gears is controlled within a 3-fold scatter band. With an increase in the number of augmented data to 1,000, the A^* of the CTAB-GAN+XGBoost model increases to 50% for the contact fatigue life of the PET and POK gears. It is challenging for the CTAB-GAN+XGBoost model to accurately predict the contact

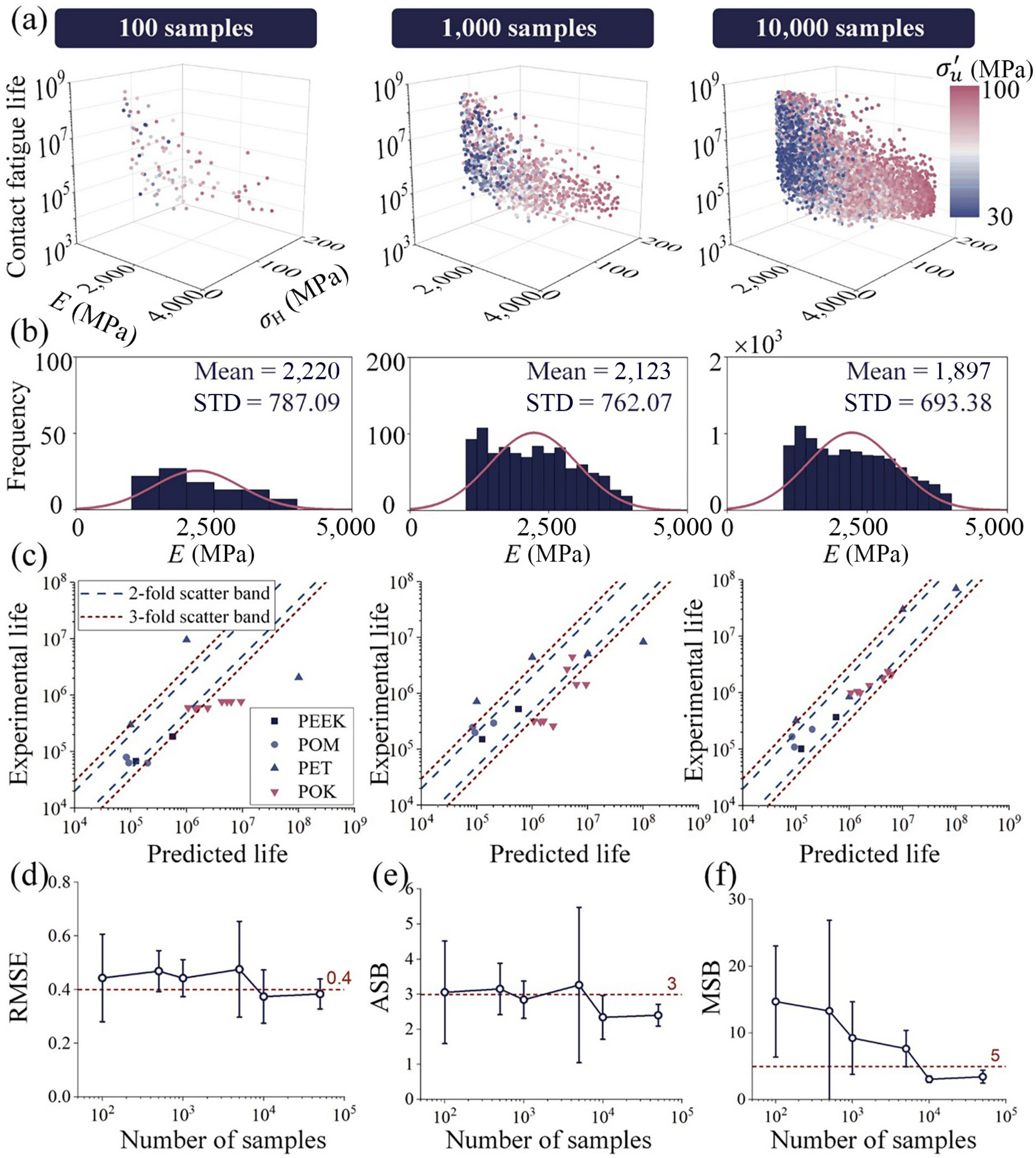


Fig. 11 Effects of the number of augmented data on model performance: (a) fatigue life data spaces; (b) elastic modulus distributions; (c) comparison of model prediction accuracies; (d) comparison of RMSEs; (e) comparison of ASB; (f) comparison of MSB.

fatigue life of polymer gears with ultra-limited augmented data. In addition, Figs. 11(d)–11(f) demonstrate the correlation between the prediction accuracy and the number of augmented data.

