A High-Temperature Sliding Contact Fatigue Life Prediction Model Based on Surface Integrity

Beibei Liu*

Xi'AN MINGDE INSTITUTE OF TECHNOLOGY, Xi'an, 710124, China
Corresponding author: 18502973340@163.com

Abstract: *Fatigue failure of mechanical components at high temperatures and in the presence of sliding contact is a serious problem in the field of aero-engine and advanced gas turbine main equipment. Traditional life prediction models generally assume that the material is a homogeneous body and do not fully consider the significant influence of the surface integrity state, which is determined by the joint effect of manufacturing and service, on fatigue behaviour; thus, there is a large spread in the prediction results. Therefore, the aim of this paper is to construct a physics-mechanism-based life prediction model that can take into account all kinds of attributes of surface damage. The above are the ways to determine the shape of the surface, residual stress and microstructure, as well as the change laws of these under high-temperature sliding contact. It has been shown that many kinds of damage occur in thermo-mechanical coupling and creep-fatigue interaction. According to the theory of multi-scale damage coupling, this paper proposes a general system to quantify all kinds of parameters in the damage evolution equation related to the surface condition and develops a numerical solution algorithm that combines the finite element method with damage mechanics. The model can simulate all the changes in the bad surface condition and the final failure, so it is a very good theoretical tool for studying anti-fatigue design and the reliability of high-temperature sliding contact parts.*

Keywords: *Surface Condition, High-temperature sliding contact, Fatigue life prediction, Damage mode, Multi-scale modelling, Thermo-mechanical coupling*

Introduction

At high temperatures and under the influence of sliding contact, the fatigue life of mechanical components will be reduced to some extent due to the condition of the surface at the joint caused by manufacturing processes and the initial service load. Roughly speaking, the condition of the surface is a combination of roughness and residual stress at the surface, and this directly affects where and how quickly a fatigue crack will start and spread. Most of the previous research has focused on only one factor or room-temperature conditions, and there is no systematic list of the multiple parameters that affect the development of surface integrity at high temperatures and how they interact with all kinds of damage mechanisms. Therefore, the old life-prediction model based on the stress-strain method is not very suitable for high-temperature contact fatigue. Therefore, we need to build a physically predictive model that can consider all kinds of attributes at different times and from all sides, properly account for the special thermo-mechanical coupling and time-dependent damage mechanisms in high-temperature environments, advance the basic knowledge of fatigue theory, and meet the urgent engineering demand for accurate life design and improved reliability of key components in high-parameter equipment. The purpose of this paper is to achieve the above goal.

1. Characterization of Surface Integrity and Its Effect Mechanism on Fatigue Behaviour

1.1 Characterization Parameters of Surface Topography Features under High-Temperature Sliding Contact

High-temperature sliding contact is a process that changes the shape of the material considerably, so we need to know its properties accurately to determine the condition of the surface. The first three are: three-dimensional surface roughness, arithmetic mean deviation of the profile, and maximum height of the profile; the fourth is: directionality of the surface texture. Due to the high-temperature environment, there is a combined effect of plastic deformation, material transfer and oxidation at the contact surface, so the traditional two-dimensional roughness parameters are not very representative. According to the scale-dependent parameters of fractal theory and analysis of the bearing area curve, we can better quantify the distribution and development of surface roughness under thermal-mechanical load. Topographical features directly affect the local concentration degree of the contact stress field, the retention capacity of the lubricating medium and the start location of microcracks; thus, they need to be taken into account in the construction of the initial conditions for fatigue damage. In order to obtain accurate dynamic data on changes in the high-temperature environment, it is necessary to combine the method of in-situ measurement with a non-contact 3D optical profilometer to obtain real-time topographical development data of wear. Based on characterisation, it is determined how much the topography and local stress concentration contribute; an approximate geometry is obtained and the starting point for damage in the following multi-scale models is set[1].

