REVIEW

A Review of the Rolling Contact Fatigue of Rail Wheels Under Various Stresses

M. Nur Tawfik (iD) **· M. Md Padzi · S. Abdullah · D. Harmanto ·**
M. N. Firdaws · H. Hapaz

Abstract This research paper reviews the recent works on the rolling contact fatigue of rail wheels. The topic covered includes the mathematics equations of rail-wheel contact stress and fatigue, RCF defects in the rail, the development of cracks, and the strategies used to reduce the rate of RCF defects. Therefore, the main understandings of RCF as they currently stand are crucial for the improvement and mitigation on rail fatigue failures in the future. This research paper has not been published yet and has not been discussed yet by the Malaysia Railway Network.

Keywords Rolling contact fatigue · Rail–wheel contact · Rail defects · Crack propagation · Stresses
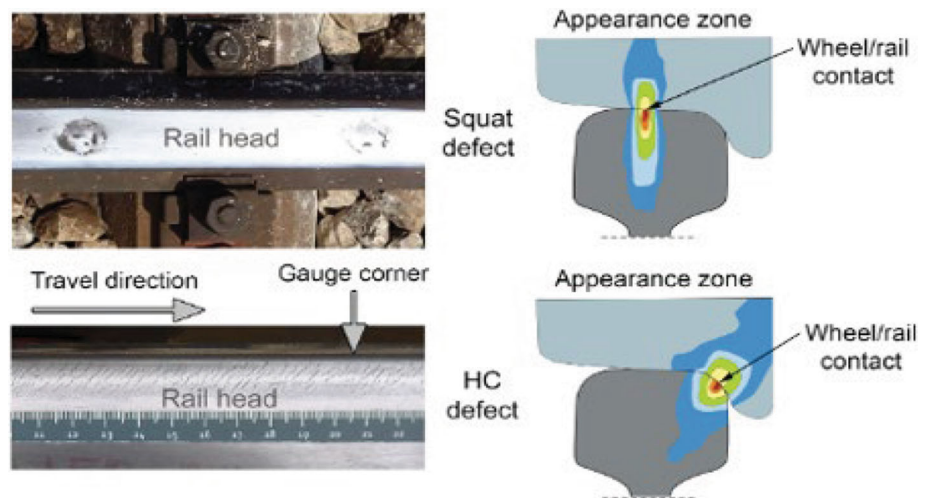
M. N. Tawfik (✉) · M. M. Padzi · H. Hapaz
Department of Mechanical Engineering, Malaysia France Institute, University Kuala Lumpur, 43650 Kuala Lumpur, Malaysia
e-mail: tawfik.badruddin@s.unikl.edu.my

S. Abdullah
Department of Mechanical and Manufacturing Engineering, Universiti Kebangsaan Malaysia, 43600 Bangi, Malaysia

D. Harmanto
Faculty of Computing Engineering and Media, School of Mechanical and Aeronautical Engineering, De Montfort University, Leicester LE1 9BH, UK

M. N. Firdaws
Department of Track Network Maintenance, Rapid Rail - Amp. Line, 5th Mile, Jalan Ampang, PO Box 39, 68000 Kuala Lumpur, Malaysia

Introduction

In today's engineering field, fatigue is one of the leading major failures caused by various loading amplitudes. The term fatigue often refers to the load applied to the subject (component). An insufficient level may cause failure in a single application [1]. Cui [2] also added that a component receives a force in continuous loads, thousands or millions of times in variable or constant ways, allowing it to develop into cracks and increasing the likelihood of the failure of the components. The failure always appears gradually, locally and permanently by relying on the repeated stress coming out of dynamic stress in the critical area where the number of stresses subjected are less than the ultimate tensile strength (UTS). Recent publications by Ye and Cini [3, 4] state that fatigue failure can be traced back to the nineteenth century during the Industrial Revolution in Europe, as featured by German engineer, W. A. J Albert in 1829. In Albert's paper, he mentioned that the repeated stress of lower amplitude loads and the loss of durability of the metal [3]. The word 'fatigue' was used a decade later by Poncelet in his book titled 'Introduction à la Mécanique Industrielle', also known as the 'Introduction to Industrial Mechanics' [3]. Later on, a disaster occurred in 1842 when a heavy locomotive axle failed due cyclic loads. The broken axle demonstrated the presence of brittle cracking across the diameter, as found by British Railway Engineer, W. J Rankine. He recognized that fatigue has an effect on the concentration of stress in mechanical components, as shown in Fig. 1 that the highest concentration of stress occurs at rail-wheel contact as able to form the defect as RCF as squat and head checking.

In 1850, the research engineer Wohler designed a method for determining the characteristics of fatigue based

Fig. 1 Zone appearance and stress concentration of RCF [8]

on a myriad of experiments focusing on the failure of railway axles. To determine the fatigue life of the components, some of the specimens with cylindrical gauge lengths were tested using constant conditions. The results from the cyclic test were plotted on the stress–life curve. Ultimately, some amendments to the mechanical design and properties are required to reduce the stress concentration areas and to control the fatigue failure of the component.

In terms of rail–wheel contact, [5] and [6] underlined that Rolling Contact Fatigue (RCF) is a term that has been used broadly in railway engineering. It is well known to be a common cause of failure because of the repeated stresses due to rail–wheel interactions when moving. The types of defect due to RCF such as head checking, squat, and spalling/shelling may be classed as dangerous as they can cause the train to derail from the track and contact patch [7, 8 and 9]. For instance, Fig. 1 shows the zone appearance and stress concentration of RCF.

According to [10, 11], rails are one of the key components of railways, involving a heavy haul track and high circle of fatigue with a cover in between 10^6 and 10^8 cycles of loading across its rail life. Following to the recent publication by Murray [11], a number of circles of 10^7 from steel can develop the endurance limit. Failure never occurs if the total number of stresses does not exceed the limit. To control the rail fatigue, the selection of rails is highly important in relation to the track limit. It is often involved in the percentage calculation of uniaxial tension strength, also known as UTS [11, 12 and 13] (Figs. 2 and 3).

Where the stress number is below the endurance limit, plastic flow may occur on the rail, but it raises the strain hardening in response to the yield stress. This is where the steel carries additional load cycles and the plastic flow ceases without further damage. This process is known as
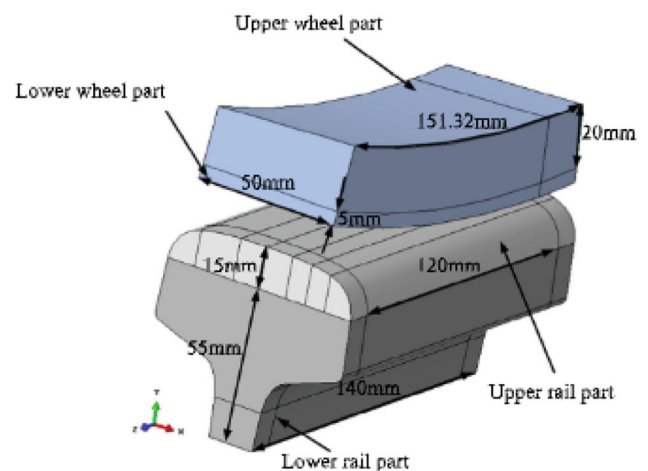

Fig. 2 FE model of rail–wheel elliptical contact [16]

shakedown, and it is one of the rails' characteristics [11]. For instance, cracks appeared at present in the surface of rail, but the stresses are below the endurance limit. Here, the plastic flow might be the action needed for the propagation of crack ceases, appearing at the tip of the crack.