3.3 Prediction formula for fatigue life

To aid engineering applications, a formula for predicting the contact fatigue life of polymer gears has been derived with the help of the CTAB-GAN+XGBoost model. The influence and contribution of material parameters and working conditions on the contact fatigue life of polymer gears were specified using 10,000 augmented data with a 1:3 mixture ratio. Figure 12(a) shows the influence of elastic modulus, maximum contact stress, friction coefficient, tensile strength, and Poisson's ratio on the fatigue life of polymer gears, as per the SHapley Additive exPlanations (SHAP). Among these factors, the maximum contact stress is the leading factor in determining the fatigue life of polymer gears. Moreover, the maximum contact stress has a negative effect on the fatigue life of gear contact, which corresponds to the observations in engineering practice. Another critical effect on contact fatigue life is the elastic modulus and tensile strength. The elastic modulus represents a negative effect on the contact fatigue life of polymer gears, while the tensile strength has a complex non-linear effect. As shown in Fig. 12(b), the contributions are ranked following the SHAP value. The contributions of the maximum contact stress, elastic modulus, and tensile strength are 0.483, 0.198, and 0.147, respectively. Conversely, the contributions of coefficient of friction and Poisson's ratio are 0.0944 and 0.0779, respectively, indicating that these impacts are negligible.

An empirical formula for the fatigue life of polymer gears was derived, ignoring the effects of Poisson's ratio and coefficient of friction. Owing to its simplicity, Basquin's formula (Eq. (12)) is widely utilized in fatigue life calculations, like ISO 12107.2021 [37] and ISO 6336.1-2019 [38], to describe the S–N curves:

$$S^m N = C^* \tag{12}$$

where S represents the stress, MPa; N refers to the fatigue life; m and C^* represent the material parameters.