1.2 Formation and Evolution Law of the Residual Stress State in the Surface Layer

The state of residual stress at the surface is an internal stress equilibrium system caused by the joint effect of non-uniform plastic deformation, thermal gradient and phase change in high-temperature sliding contact. It is the plastic strain gradient caused by a mechanical load and the thermal strain mismatch due to a difference in the coefficient of thermal expansion at high temperatures. Under cyclic sliding contact, the residual stress is not constant, and the law of change for it is affected by the combined influence of load, sliding velocity, ambient temperature and the number of cycles. The first residual compressive stress can delay the start of a fatigue crack, but due to repeated thermo-mechanical coupled loads, there will be stress relaxation and redistribution; therefore, the sign of the stress field may change or its magnitude may decrease. In order to present an accurate description of the development process of this phenomenon, we need to introduce a constitutive model and the effect of cyclic softening and hardening to improve the prediction accuracy of fatigue life.

1.3 Correlation Analysis of Microstructure State and Surface Integrity

The shape and mechanical condition of the surface, along with the microstructure of the material in the subsurface region, are collectively known as surface integrity. High-temperature sliding contact can result in grain refinement or grain coarsening of the microstructure in the area of the surface, changes in dislocation structure, precipitation or dissolution of second-phase particles, dynamic recrystallisation, etc. The following are changes in the local mechanical properties of the material, such as hardness, yield strength and fracture toughness. There is a close connection between the change in the gradient of the subsurface microstructure and the residual stress field and the depth of the plastic deformation layer. For example, a severely plastically deformed layer is usually in a state of high dislocation density and compressive stress; at a high temperature, there may be recrystallisation softening of the surface layer, and thus the material's resistance to crack propagation will be reduced. Therefore, to find out why there is a problem of high-temperature contact fatigue, we need to know how it is related to changes in the surface microstructure.

2. Fatigue Damage Mechanism of High-Temperature Sliding Contact

2.1 Temperature-Dependent Mechanical Properties and Surface Response Characteristics of Materials

2.1.1 Temperature-dependent Constitutive Behavior and Cyclic Softening/Hardening Mechanism

The Structure of Materials at High Temperatures is strongly dependent on temperature and rate. A viscous plastic term should be added to the constitutive model to account for dislocation motion and thermally activated diffusion creep. Cyclic sliding contact loads will cause cyclic softening or hardening of the subsurface material because of the non-proportionality of the loading path. Cyclic response characteristics are associated with dynamic strain aging, the formation of a dislocation cell structure, interaction with second-phase particles, etc., and directly affect both the speed and distribution of plastic strain accumulation; thus, this is a necessary condition for the occurrence of damage.

2.1.2 Time-Temperature Equivalence Effect and Creep-Fatigue Interaction

In a hot place, the load will be maintained for a long time and creep damage may occur. Creep-fatigue interaction is the joint occurrence of damage and stress relaxation at different times. One can use the principle of time-temperature equivalence as a reference, but due to the local and instantaneous high-temperature and high-stress gradient caused by sliding contact, the classical steady-state creep model is not applicable. A constitutive law should be developed that can describe transient creep and its interaction with cyclic plasticity to quantify the additional damage caused by creep voids or grain boundary sliding; this is one of the main differences between high-temperature contact fatigue and room-temperature fatigue[2].

2.1.3 Interaction of Environment and Surface Degradation Mechanism

High-temperature operation is generally caused by oxidation or other special environments. Environmental interaction is the formation and decomposition of the surface oxide film, as well as the diffusion and penetration of oxygen along grain boundaries or defects; thus, it can cause damage to the internal structure and mechanical properties of the material. The reason for environmentally assisted crack initiation is that the type of surface response has changed from purely mechanical to a combination of mechanical and chemical. In order to fully describe the effect of the environment on the ductility, fracture toughness and recrystallisation temperature at the surface of a material, one needs to know the adverse effects caused by the environment.