Generally, the rail fatigue experiences were the outcome of the highest bending stresses, lateral loading, and rail–wheel contact stresses observed, etc., such that if the stress values exceed the endurance limit of the rails, then shakedown may not occur and the fatigue damage might accumulate with every cycle of stress [11]. In the initial phase of crack initiation, there is propagation at the surface and finally, the rail breaks which can cause a derailment accident. To respond to the rail fatigue problem, engineers need to follow the theory of contact stress by Hertz to determine how to multiply the wheel–rail dynamic load and static load.
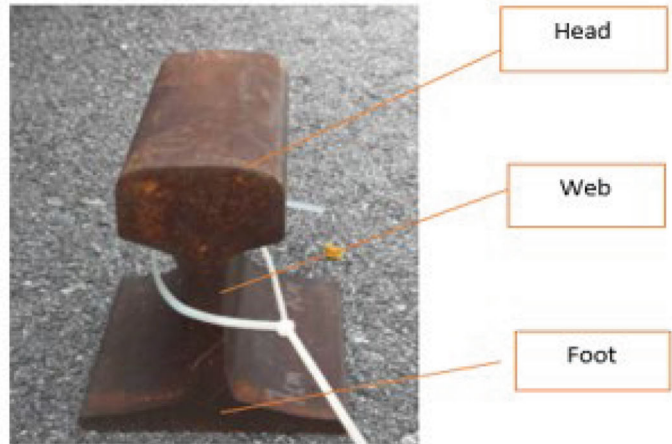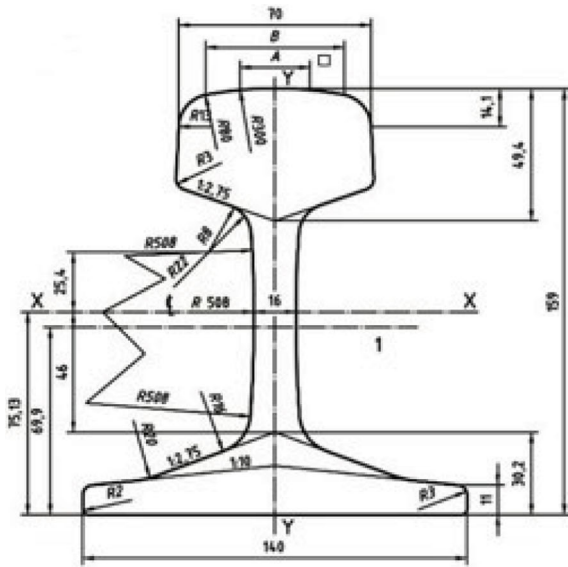
Fig. 3 Left: Actual dimensions in millimeters; right: damaged specimen of R260 for LRT [29 and author] [27] mentioned that the Hertz contact theory considers the radius of the curvature of the rail and wheel profiles when they are in contact as assumed. This is where there is a constant and no plastic deformation in the contact area.

Mathematic Equations for the Rail–Wheel Contact

To predict the fatigue of the rails, it is vital to engage in the accurate stress analysis of the rail wheel. Generally, the main backbone of the Hertz contact theory is the rolling contact stress predictions. Sometimes this method incurs less accuracy in terms of the results of the rolling contact stress when the plastic deformation develops in response to a large amount of stress.

To find the value of the normal pressure distribution, $p(x,y)$ has been expressed below by Srivastava and Jagadeep and Sayed [14, 15 and 16]:

$$P = p_o \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} \qquad \text{(Eq 1)}$$

The formula represents the following:

a = half-width contact area for the x longitudinal direction.

b = half-width contact area for the y lateral direction

$$a = m\left(\frac{3\pi}{4} * p\frac{K_1 + K_2}{(A + B)}\right)^{1/3} \qquad \text{(Eq 2)}$$

$$b = n\left(\frac{3\pi}{4} * p\frac{K_1 + K_2}{(A + B)}\right)^{1/3} \qquad \text{(Eq 3)}$$

To calculate the contact areas of the rail wheel, it is essential to obtain the real measurement of the geometric constants used in the above mathematical formula. The following are the combinations of curvature proposed [14–16]:

Figure 2 shows the elliptical rail-wheel contact with labelled dimensions on the x, y, and z axis. Figure 3 on the left presents the actual dimensions of the rail at an LRT of R260

$$A + B = 0.5 * \left(\frac{1}{R_{11}} + \frac{1}{R_{12}} + \frac{1}{R_{22}} + \frac{1}{R_{21}}\right) \qquad \text{(Eq 4)}$$

$$B - A = 0.5 *$$
$$\left[\left(\frac{1}{R_{11}} - \frac{1}{R_{12}}\right)^2 + \left(\frac{1}{R_{22}} - \frac{1}{R_{21}}\right)^2 + 2\left(\frac{1}{R_{11}} - \frac{1}{R_{12}}\right)\left(\frac{1}{R_{22}} - \frac{1}{R_{21}}\right)^2 * \cos 2\phi\right]^{1/2}$$
$$\text{(Eq 5)}$$

The values of A and B as constants are where the R_{11} is the rolling radius of the wheel curvature, R_{12} is the radius of the wheel profile, R_{21} is the radius of the runaway of infinity, and $R_{22:}$ is the radius of the rail curvature in the cross section [14].

To calculate the principal stresses at the surface of the rail–wheel contact, the following equations were used:
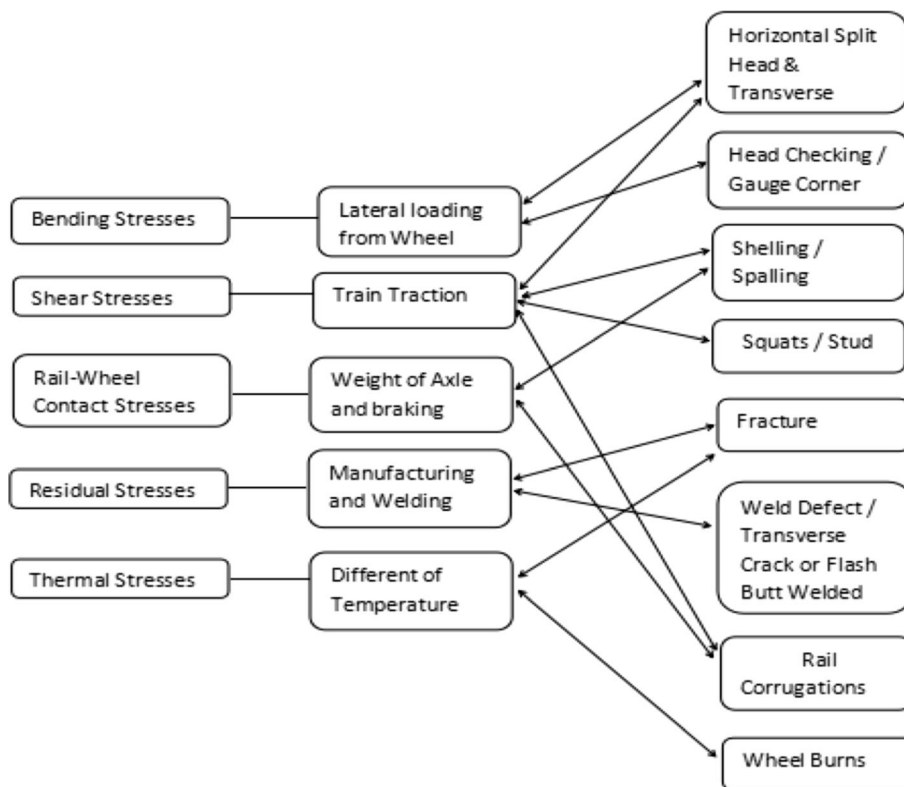
$$\sigma_1 = -2\mu p_0 - (1 - 2\mu)p_0\frac{1}{R_{11}} + \frac{1}{R_{12}} + \frac{1}{R_{22}} + \frac{1}{R_{21}}\right)$$
$$\text{(Eq 6)}$$