Basquin's formula does not consider the effects of the elastic

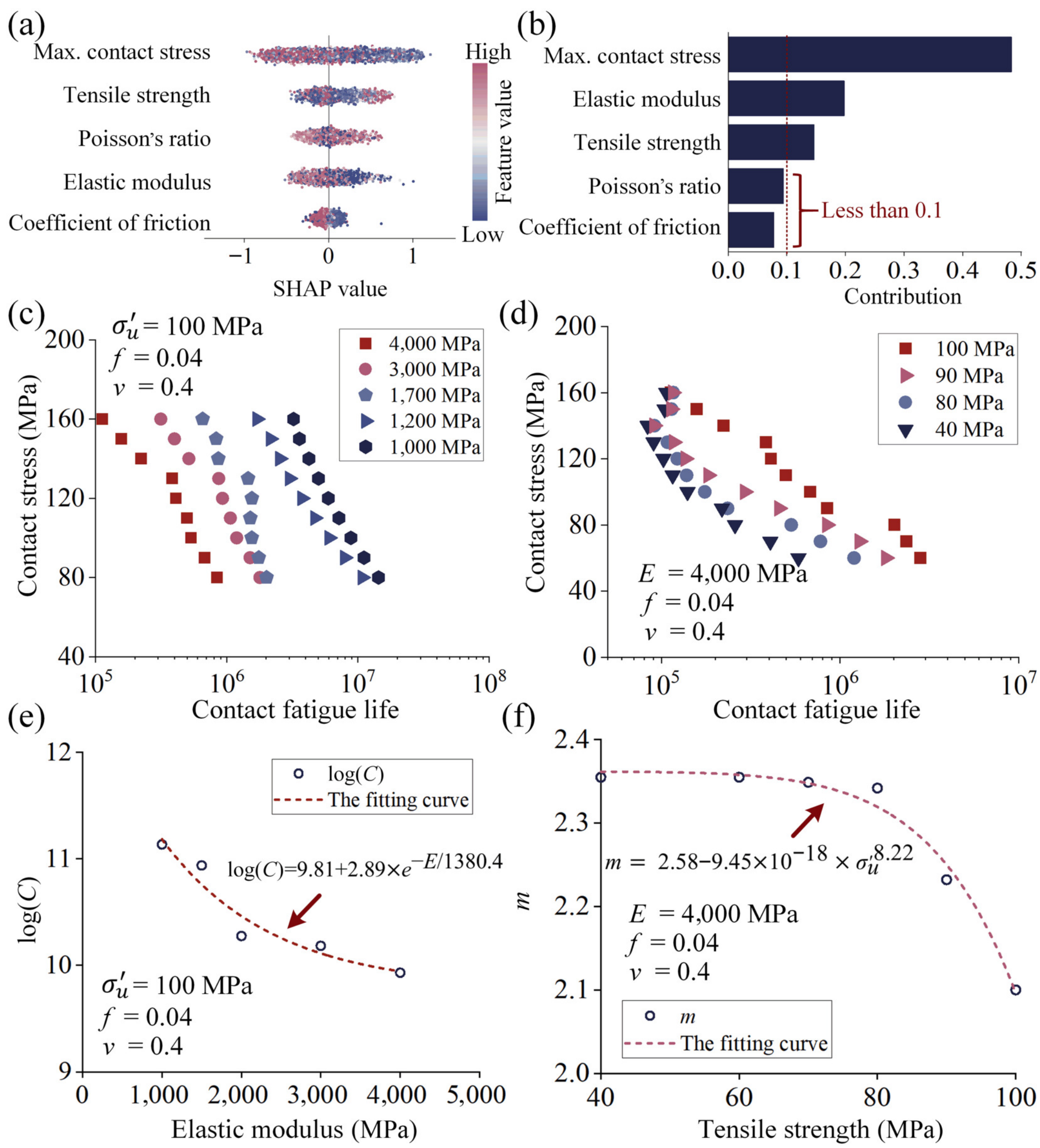


Fig. 12 Influence of material properties and loading conditions: (a) SHAP analysis; (b) feature contributions; (c) effect of elastic modulus; (d) effect of tensile strength; (e) fitted equation for material coefficient log(C); (f) fitted equation for material coefficient m.

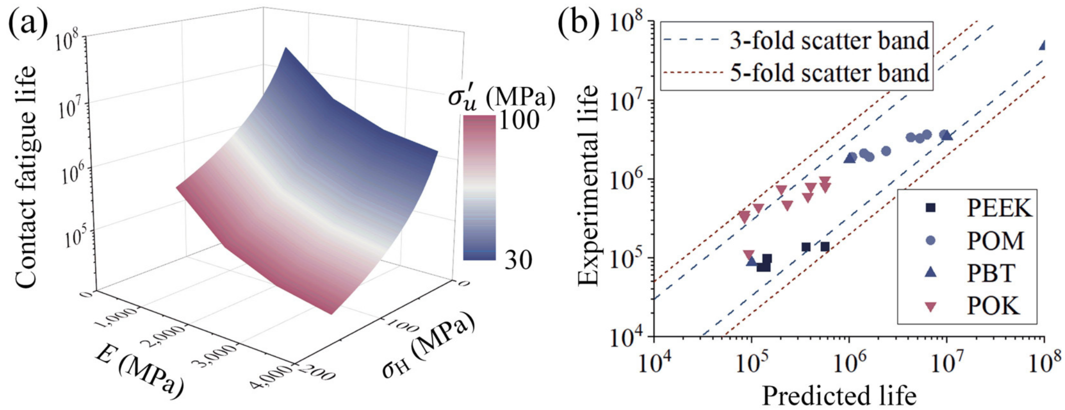


Fig. 13 Prediction formula for contact fatigue life of polymer gears: (a) distribution and (b) analysis of prediction accuracy comparison.

modulus and tensile strength on the fatigue life. Figures 12(c) and 12(d) present the effects of the elastic modulus and tensile strength on the *S*–*N* curves according to the CTAB-GAN+XGBoost model. With a specific combination of tensile strength, friction coefficient, and Poisson's ratio, the contact fatigue life of polymer gears with an elastic modulus of 1,000 MPa is approximately 10^7 cycles, whereas that with an elastic modulus of 4,000 MPa is less than 10^6 cycles at a contact stress of 80 MPa.