2.2 Thermo-Mechanical Coupling and Stress Field Analysis of the Sliding Contact Process

2.2.1 Model of Frictional Heat Source and Calculation of Unsteady Temperature Field

Friction at the sliding contact surface is a type of heat generation caused by motion. The exact shape of the friction coefficient is not fixed, and it will change with changes in temperature, speed, etc., at the surface. According to the theory of heat conduction and considering the temperature dependence of the thermophysical properties of the material, the unsteady temperature field in the contact zone and the subsurface layer can be determined. Based on the calculation results, both the spatial gradient and the time fluctuation of this temperature field are very large; the maximum value is close to or even exceeds the phase transition temperature of the material, so there will be a significant change in the local state of the material.

2.2.2 Coupling Analysis of Thermoelastic Stress and Thermoplastic Stress

Thermal stress is the result of preventing uneven thermal expansion in a material. If the local stress exceeds the upper limit of the high-temperature yield, there will be permanent thermoplastic deformation and then a new residual stress field will be formed. The first goal of the thermo-mechanical coupling analysis is to find the distribution of elastic-plastic stress-strain caused by the joint action of mechanical loads and thermal loads. The main method is the finite element method with incremental iteration, and it is necessary to consider the temperature non-linearity of material properties and deformation history to reveal the evolution law of the amplitude, distribution and mean stress of the stress field under cyclic loading.

2.2.3 Evolution of Stress State and Determination of Potential Damage Locations

The first purpose of the thermo-mechanical coupling analysis is to find the places with the highest equivalent stress, maximum shear stress and maximum hydrostatic tensile stress, and how much these values change due to temperature fluctuations. At high temperatures and in the region of the sliding contact, the maximum shear stress will be relatively small because of thermal softening of the material. The part of the hydrostatic tensile stress is related to creep damage and the start of cracks in the opening mode. Based on the analysis of multiaxiality and non-proportionality in the stress state, the weakest point for the start of fatigue damage and its main mode of stress have been determined[3].

2.3 Fatigue Crack Initiation and Propagation Paths According to Surface Condition

2.3.1 Competition Mechanism of Multi-mode Crack Initiation

Crack initiation is not caused by a single reason; rather, it is the result of the combined action of many factors at that time. Transgranular initiation may occur at stress concentrations caused by the roughness of the surface; intergranular initiation can happen in the region of surface residual tensile stress or in areas with microstructural defects (such as coarse grain boundaries or precipitate bands); and at the subsurface location of the maximum shear stress amplitude, crack initiation may be due to extrusions and intrusions caused by dislocation accumulation. The first mode of initiation is determined by the ratio of the local driving force to the material resistance at that place, so it is a multi-site competitive damage initiation problem.

2.3.2 Short-Range Barrier Effect of Microstructure on Short Crack Propagation

The early development of cracks after initiation (the short-crack stage) is closely related to the microstructure and is very irregular and scattered. Grain boundaries, phase boundaries and second-phase particles are all short-range defects, so at this time, the speed of crack propagation is greatly influenced by the local microstructure. Microstructure gradients are caused by changes in the surface environment, such as plastic deformation zones, recrystallisation areas, oxidation-damage zones, etc., and all of them are relatively impermeable to the growth of short cracks. To know how a short crack propagates in a gradient microstructure, we first need to know when and where it started.

2.3.3 Main Crack Path Selection and Multi-crack Interaction Model

A small crack will gradually widen and become a serious main crack, and the path of this main crack will be determined by the near-field stress intensity factor and the anisotropy of fracture resistance in the material. Due to the complex stress field produced at the high-temperature sliding contact, a crack may be initiated and propagate in the direction of the maximum energy release rate, or it may also be deflected by the effect of material texture or the residual stress field. At the same time, there are also some cracks (such as stress-field shielding or coalescence), and the remaining service life will be greatly reduced. The last theoretical problem in achieving accurate life prediction is how to build a propagation model that considers the initial conditions of surface integrity, changes in the stress state and all kinds of crack interactions.