$$\sigma_2 = -2\mu p_0 - (1 - 2\mu)p_0\frac{1}{R_{11}} + \frac{1}{R_{12}} + \frac{1}{R_{22}} + \frac{1}{R_{21}}\right)$$
$$\text{(Eq 7)}$$

$$\sigma_3 = -p_0 \qquad \text{(Eq 8)}$$

$$\sigma_{\text{von-mises}} = \sqrt{(\sigma_1 - \sigma_2)}^2 + \sqrt{(\sigma_2 - \sigma_3)}^2 + \sqrt{(\sigma_3 - \sigma_1)}^2$$
$$* \left[\frac{1}{2}\right]$$
$$\text{(Eq 9)}$$

Fig. 4 Types of stress and their root cause

Fatigue Analysis

This section demonstrates the fatigue analysis results based on the finite element analysis. The fatigue crack life usually can be divided into 3 phases covering the initiation of the crack and its growth. In the first phase, shear stress occurs at the surface. In the second phase, there is the growth of the crack and in the final phase, there is the subsequent shear crack growth and tensile effect. The crack starts with repeated rolling contact on the surface which may occur in phase 1. To determine the life of fatigue, a formula has been introduced by Jiang [17] as follows:

$$FP = <\sigma^{\max}> (\Delta\varepsilon * 0.5) + J\Delta\tau\Delta\gamma) \quad\quad\text{(Eq 10)}$$

The mathematical formula for the fatigue life on the plane can be used as follows: Eq. 11 sets the material properties shown in Fig. 4.

$$FP_{\max} = <\sigma^{\max}> (\Delta\varepsilon * 0.5) + J\Delta\tau\Delta\gamma)_{\max} \quad\quad\text{(Eq 11)}$$

Types of Stress and the Root Cause of Rail–Wheel Stresses

Bending Stresses

Bending stresses are one of the types of stress during train movement from point A to point B. The train's movement through the wheel creates a force known as lateral loading that is also vertical to the rail [7]. There is the possibility of rail failure due of the forces involved which contributes to the fatigue failure. Usually, the vertical loading appears where there is a static axle load of 7 tonnes—23 tonnes. The dynamic effect comes from the train movement. Figure 5 shows the schematic diagram for the movement direction and crack propagation. The weight of the rails, the presence of defects, and irregularities in the running interface mean that motion sleepers can be influenced by the magnitude of the bending stresses (Fig. 6).

In the current circumstances, the bending stress can be measured using a computer based on the theory of a beam on an elastic foundation. Tensile stress shows on the rail surface a distance from the wheel's position rather than in contact with the wheel, according to [7]. Figure 7 displays that how the stress distribution on the bending, shear, contact stresses and bilk stresses in longitudinal section of rail.

Zimmermann [18] revealed that determining the bending movement of M_Y and M_Z can be undertaken as follows:

$$M_y(x) = -\frac{F_v}{4\lambda_v}(\cos\lambda_v x - \sin\lambda_v x)\cdot\exp(-\lambda_v x)$$

$$M_z(x) = -\frac{F_v}{4\lambda_v}(\cos\lambda_L x - \sin\lambda_L x)\cdot\exp(-\lambda_L x)$$

The representatives of symbol x—the longitudinal position on the rail, F_L and F_V the lateral and vertical

20

J Fail. Anal. and Preven. (2023) 23:16–29
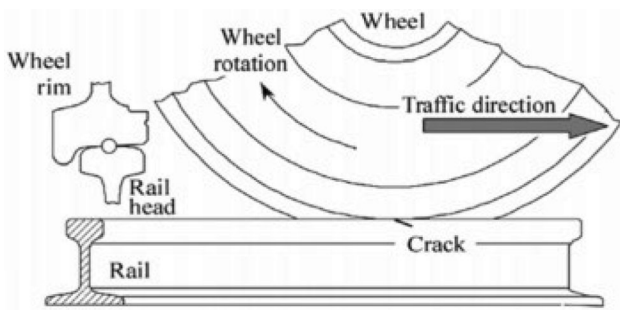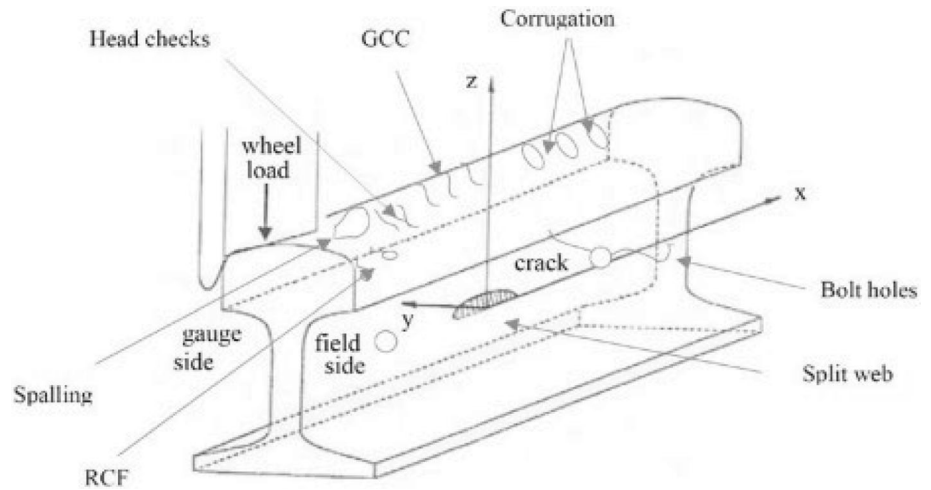
Fig. 5 Schematic of rail defects [28]

Fig. 6 Schematic diagram of a wheel rolling over a rail showing the movement direction and crack propagation by [7]

Fig. 7 The stress distributions of bending, shear, contact stresses, and bulk stresses in the longitudinal section of the rail [7, 29]

wheel loads are shown in Fig. 8. As Fig. 8 demonstrates, the bending stress at rail surface and location of F_L and F_V λL and λV are constant parameters which rely on the lateral and vertical foundation stiffness k_V and k_L and rail geometry.

$$\lambda_v = \sqrt[4]{\frac{k_v}{4E \cdot I_{yy}}}$$

$$\lambda_L = \sqrt[4]{\frac{0.85k_v}{4E \cdot I_{zz}}}$$

Shear Stresses

The contact area of the interface between the rail and wheel involves the traction of the trains which generates the shear stress [7]. The types of situation caused by shear stress can be due to the particulars below:

1. Acceleration and deceleration of trains [7]
2. Train start and stops [7]
3. Slip occurs [7]
4. The lateral movement of the wheel results in a transverse creep during the train curving or hunting [7]