Nevertheless, considering a specific combination of elastic modulus, friction coefficient, and Poisson's ratio, the contact fatigue life of polymer gears with a tensile strength of 40 MPa amounts to approximately 3×10^5 cycles, while those with a tensile strength of 100 MPa exhibit a higher contact fatigue life of 2×10^6 cycles under a contact stress of 80 MPa. As the contact stress increases, the differences in contact fatigue life between polymers with varying tensile strengths decrease. A linear relationship was observed between the elastic modulus and fatigue life, whereas a non-linear relationship existed between the tensile strength and fatigue life.

Based on the data in Fig. 12(c), alongside the determination of log(C*) and its fitting curve, m proves constant in Basquin's formula through Newton's iterative method, as shown in Fig. 12(e). Similarly, m is determined as per Fig. 12(f). The fitting equation (Eq. (13)) can be expressed as

$$\begin{cases} C^* = 10^{9.81+2.89\times \mathrm{e}^{-E/1380.4}} \\ m = 2.58 - 9.45\times10^{-18}\times\sigma_{\mathrm{u}}^{\prime 8.22} \end{cases} \tag{13}$$

Like Figs. 10(a) and 11(a), Fig. 13(a) also shows the distribution characteristics in the feature space of the proposed formula, while Fig. 13(b) indicates a comparison of predicted lives via the formula to experimental results. As the effects of Poisson's ratio and coefficient of friction on the contact fatigue life of polymer gears were neglected, the prediction error was controlled within a 5-fold scatter band. This is acceptable for design reference in engineering.

4 Conclusions

A strategy for contact fatigue life prediction was constructed following a hybrid process of experimental and simulated data of polymer gears via a CTAB-GAN+XGBoost algorithm. Moreover, an empirical formula was suggested for fatigue life prediction of polymer gears considering the effects of the maximum contact stress, elastic modulus, and tensile strength. The primary conclusions were as follows:

(1) A hybrid data-driven model via a CTAB-GAN+XGBoost algorithm was proposed according to experimental and simulated life data. It achieved an acceptable generalization capability of life prediction for polymer gears with small experimental samples, outperforming other algorithms within a 3-fold scatter band.

(2) The contact fatigue performance of polymer gears was significantly affected by the maximum contact stress, elastic modulus, and tensile strength. The contributions of these factors were 0.483, 0.198, and 0.147, respectively. Notably, the coefficient of friction and Poisson's ratio had negligible effects on the performance.

(3) In light of the elastic modulus and tensile strength of polymer materials, a fatigue life prediction formula for polymer gears was derived. For PEEK, POM, PET, and POK gears, the prediction error was maintained within a 5-fold scatter band, which was acceptable and convenient for engineering purposes.

Further work is needed to consider the effects of fiber reinforcement, humidity, and other factors on the contact fatigue performance of polymer gears. To promote the replacement of steel with polymers in the field of power transmission, it is critical to further study the bending fatigue of polymer gears.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (Nos. 52275050 and 52175041).

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

References

[1] Lu Z H, Liu C, Liao C J, Zhu J Z, Liu H J, Chen Y M. Conceptual design and optimization of polymer gear system for low-thrust turbofan aeroengine accessory transmission. *J Comput Des Eng* **11**(1): 212–229 (2023)

[2] Reitschuster S, Tobie T, Stahl K. Experimental verification of high-performance polymer gears in an electric vehicle powertrain. *Forsch Ingenieurwes* **87**: 881–890 (2023)

[3] Demšar I, Černe B, Tavčar J, Vukašinović N, Zorko D. Agile development of polymer power transmission systems for e-mobility: A novel methodology based on an e-bike drive case study. *Polymers* **15**(1): 68 (2023)