3. Construction of a Fatigue Life Prediction Model Based on Surface Condition

3.1 Theoretical Basis of Multi-scale Damage Coupling

3.1.1 Macroscopic Continuum Framework and Phenomenological Damage Mechanics Characterization

According to the framework of macroscopic continuum mechanics, fatigue damage is considered an internal variable, and changes in this damage can be expressed by a constitutive law that also considers the effect of thermo-mechanical coupling. According to the laws of irreversible thermodynamics, a free-energy potential function can be introduced to account for elasticity, plasticity, creep and damage. The damage variable is the reduction in the effective load-bearing area of the material, and its evolution equation is driven by the accumulation of plastic strain, hydrostatic stress and temperature as driving potentials; thus, at the macroscopic level, it can be seen that the load-bearing capacity of the material continues to decrease due to damage at the surface.

3.1.2 Unified Evolution Criterion for Damage Mechanisms at the Mesoscopic Scale

A single development rule for damage at the mesoscale should be established to link the response of the macro-level with the causes of micro-level damage. First of all, the beginning and early development of cracks are the result of the combined effect of competition in the local stress-strain field, microstructural resistance and environmental factors. According to the critical plane method, a multi-axial fatigue parameter is employed to construct a physical mechanism model in line with the microstructural short-range barrier model, and thus the process of dislocation accumulation, microvoid formation and finally microcrack coalescence can be described; therefore, a mesoscopic physical interpretation of the macroscopic damage variable can be obtained[4].

3.1.3 Cross-Scale Correlation and Information Transfer Method

The first general way to achieve multi-scale coupling is to build a good mode of cross-scale information transfer. A representative volume element (RVE) is selected, the average damage evolution rate and the constitutive softening effect at the mesoscale are averaged, and then homogenization is performed to obtain the results at the macroscopic scale. The non-uniform stress-strain field at the macro level is given as the boundary condition for the mesoscale analysis. Top-down and bottom-up bidirectional coupling can be used to consider the gradient characteristics of surface integrity at all scales uniformly in the model, such as the residual stress field and the microstructure gradient.

3.2 Quantify the Key Surface Integrity Parameters and Add Them to the Model

3.2.1 Dimensionality Reduction of Characterization Parameters and Extraction of Feature Vectors

The data of the raw surface integrity (e.g., 3D topography point clouds, residual stress depth profiles and EBSD orientation maps) are high-dimensional and have a large volume, so feature extraction and dimensionality reduction are necessary. Principal Component Analysis, Wavelet Transform and other methods can be used to obtain the fractal dimension of a surface and reduce the large amount of information in this type of data to a few meaningful parameters. The above vector is the collection of all statistical features of surface topography, gradient characteristics of the stress field, and topological features of the microstructure, and it serves as the first input set for the model[5].

3.2.2 Coupling Modeling of the Parameters and the Damage Evolution Law

Add the extracted feature parameters of surface integrity to the damage evolution equation according to the organisation. For example, one can relate the amplitude and distribution parameters of surface roughness to the crack initiation driving force term; one can add the principal components of the residual stress tensor as a mean stress correction term to the fatigue life equation; and one can correlate microstructural parameters such as grain size and dislocation density with the material cyclic hardening/softening law and crack propagation resistance. A weighted coupling relationship matrix is introduced to add weight coefficients and interaction terms, and it is determined how much each element of surface integrity contributes to the different stages of damage (initiation, short crack and long crack).

3.2.3 Parameter Sensitivity Analysis and Model Order Reduction Based on Machine Learning

Given that the model has many parameters and nonlinear coupling, the method of global sensitivity analysis used here is surrogate model technology based on machine learning. The first is Latin Hypercube Sampling, and then either Gaussian Process Regression or a deep neural network can be employed to construct a very accurate mapping function from the input parameters of surface integrity to the output of fatigue life. Based on the above, it can be determined which parameter combination is most sensitive to changes in life prediction, reduce the order of the model, and find the direction for optimisation of process control to improve the engineering practicality of the model.