The creep traction or forces can result in fatigue failure as it generates a high level of tangential stress between the train and rail interface layer which is a mixture of wear debris, dust, rust, and environmental contaminants [7]. To minimize the creep traction that controls the traction from the train, a lubricant is used at the interface, in comparison with a dry rail surface [7, 19]. Ultimately, shear stress can result in a crack at the surface. It plays a salient role in vertical crack propagation with the involvement of major fracture mode II as well as knowing how the wheel rolls over the rail surface.
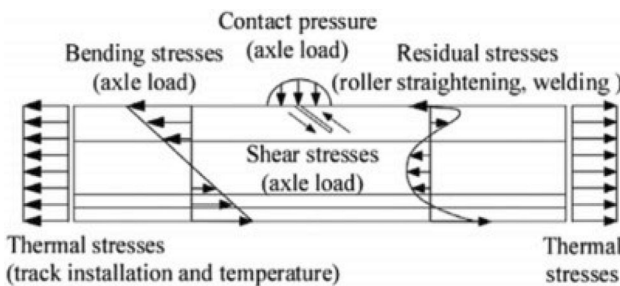
Rail–wheel Contact Stresses (RWC)

One of the major problems that contribute to fatigue failure come from rail–wheel contact stress. It forces the wheel rail at the contact patch to induce high contact stresses in the initial stage of crack propagation. The modern researcher has been sketch out the position of contact stress zone and typically loads in three axis by Singh [20] as Fig. 9. A crack from RWC stress can be traced at the rail surface, and it has the potentiality to propagate in two stages as follows:

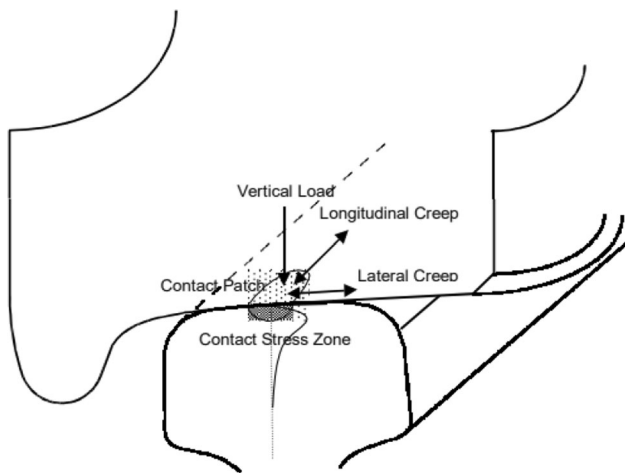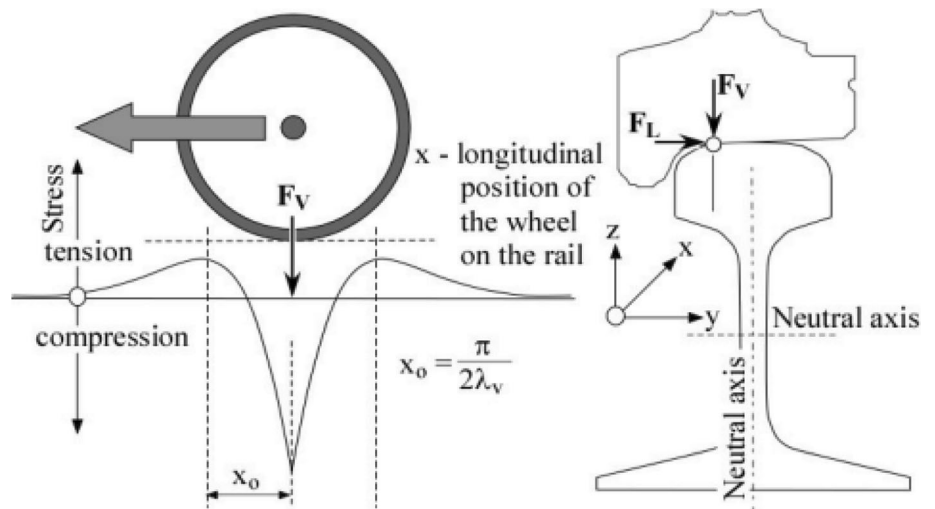**Fig. 8** Bending stress at the top of the rail surface due to the wheel and position of F_L and F_V by [29, 30]

Fig. 9 The position of the contact stress zone, vertical load, longitudinal load, and lateral load [20]

Fig. 10 Schematic diagram governed by contact stresses and bulk stresses [7]

Stage I Cracks can grow at a shallow angle of 9–39.5° into the rail. This is where the crack can change the propagation direction after reaching a critical depth. The stress field dominates crack propagation [7, 21, 22].

Stage II Upward into the rail surface, causing surface spalling. When downward, it is at a steep angle of 60–80°, potentially causing the rail to break [7, 21, 22].

The RWC stress relies on the contact area profile, train braking/steering, the presence of welds, and dynamic axle loads which all propagate cracks at the rail, including when the direction and speed rapidly decreases. To reduce the crack propagation, it is essential to use grease or lubricant to control it. According to recent publications, the Hertzian model is the right theory of calculation to use for rail–wheel contact stresses as it highlights elastic contact [23, 24]. The rail–wheel contact stresses are decided by the transverse wheel profile, transverse rail profile, wheel load, and wheel diameter. Furthermore, a double wheel load can
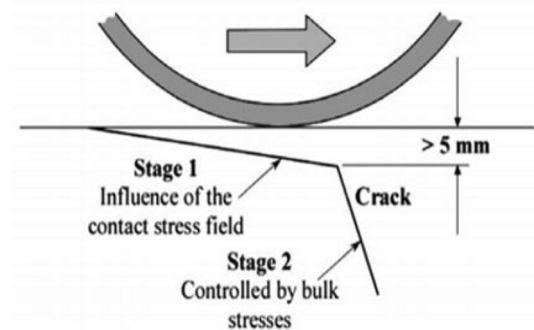
increase the contact stress from 27–44% depending on weight, but it can be reduced the acceleration level of crack in propagating with supporting from bulk stresses due to contact stress. To get more understadings about these, Shen [7] has drawn the schematic diagram as Fig. 10.

Residual stresses

Residual stresses can be traced at any rail location due to the welding or manufacturing process during either the installation or in rail production. During the manufacturing process, there are residual stresses generated by the heating and roller straightening treatments.

Thermal Stresses

Thermal stresses can be generated by the rails due to the difference between the rail's neutral temperature and the ambient temperature. The rail's neutral temperature means that there is no thermal stress on the rail during installation or during the running of the train. The ambient temperature is higher than the neutral temperature. For example, at
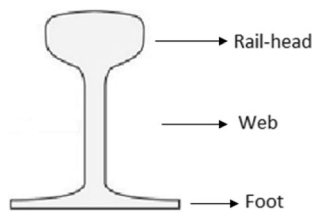
22

J Fail. Anal. and Preven. (2023) 23:16–29

Fig. 11 Rail profile

noon and on hot days, the rail is compressed and prevented from expanding in the axial direction [7]. Therefore, the risk of derailment may be increased because of the buckling of the rail. Recent reports also mention that when the tensile thermal stress reaches the maximum at night or in cold weather, the possibility of the rails fracturing is high.

Myriad Rail Defects

A lot of defects can be detected on the rails either at the head, web, or foot depending on the particular circumstances, example: Fig. 11 about the rail profile. Rail defects can be classified according to their position:

- Railhead—Cracks will occur at the surface point of origin caused by shelling/spalling, squats, and head checking due to rolling contact fatigue [7].
- Web—Horizontal cracks and vertical longitudinal cracks (piping) on the web. These are normally caused by welding where cracks occur at the fish-bolt holes used for the joining of rail sections and in the manufacturing process. All of the cracks are able to fracture [7].
- Foot—Cracks can be transverse or longitudinal as shown in Fig. 12. However, the location when inspecting the foot is the most difficult because it regularly causes rail fractures [7].