[4] Mao K, Hooke C J, Walton D. Acetal gear wear and performance prediction under unlubricated running condition. *J Synth Lubr* **23**(3): 137–152 (2006)

[5] Mao K, Li W, Hooke C J, Walton D. Friction and wear behaviour of acetal and nylon gears. *Wear* **267**(1–4): 639–645 (2009)

[6] Senthilvelan S, Gnanamoorthy R. Damage mechanisms in injection molded unreinforced, glass and carbon reinforced nylon 66 spur gears. *Appl Compos Mater* **11**(6): 377–397 (2004)

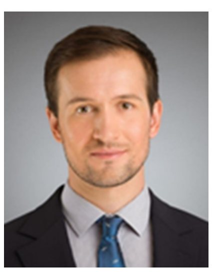
[7] Senthilvelan S, Gnanamoorthy R. Effect of gear tooth fillet radius on

the performance of injection molded nylon 6/6 gears. *Mater Des* **27** (8): 632–639 (2006)
[8] Hasl C, Illenberger C, Oster P, Tobie T, Stahl K. Potential of oil-lubricated cylindrical plastic gears. *J Adv Mech Des Syst* **12**(1): JAMDSM0016 (2018)
[9] Lu Z H, Liu H J, Zhu C C, Song H L, Yu G D. Identification of failure modes of a PEEK–steel gear pair under lubrication. *Int J Fatigue* **125**: 342–348 (2019)
[10] Zhong B B, Song H L, Liu H J, Wei P T, Lu Z H. Loading capacity of POM gear under oil lubrication. *J Adv Mech Des Syst* **16**(1): JAMDSM0006 (2022)
[11] Illenberger C M, Tobie T, Stahl K. Damage mechanisms and tooth flank load capacity of oil-lubricated peek gears. *J Appl Polym Sci* **139** (30): e52662 (2022)
[12] Illenberger C, Tobie T, Stahl K. Operating behavior and performance of oil-lubricated plastic gears. *Forsch Ingenieurwes* **86**: 557–565 (2021)
[13] Lu Z H, Liu H J, Zhang R H, Zhu C C, Shen Y Q, Xin D. The simulation and experiment research on contact fatigue performance of acetal gears. *Mech Mater* **154**: 103719 (2021)
[14] Lu Z H, Liu H J, Wei P T, Zhu C C, Xin D, Shen Y Q. The effect of injection molding lunker defect on the durability performance of polymer gears. *Int J Mech Sci* **180**: 105665 (2020)
[15] Chen D F, Zhu J Z, Liu H J, Wei P T, Mao T Y. Experimental investigation of the relation between the surface integrity and bending fatigue strength of carburized gears. *Sci China Technol Sc* **66** (1): 33–46 (2023)
[16] Mao T Y, Liu H J, Wei P T, Chen D F, Zhang P, Liu G Q. An improved estimation method of gear fatigue strength based on sample expansion and standard deviation correction. *Int J Fatigue* **161**: 106887 (2022)
[17] Zhang X H, Wei P T, Parker R G, Liu G L, Liu H J, Wu S J. Study on the relation between surface integrity and contact fatigue of carburized gears. *Int J Fatigue* **165**: 107203 (2022)
[18] Zhang X C, Gong J G, Xuan F Z. A deep learning based life prediction method for components under creep, fatigue and creep-fatigue conditions. *Int J Fatigue* **148**: 106236 (2021)
[19] He L, Wang Z L, Akebono H, Sugeta A. Machine learning-based predictions of fatigue life and fatigue limit for steels. *J Mater Sci Technol* **90**: 9–19 (2021)
[20] Kishino M, Matsumoto K, Kobayashi Y, Taguchi R, Akamatsu N, Shishido A. Fatigue life prediction of bending polymer films using random forest. *Int J Fatigue* **166**: 107230 (2023)
[21] Zhang Z Y, Liu Q Y, Wu D Z. Predicting stress–strain curves using transfer learning: Knowledge transfer across polymer composites. *Mater Des* **218**: 110700 (2022)
[22] Li Y, Wei P T, Xiang G, Jia C F, Liu H J. Gear contact fatigue life prediction based on transfer learning. *Int J Fatigue* **173**: 107686 (2023)