3.3 Framework and Model Validation of Life Prediction Algorithm

3.3.1 Numerical Solution Framework for Coupling Finite Element Method and Damage Mechanics

The general numerical solution framework of this paper is shown in the figure above, and at the core of this framework is the user material subroutine interface for commercial finite element software, such as ABAQUS or ANSYS. At each step of increasing the load, the Finite Element Method is employed to solve the problem of thermo-mechanical coupling in large structures. The user-defined material subprogram will add the multi-scale damage coupling constitutive model and update the material state (stress, strain and damage) in accordance with the current field variables. To ensure the accuracy of the calculation, an implicit integration method will be employed, and then cyclic iteration will be carried out until the damage reaches its upper limit; thus, the entire life cycle of the component will be modeled numerically.

3.3.2 Hierarchical Strategy for Model Validation and Uncertainty Quantification

There is a hierarchy of model verification. First, at the level of materials, it has been verified that the constitutive model can reproduce the results of uniaxial tension-compression, creep and fatigue tests. Second, at the level of components, this paper compares the predicted and measured positions of crack initiation, propagation paths and total life after simulating sliding contact fatigue tests on simplified specimens with a known condition of surface integrity. To determine the accuracy of the prediction, uncertainty quantification analysis is also carried out in this paper; that is, using the Monte Carlo method, the statistical scatter of surface integrity parameters, material properties, loading conditions, etc., is considered to obtain a prediction interval for the life rather than just a single value[6].

3.3.3 Adaptability Assessment and Generalisation Capability Discussion of the Prediction Model

Finally, it should be confirmed that the model is in good agreement with all the material systems, process conditions and service environments. Based on the above model, predict the fatigue life of the new data (e.g., data obtained after all kinds of surface treatment processes or at slightly different temperatures) and determine how well it can be generalised. The main reasons for the prediction deviation in this paper have been identified, and it can be concluded that they are either due to an insufficient description of a particular damage mechanism or a lack of characterization parameters for surface integrity; thus, a clear technical path for continuous iteration and improvement of the model has been established.

Conclusion

Organise systematically the study of the mechanism of the effect of surface integrity on high-temperature sliding contact fatigue behaviour and develop a life prediction model based on multi-scale damage physics in this paper. To construct a good model, we need some quantitative data on the surface shape, residual stress field and gradient microstructures, as shown in this paper. Based on the above studies, many types of damage initiation and propagation in the presence of thermo-mechanical coupling and creep-fatigue interaction are caused by the combination of surface conditions and local stress fields. Periodically add the characteristic parameters of surface integrity to the cross-scale damage evolution law in the existing prediction model, and then, with the help of a numerical solution framework and the finite element method, simulate the physical mechanism of fatigue life and quantify the prediction interval. The first problem of the old empirical method was that it did not consider surface effects. The direction of the next study will be to develop a high-fidelity cross-scale computational scheme that can capture the dynamic evolution of microstructure, expand the application of this model to more complex multiaxial non-proportional loading and severe environments, and explore an inverse optimisation design method for surface integrity process parameters based on this model to build a closed-loop design theoretical system for manufacturing processes and service performance.

References

[1] Ning Shunlu. Research on the Prediction Method of Interlayer Fatigue Life of Cord-Rubber Composite Materials. 2025. Harbin Institute of Technology, MA thesis.

[2] Han Hongli. Fatigue Life Prediction of Double-Layer Armored Dynamic Cables Considering Interlayer Friction and Wear. 2023. China University of Petroleum (Beijing), MA thesis.

[3] Ding Xianwang. Study on the Thermo-Mechanical Coupling Behavior and Key Influencing Factors of Wheel-Rail Sliding Contact. 2023. Southwest Jiaotong University, MA thesis.

[4] Sun Yaoliang. Research on a Fast Prediction Method for Wheel-Rail Rolling Contact Fatigue Based on the Shakedown Map and Potential Function. 2022. Southwest Jiaotong University, MA thesis.

[5] Li Hailong. Study on the Influence of Medium on Thermal-Mechanical Damage during Wheel-Rail Sliding Contact. 2021. Lanzhou Jiaotong University, MA thesis.

[6] Chen Jinjing. Fatigue Life Analysis and Experimental Study of the Transmission Mechanism of a Column-Type Electric Power Steering System. 2020. Chongqing University of Technology, MA thesis.