The improvement in the quality of the rails is important to reduce the defects associated with RCF resulting from high train speeds, high axle loads, and inappropriate manufacturing and installation.

Horizontal Split Head

According to [25], a defect in the horizontal split head can be traced from horizontal separations in the railhead where it shows the division of the two parts of the rail head itself. Generally, this defect can be considerable in length, specifically between more than 50–100 mm. Figure 13 presents that a real defect on horizontal split head at top of rail surface from rail company. The dark point of small and
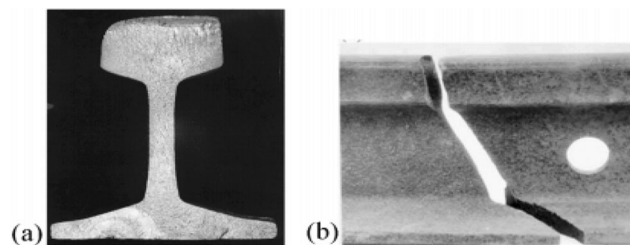

Fig. 12 (**a**) Transverse rail foot crack and (**b**) longitudinal crack in the foot by [7]
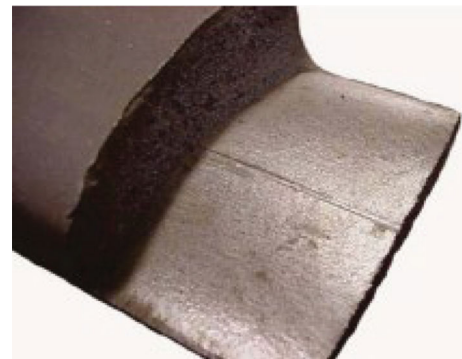

Fig. 13 Horizontal split head defect from NSR [26]

medium defects is that they are unable to be traced visually. However, larger defects can be looked at as shown in the figure below:

Some of the previous studies argue, regarding this defect, that the inclusion stringers or elongated seams of the horizontal split head which present at the rail and initial crack have the possibility to grow by an uncertain amount in actual stress conditions [25]. Nevertheless, it is pertinent to consider whether the crack development entails a shear stress component with strong evidence involved as follows:

- Lateral creep force occurs at the railhead that is produced as the wheel oscillates from one side of the track to the other [25]
- The field side localizes the loading of the rails which occurs where there are worn rails and hollow wheels [25]

The horizontal split head is able to be treated through newer rails which reduces the risk development normally occurring in older rail steels [25].

Gauge Corner Cracking (Head Checking)

This is one of the names for surface breaking that can be detected at the gauge corner of the rail head. The dominant location of this defect is mostly at tangent rails and the outer rails of curves. The spacing of the cracks is able to
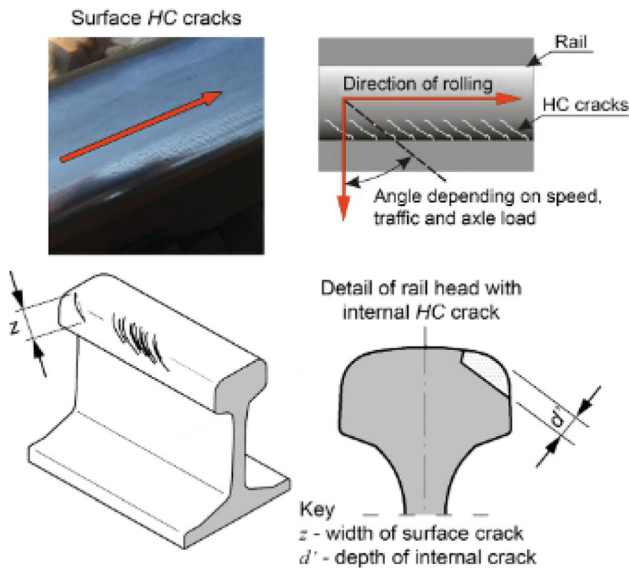
Fig. 14 Head checking of the gauge side and real dimensions [8]

threaten the integrity of the rail as the loads can steeply turn into the rail, causing transverse cracks [7]. In addition, there is the potential to fracture the rail and cause the crack to propagate upward [7]. The academicians from railway have drawn the defect of head checking and real dimensions in Fig. 14 [8].

Shelling/Spalling

The defects here start with an internal defect at a depth of 2.8 mm below the gauge corner. This mostly occurs on the high rail of the track as of Fig. 15. However, some researchers have opined that the shelling defect becomes noticeable when there are dark spots on the gauge corner of the rails. Regularly, the shelling defect does not appear and form at the gauge corner. It can develop into a crack on either the longitudinal or horizontal plane consistent with the shape of the rail.

Squat/Stud Defect

Squat defects occur at the surface or subsurface of the rails where it is possible to propagate the crack longitudinally and laterally when caused by RCF at the rail-head as shown in Fig. 16. Several characteristics of this type of defect include a shallow depression on curved rails and a tangent [7]. The root cause of squat defects comes from the train speed at either low, medium or high traction. The cracks from the squat can be divided into two major cracks as follows:

- Longer trailing crack propagating in the opposite direction.
- One short leading crack in the direction of the train.

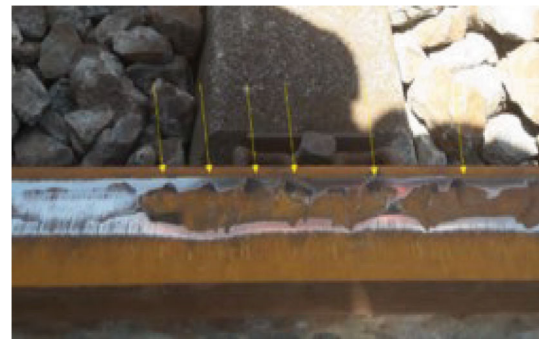Fig. 15 Shelling/Spalling defect at rail surface [31]

Fig. 16 Squad defect of the running surface by [32]

Moreover, the cracks are able to propagate 3–5 mm below the rail surface and occur 10–30° to the horizontal. To detect small cracks from the squat, it is essential to use an ultrasonic signal to ensure the location of the defect. As stated by 66–68, there are two different types of squat mechanism as stated below:

- The squat defect, also known as stud, which can develop without any major plastic deformation of the rail surface.
- The classic defect of the squat is able to form due to plastic deformation at the surface due to the wheel-rail traction. The liquid element can be influenced by crack propagation where there is a shear mode and tensile crack growth.

Rail Corrugations

This type of defect is cyclic (wake like) and involves irregularities and vertical running on the surface of the rails [25]. There are two types of rail corrugation:

(1) Long pitch—the number of measurements above 300 mm in wavelength [25]. Ker [25]

24

J Fail. Anal. and Preven. (2023) 23:16–29

Fig. 17 Rail corrugations by [33]

Fig. 18 Wheel burns defect [34]

states that the long pitch can be developed by more than > 20 tonnes of train operation. The depth of length can range from 0.1 mm to over 2.0 mm. The original defect for these are in Fig. 17.

(2) Short pitch—the number of measurements ranging between 30 and 90 mm in wavelength [25]. Ker [25] highlights that the short pitch is able to grow below < 20 tonnes of nominal load and the length of depth of the rail corrugations is normally less than 0.2–0.3 mm. Further phenomena regarding rail corrugation are as follows [25]:

- Rail support is very weak—timber sleepers, concrete sleepers, and continuous support.
- Different grades of track.
- Increase in the dynamic loads of the train.
- Higher vehicle speeds.