[23] Jia C F, Wei P T, Lu Z H, Ye M, Zhu R, Liu H J. A novel prediction approach of polymer gear contact fatigue based on a WGAN-XGBoost model. *Fatigue Fract Eng M* **46**(6): 2272–2283 (2023)
[24] CN-GB. GB/T 10095.1 Cylindrical gears—ISO system of flank tolerance classification—Part 1: Definitions and allowable values of deviations relevant to flanks of gear teeth. GB, 2022.
[25] CN-GB. GB/T 14229 Test method of surface contact strength for gear load capacity. GB, 2021.
[26] Zorko D. Investigation on the high-cycle tooth bending fatigue and thermo-mechanical behavior of polymer gears with a progressive curved path of contact. *Int J Fatigue* **151**: 106394 (2021)
[27] DE-VDI. VDI 2736-2: 2014 Thermoplastic gear wheels–cylindrical gears–calculation of the load-carrying capacity. Association of German Engineers (VDI), 2014.
[28] Chen F, Ou H G, Lu B, Long H. A constitutive model of polyether-ether-ketone (PEEK). *J Mech Behav Biomed* **53**: 427–433 (2016)
[29] Zhang B Y, Liu H J, Zhu C C, Ge Y B. Simulation of the fatigue-wear coupling mechanism of an aviation gear. *Friction* **9**(6): 1616–1634 (2021)
[30] Baümel A, Seeger T, Boller C. *Materials Data for Cyclic Loading—Supplement 1*. New York (USA): Elsevier, 1990.
[31] Gulrajani I, Ahmed F, Arjovsky M, Dumoulin V, Courville A. Improved training of wasserstein GANs. *arXiv preprint* arXiv: 1704.00028 (2017)
[32] Zhao Z L, Kunar A, Birke R, Scheer H V D, Chen L Y. CTAB-GAN+: Enhancing tabular data synthesis. *Front Big Data* **6**: 1–12 (2022)
[33] Chen T Q, Guestrin C. XGBoost: A scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016: 785–794.
[34] Zhong J C, Sun Y S, Peng W, Xie M Z, Yang J H, Tang X W. XGBFEMF: An XGBoost-based framework for essential protein prediction. *IEEE T NanoBiosci* **17**(3): 243–250 (2018)
[35] Ishfaq H, Liu R. TVAE: Deep metric learning approach for variational autoencoder. *arXiv preprint* arXiv: 1802.04403 (2018)
[36] Zhang B Y, Liu H J, Zhu C C, Li Z. Numerical simulation of competing mechanism between pitting and micro-pitting of a wind turbine gear considering surface roughness. *Eng Fail Anal* **104**: 1–12 (2019)
[37] CH-ISO. ISO 12107: 2012 Metallic materials—Fatigue testing—Statistical planning and analysis of data. ISO, 2012.
[38] CH-ISO. ISO 6336-1: 2019 Calculation of load capacity of spur and helical gears—Part 1: Basic principles, introduction and general influence factors. ISO, 2019.

Zehua Lu is currently working as an assistant researcher at the State Key Laboratory of Mechanical Transmission for Advanced Equipment, Chongqing University, China. He received his Ph.D. degree in mechanical engineering from Chongqing University, China, in 2024. His current research focuses on the lightweight design of gear transmissions via polymers.

Huaiju Liu is currently working as a professor in Chongqing University, China. He received his Ph.D. degree from the University of Warwick, UK, in 2013. His research fields include intelligent design of mechanical transmissions.

Peitang Wei is currently working as a professor in Chongqing University, China. He received his Ph.D. degree from the University of Wollongong, Australia, in 2015. His research fields include tribology and fatigue behaviors of mechanical elements.



Damijan Zorko is currently working as a CEO at RD Motion d.o.o., Ljubljana, Slovenia. He received his Ph.D. degree from the University of Ljubljana, Slovenia, in 2019. His research fields include tribology performance of polymer gears.