Wheel/Engine Burns

Wheel burns or engine burns are a known type of defect formed by the running of the rail surface. It is caused by the continuous slipping of the train or locomotive wheel movement. During the wheel slipping that occurs, there is an increase in temperature on the surface of the rail to a high value. Subsequently, the rail can be deformed by more than 50 mm in length, in addition to small defects (equivalent to a squad defect) at the rail surface as Fig. 18 occurs in India and Fig. 19 occurs in Malaysia.

A recent study has determined the key factors that increase the likelihood of the wheel movement to slip as stated below:

- The train driver not following procedure to accelerate the locomotive vehicle [25]
- The type of wheel is stronger than the track grades [25]
- Mistake made in the selection of lubricant to assist in the avoidance of contamination when running the surface of the rail [25]
- Insufficient locomotive power in the gradient [25].

Furthermore, this type of damage has the potential to form a transverse defect as it develops from the defects associated with wheel burn when the surface condition reaches a critical displacement stage.

Development of Crack Initiation

Singh [20] mentioned that the level of cracks relies on the volume of stress from the different types of loading, friction conditions, and contact geometry between the rail and wheel. The high friction conditions of shear stresses have quite a big effect but are very shallow. Otherwise, the low friction conditions of the peak shear stress are a little less but extend deeper into the railhead. Various publications regarding RCF demonstrate that the defects come from RCF initiated at the surface and subsurface of the railhead.

Initial Defects on the Surface

In today's rail operations as part of an urban network, specifically a Light Train Transit, the railway track experiences more than 4000 cycles of rail–wheel contact due to the present day axle loads in Kuala Lumpur Railway. However, these contacts do not massively damage the rail

Fig. 19 Small defect consisting of wheel burns at Urban LRT

and several plastically deform the steel in the direction of the applied traction. Eventually, plastic deformation occurs at the surface layer until the rail exhausts its available ductility. Azmalea [26] also provided some viewpoints on this in that the defects from RCF are not solely based on single normal contact stress. Instead, they proposed that there might be an interdependence between material strength and traction forces. In the context of rolling contact, fatigue can be divided into three parameters as shown in the next section.

Normal Stress at Rail–Wheel Contact

The normal stress experienced by rail wheels can be categorized according to four factors as the wheel load and wheel diameter include a dynamic augment. There is also the transverse profile of the wheel and transverse rail profile to take into account. An experiment focused on wheel load found that doubling the load of the wheel may increase the contact stress by 27%. Triple the load can increase the contact stress by 44% [20]. Nevertheless, the effect of the diameter of the wheel is limited to 1/3 in terms of power function [20]. Other factors might influence the actual rail–wheel position highlighted as follows:

1. Track gauge—This involves changing the geometry of the rail–wheel contact and the distance between two rails. A straight track on a tight gauge promotes RCF and the corner contact where there is a nominal gauge with more contact is able to generate a crown. For the curves, reducing the RCF defects involves controlling them with a wide gauge to mitigate and minimize the rail damage.
2. Welds and Profile Irregularities—The type of welded steel can have different hardnesses according to the parent of the rail steel. The harder welds generate high spots that play a role in controlling RCF damage and increasing the dynamic augment. Softer welds can produce cupping or a dip which speeds up the development of RCF.
3. Other Aspects—The other aspects that have an effect on contact stress include track geometry errors, mismatched rail–wheel diameters, cant excess, uneven wheel loads, the hunting of wheel sets in mild curves and tangent track, and cant deficiency.

Tractions of the Rail Wheel

According to [20], rail–wheel traction is the development of a minor relative slip between the rail–wheel contact where shear occurs at the layer in the contact area. Another term for slip is known as creep, and it relies on the traction requirements and curving movement. The forces that come from creep must not exceed the available adhesion as the vertical forces at the wheel are a coefficient of friction. To control creep, controlling traction is the right process, as well as controlling the properties of the interfacial layer. Shen [7] states that the application of a lubricant is one of the alternatives used to reduce the peak traction force to compare it with a dry surface so then it is eventually in a curve rather than a straight line. This effectively controls the wear at the rail gauge face at the top of the rail as a friction coefficient of the 0.33 required to facilitate traction and braking [20]. A lot of recent researchers agree that traction produces a longitudinal F-long (force component) and lateral F-flat as $T = sqrt (Flat 2 + Flong2)$. This is where the approaches limit the rail–wheel friction coefficient. Counting the real traction ratio T/N for any given rail–wheel combination depends on some of the parameters underlined by Singh [20] specifically the suspension characteristics of bogie, the wheel base, curvature, friction coefficient, and cant deficiency.

Rail Metallurgy

Rail metallurgy is a subject focused on the level of steel hardness and pearlitic steels so as to create more resistance to RCF propagation than other structural materials such as martensite and bainite. To control RCF using hardness and prior experience also reveals that the high strength of steel can result in RCF due to the failure of harder steels to wear down to a lower stress shape. As a result, the hard rail is the most effective compared to softer steel when it comes to reducing RCF. [7, 20] argue that to reduce the rate of RCF defects, to control the maintenance cost, and to improve the safety of train movements on the rails, there is a need to

26

J Fail. Anal. and Preven. (2023) 23:16–29

focus on a work-hardened layer that is 8–10 mm thick with a high shear yield strength and high tolerance to compressive residual stress. The advantage of being resistant to high compressive residual stress is that it generally prevents the growth of shallow cracks and therefore responds to fatigue and crack growth.

Surface Crack Propagation

In terms of fatigue failure at the point of rail–wheel contact, the start of the crack may be initiated on the rail surface where the rate of progression becomes bigger and more serious. The current researcher has argued that crack propagation can lead to failure, but it depends on the stress propagation and the ability of the steel to resist propagation. To tackle this problem, the presence of water and grease at the rail–wheel contact point plays a vital role in the depth of the surface crack propagation and its rate of progression. The function of grease in the surface of the cracks is that it is able to reduce the crack face friction, contributing to moderate crack growth rates and the ability of the crack faces to slide past each. For the water, it has a high surface tension and low viscosity, meaning that it is drawn into the cracks by a capillary action. Singh [20] explained that if the surface crack is oriented in a direction where it drops away from the approaching load, the rolling contact will first seal the crack entrance and then hydraulically pressurize the crack tip.

Subsurface of RCF Defects

In rail–wheel contact, not only can a defect happen on the surface of the rail but also on the subsurface of the rails. This is considered to be rolling contact fatigue. This comes from horizontal split heads, as well as the vertical and tache ovales categorized as RCF defects. Shen [7] and [20] stated that deep-seated shells occur in direct straightforward through to excessive contact loads at the gauge corner, creating a failure in the rail along the shear-slip line. They emphasized the critical point on the surface of the rail according to RCF.

Subsurface Crack Propagation

It is necessary to understand how the distribution of cracks propagate according to the type of stress. The propagation of cracks can be determined through the rolling stock and environmental conditions as either the main phenomenon or as one of the influences on the parameters [20]. Ultimately, the characteristics of the cracks allow for their propagation due to the gravity of the situation and the propagation speed. The figure below shows the crack propagation process divided into 4 phases.

To gain a strong understanding of the phases of crack propagation as shown in Fig. 20, some of the summaries that are easy to understand have been included as follows [20]:

Region	Phenomena
A	Crack initiation by low-cycle fatigue; deformation and the propagation process starts
B	As crack lengthens, stresses near the crack tip increases and the crack growth rate increases
C	Beyond a certain critical crack length, the tip of the longest crack, moves away from the highly localized contact stress field and the stress intensity drops, leading to a reduction of crack growth rate
D	Crack is subjected to flexion and the residual traction stress facilitates the crack growth rate

| Wear rate level da/dN	Characteristics
Level 1	The very high wear rates do not permit crack formation; crack mouth truncation is higher than the rate of advance of the crack
Level 2	The slightly lower wear rate rubs out the initiating crack faster than them from; hence cracks are hindered from growing
Level 3	Very low wear rates; thus that is not affecting crack growth at all

Strategy to Reduce the Defects in the RCF of Rail Wheels

RCF defects are a vital focus in the railway industry due to the threat to the safety of passengers, the potential for a loss of assets, and the effect on the maintenance program. RCFs also can provoke rail fractures when in an extremely critical situation from the safety perspective. Several methods and procedures can be used to reduce the risk of RCF development and subsequent rail failure which has been summarized as follows.

Rail Grinding

As pointed out by [20], rail grinding is an important process that is used to control RCF defects, particularly surface cracking and the gauging of corners. It also plays a salient role in establishing a balance between the

Fig. 20 Phases of crack propagation [20]
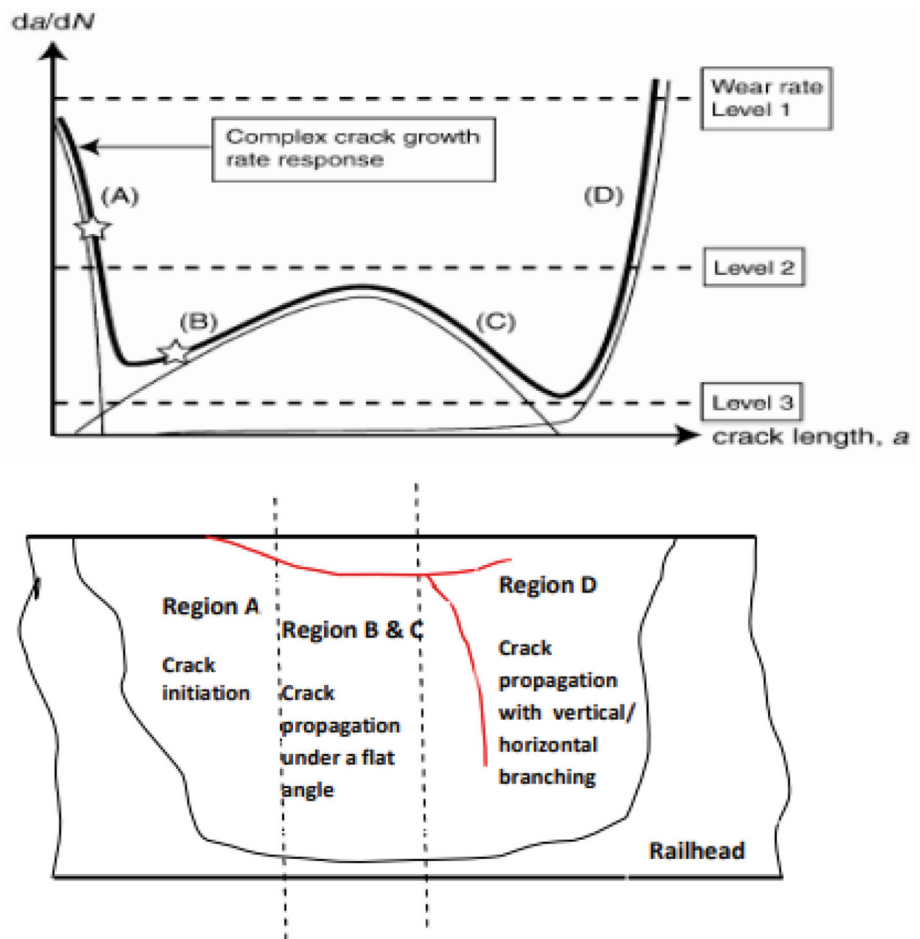
development of RCF and rail wear. This method has been used to prevent rail and wheel material waste but not a sufficient amount of wear as it allows fatigue cracks to initiate and propagate, increasing the risk of rail failures and reducing the life of the component [20]. Therefore, the completion of the rail grinding process can use the concept of magic wear rate of Fig. 21 which is where the material surface wear is removed by grinding. This removes the small fatigue cracks from propagating [20], for example Fig. 22 shows the schematic diagram of rail grinding process and the current approaches might have to uses a rail grinding train to accelerate the process to control the defects of RCF as Fig. 23.

Higher Strength Rail Steel Material

Regarding the collected results of the experiments including the modeling, visualization, and simulation that use the types of rail material with a higher strength in critical track locations such as gradients and curves, this may help to increase the allowable shear stress limit. In addition, the

hardened rails have found optional ways to reduce the development of RCF defects that are likely to include, transverse, spalling/shelling or higher axle loading circumstances. Consequently, when hardened rails have been used, it demonstrates that they are able to minimize the level of wear and reduce the degree of plastic deformation in comparison with a standard rail. The preventative maintenance regular method can therefore be applied in this manner.

Re-Profiling of Rails

This involves a correction of the railhead geometry using planning, milling, and grinding to remove defects at the rail surface and subsurface. It is one of the components of the preventative exercise part of the maintenance program used to control and respond to rail defects. As the current researchers explain, according to [8], profiling is a modern strategy with a high reliance on knowledge, experience, equipment, and technology to engage in the successful management of RCF defects. This approach makes it

28

J Fail. Anal. and Preven. (2023) 23:16–29

Fig. 21 Magic rate of wear [20]
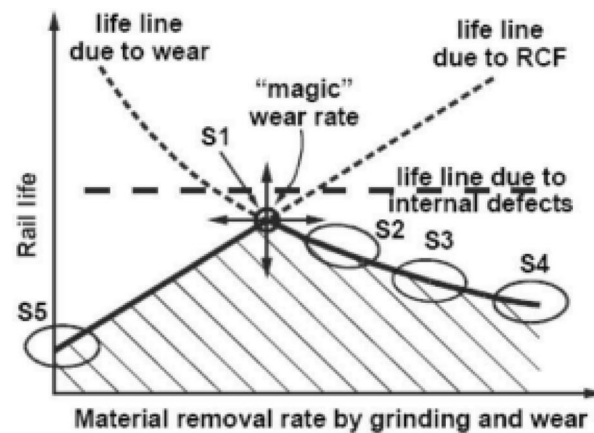
Strategy S1 : initial preventive; Strategy S2 : systematic preventive ;
Strategy S3 : maintenance; Strategy S4 : corrective; strategy S5 : no grinding
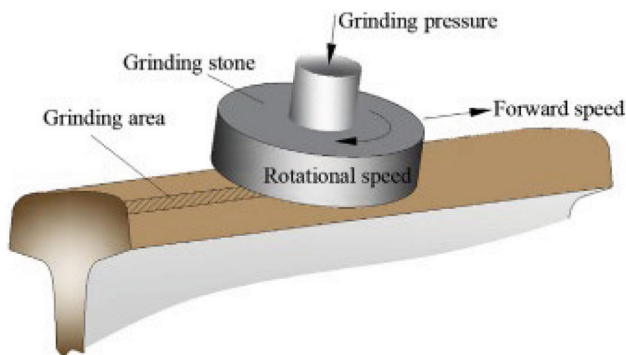

Fig. 22 Schematic diagram of rail grinding process [35]

Fig. 23 Rail grinding train on the rail surface [36]

possible to make corrections part of the long-term process and to maintain the track section in the RCF regime (squad, wheel-burn, braking, etc.).

Conclusion

In this research study, recent reviews of rolling contact fatigue have been explored according to the appearance of various defects found on rails by referring to recent publications in journals, books, articles, and various other sources. It has been noted that the highest stress occurring as part of the rolling between the rail and wheel contact can be involved in the creation of defects and fatigue failure at the surface area of the rails. In addition, every single type of stress may contribute to failure, defects, and fatigue, but it can be prevented if a strategy is undertaken to reduce the RCF defects and control the crack propagation. Therefore, this paper informs future research to enable the investigation of the relationships possible with material properties, wear reports, maintenance programs, a new design of rail, and the coating of the rail to increase its longevity.

Conflict of interest The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. I.R. Memon, X. Zhang, D. Cui, Fatigue life prediction of 3-D problems by damage mechanics with two-block loading. Int. J. Fatigue. **24**, 29–37 (2002)
2. W. Cui, A state-of-the-art review on fatigue life prediction methods for metal structures. J. Marine Sci. Technol. **7**, 43–56 (2002)
3. X.W. Ye, Y.H. Su, J.P. Han, A State-of-the-art review on fatigue life assessment of steel bridges. Mathematical Problems in Engineering (2014)
4. A. Cini, Scribe marks at fuselage joints: initiation and propagation of fatigue cracks from mechanical defects in aluminium alloys. Ph.D. thesis. Cranfield University (2012)
5. C. Lu, J. Melendez, J.M. Martínez-Esnaola, A universally applicable multiaxial fatigue criterion in 2D cyclic loading. Int. J. Fatigue. **25**, 1364 (2018)

6. L. Reis, B. Li, M. De Freitas, A multiaxial fatigue approach to rolling contact fatigue in railways. Int. J. Fatigue. **67**, 191–202 (2014)

7. J. Shen, Responses of alternating current field measurement (ACFM) to rolling contact fatigue (RCF) cracks in railway rails. PhD thesis, University of Warwick (2017)

8. M. Micic, L. Brajovic, L. Lazarevic, Z. Popovic, Inspection of RCF rail defects—review of NDT methods. Mech. Syst. Signal. **182**, 109568 (2023)

9. R. Lewis, U. Olofsson, Wheel–rail interface handbook. (2009). ISBN 978–1–84569–412–8

10. V. Kazymyrovych, *Very high cycle fatigue of engineering materials*. (Karlstad University Studies, Sweden, 2009)

11. M. Murray, Rail fatigue and the roles of impact forces. in *IHHA 2015 Conference, Perth-Australia* (2015)

12. F.C. Campbell, *Elements of metallurgy and engineering alloys*. (ASM International, Ohio, 2008)

13. S. Marich, *Rail and related track structures*. (Study Notes. Queensland University of Technology, Brisbane, 2007)

14. J. Srivastava, V. Ranjan, P. Sarkar, Contact stress analysis in wheel-rail by Hertzian method and finite element method. J. Inst. Eng. India Ser. (2014). https://doi.org/10.1007/s40032-014-0145-x

15. B. Jagadeep, P.K. Kumar, V. Subbaiah, Stress analysis on rail wheel contact. Int. J. Res. Eng. Sci. Manag. **1**(5), 47 (2018)

16. H.M. Sayed, M. Lotfy, H.N. Zohny, H.S. Riad, Prediction of fatigue crack initiation life in railheads using finite element analysis. Ain Shams Eng. J. **254**, 2332 (2018)

17. Y. Jiang, H. Sehitoglu, A model for rolling contact failure. Wear. **224**(1), 38–49 (1999). https://doi.org/10.1016/S0043-1648(98)00311-1

18. H. Zimmermann, *Die berechnung des eisenbahnoberbaus*. (Verlag von Wilhelm Ernst and Sohn, Berlin, 1888)

19. E. Magel, P. Sroba, K. Sawley, J. Kalousek, Control of rolling contact fatigue of rails, center for surface transportation technology, National Research Council Canada. (2005)

20. L. Singh, Rolling contact fatigue in rails—an overview. Access: file:///C:/Users/B013A/Downloads/paper_2_lalloosingh.pdf 17 June (2022)

21. S. Bogdański, M.W. Brown, Modelling the three-dimensional behaviour of shallow rolling contact fatigue cracks in rails. Wear. **253**(1), 17–25 (2002)

22. D. Fletcher, P. Hyde, A. Kapoor, Modelling and full-scale trials to investigate fluid pressurisation of rolling contact fatigue cracks. Wear. **265**(9), 1317–1324 (2008)

23. E.K. Gebretsadik, A. Johansson, B. Pålsson, M. Ekh, J. Nielsen, M. Ander, J. Brouzoulis, Simulation of wheel-rail contact and damage in switches and crossings. Wear. **271**(1), 472–481 (2011)

24. A. Kapoor, F. Franklin, S. Wong, M. Ishida, Surface roughness and plastic flow in rail wheel contact. Wear. **253**(1), 257–264 (2002)

25. M. Ker, TMC 226: rail defects handbook. Ver:1.2, Rail Corp Engi. Manual (2019)

26. T.H. Azmalea, Interview with M.N.Tawfik: rail defects, (Ampang, 2021)

27. T.Y. Kim, H.Y. Kim, Three-dimensional elastic plastic finite element analysis for wheel-rail rolling contact fatigue. Int. J. Eng. Technol. (IJET). **6**(3), 1593–1600 (2014)

28. M. Ciotlausa, G. Kolloa, V. Marusceaca, Z. Orbana, Rail-wheel interaction and its influence on rail and wheels wear, in *12th International Conference Interdisciplinary Engineering*, (2019) p.896

29. U. Zerbst, R. Lundén, K.O. Edel, R.A. Smith, Introduction to the damage tolerance behaviour of railway rails–a review. Eng. Fract. Mech. **76**(17), 2563–2601 (2009)

30. O. Orringer, Y.H. Tang, J.E. Gordon, D.Y. Jeong, J.M. Morris, A.B. Pertman, Crack propagation life of detail fractures in rails. U.S. Department of Transportation, FRA (1988) DOT/FRA/ORD-88/13

31. Q.Y. Li, Z.D. Zhong, M. Liu, W.W. Fan, Chapter 14–Smart railway based on the internet of things. Big Data Analytics for Sensor-Network Collected Intelligence. pp. 209–306 (2017)

32. Á.D. Bedoya-Zapata, S. Rojas-Parra, J.H. Díaz-Mazo, J.A. García-Jiménez, J.E. López-Londoño, R.A. Vergara-Puello, M. Palacio, Case study: understanding the formation of squat-type defects in a metropolitan railway. p. 6 (2021)

33. J. Valehrach, P. Guziur, T. Riha, O. Plasek, Assessment of rail long-pitch corrugation. Ser. Mater. Sci. Eng IOP Conf. (2017). https://doi.org/10.1088/1757-899X/236/1/012048

34. J. Srivastava, P. Sarkar, R.K. Meesala, K. Ranjan, A numerical study on effects of friction-induced thermal load for rail under varied wheel slip conditions. Simul. Trans. Soc. Model. Simul. Int. (2018). https://doi.org/10.1177/0037549718782629

35. K. Zhou, H.H. Ding, S.Y. Zhang, J. Guo, Q.Y. Liu, W.J. Wang, Modelling and simulation of the grinding force in rail grinding that considers the swing angle of the grinding stone. p. 137 (2019)

36. W. Zhang, C. Liu, Y. Yuan, P. Zhang, X. Fan, M. Zhu, Probing the effect of abrasive wear on the grinding performance of rail grinding stones. p. 494 (2021)