A review of non-destructive evaluation techniques for axial thermal stress and neutral temperature measurement in rail: Physical phenomena and performance assessment

Chi-Luen Huang, Yuning Wu, Xiangdong He, Marcus Dersch, Xuan Zhu, John S. Popovics

PII: S0963-8695(23)00047-6
DOI: https://doi.org/10.1016/j.ndteint.2023.102832
Reference: JNDT 102832

To appear in: NDT and E International

Received Date: 1 November 2022
Revised Date: 17 February 2023
Accepted Date: 4 March 2023


This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2023 Published by Elsevier Ltd.

Chi-Luen Huang¹*, Yuning Wu², Xiangdong He², Marcus Dersch¹, Xuan Zhu², and John S. Popovics¹

¹ The University of Illinois at Urbana-Champaign
² The University of Utah
* Corresponding author

Abstract

Thermal buckling of continuous welded rail (CWR) has been a long-standing challenge for the railroad industry because of the high derailment rate and the associated social, economic, and environmental impacts it causes. Rail buckling is generally attributed to excessive thermally-induced axial compressive stress developed in the rail from high temperatures. Knowing the rail thermal stress or its rail neutral temperature (RNT) is critical for safe and efficient rail system operation. There has been great interest and much work on the development of nondestructive evaluation (NDE) techniques to estimate rail thermal stress and RNT in situ. This paper reviews the findings and conclusions from research about NDE approaches for estimating rail thermal stress or RNT, emphasizing the physical phenomena and performance interpretation related to each of the approaches. We identify the type of reference measurements each technique relies on and tabulate this information showing key assumptions, performance, and limitations for each technique.

Keywords: Rail neutral temperature, thermal buckling, NDE, RNT, CWR

1 Introduction

1.1 Problem Significance

Modern railways have widely adopted continuous welded rail (CWR) because they support higher transport speeds and require less maintenance compared with jointed track [1], [2]. However, due to the lack of expansion joints, CWR is prone to develop internal stresses as a result of restrained free thermal expansion and contraction in the axial direction in response to rail temperature changes. When CWR is exposed to sunlight and develops high rail temperature, especially during summer months, rail buckling may occur owing to the buildup of compressive stress that exceeds the lateral track strength. Conversely, when rail temperature drops, for example during winter months, a rail pull-apart may occur owing to the buildup of tensile stress.
that exceeds the axial track strength. The sense and magnitude of the built-up rail axial stress depend on the rail temperature relative to the set rail neutral temperature (RNT). The RNT is the temperature at which the rail is free of axial stress, also known as the stress-free temperature (SFT).

Rail buckling or pull-apart can seriously jeopardize railway safety and operation. Several sources identify “track alignment irregularities (buckled/sun kink)” as a leading cause among track-related factors that cause train accidents [3]–[5]. An analysis of the FRA accident database between 2006 to 2015 for mainlines and sidings further identified the causes of derailment associated with tracks or mechanical systems [6]. The results indicate “broken rails or welds” and “buckled track” cause both high-severity accidents (above-average number of cars derailed per accident) and occur relatively frequently (Figure 1). Furthermore, between January 2007 and December 2017 the damage resulting from excessive rail deformations is estimated to be over $131 million. The primary cause of rail buckling or pull-apart events is insufficient rail stress management. Consequently, minimizing risk of rail buckling and pull-apart due to thermal stress is important to ensure rail safety.

Climate change, which implies rising average temperatures, more frequent extreme heat events, and rapid and extreme temperature fluctuations over the short term, can aggravate restrained rail thermal expansion and this adds more challenges to efficient and safe rail operation. The climate change indicators published by Environmental Protection Agency (EPA) [7] suggest that the average temperature over 48 contiguous states in the U.S. has risen at a rate
of 0.31-0.54°F (0.17-0.3°C) per decade since 1979, which represents a nearly tripled rate of that prior to 1970, unusually hot summer days have become more common over the last few decades, and the occurrence of extreme heat events in major U.S. urban areas has risen from 2 per year during the 1960s to 7 per year during the 2010s. A study on transport policies [8] analyzed the potential vulnerability of the U.S. rail system to projected temperature increases from climate change. It concluded that rising temperatures will lead to long train delays under current operating policies, and more frequent thermal buckling in extreme cases. Depending on the climate models considered, projected cumulative impacts based on delay-minute costs range from $103 to $138 billion by 2100. In other words, climate warming is a foreseeable contributor to rail thermal buckling in the future and will exacerbate the issue.

### 1.2 Management of Rail Thermal Stress and Neutral Temperature

In response to the need to manage thermal stresses in rails, techniques for measuring rail thermal stress and RNT have been developed. Assuming axial deformation of a rail is fully constrained, the developed thermal load is proportional to the difference between RNT and the in situ rail temperature following the linear thermal expansion relation [9]

\[
P = AE\alpha(T - T_{RNT}),
\]

where \( P \) is the thermally-induced axial force, \( E \) the Young’s modulus, \( A \) the cross-sectional area of the rail, \( \alpha \) the coefficient of linear thermal expansion of the rail steel, \( T \) the in situ rail temperature, and \( T_{RNT} \) the RNT. The initial RNT is a design parameter that is established during track installation; it equals the installation rail temperature if no prestress is involved, or the computed stress-free temperature when a controlled pretension is introduced.

The RNT evolves after initial construction because of the interaction of a rail with the ambient environment and supporting track structure (fasteners, seats, ties, and ballast) and disturbances by maintenance activities [10]–[12]. Continuously changing RNT over time implies the presence of additional structural resistance other than thermal expansion. As demonstrated in Figure 2, the evolution of RNT over two years on a revenue-service CWR track exhibits global decline after the rail destressing procedure applied in August 2019 atop some mild seasonal variations. This type of behavior has been reported elsewhere [10], [13]. CWR rails with a low RNT would develop high thermal stresses on hot days that could possibly trigger buckling [14]. Recognizing this fact, rail maintenance engineers typically set the initial RNT, or re-establish the RNT in an existing track with a destressing activity, to be within the range of 90 to 110°F (32.2 to 43.3°C) [1], [15] in order to reduce the potential of excessive axial stresses.
Figure 2: The RNT record of an instrumented revenue-service CWR track at Streator, Illinois showing RNT variations over a two-year period. The initial RNT soon after rail destressing is about 92 °F (33°C). The RNT is calculated based on the measured rail temperature and axial strain provided by an RNT monitoring system.

Rail destressing [16] is a common approach to manage RNT that encompasses rail cutting, de-anchoring (releasing the rail from constraints), rail pulling (pretensioning to close the gap when it is large), reconnecting by welding, and re-anchoring. Cutting a rail would, in theory, release the axial stress to zero nearby the cut location, and the rail temperature at this moment would define the RNT. Strain-gauge-based RNT monitoring systems [17] are typically used to provide accurate and reliable estimates of RNT and rail stress over time. The systems generally rely on strain gauges mounted on the rail web and adopt a Wheatstone bridge circuit, where gauge outputs caused by lateral bending and temperature are compensated for. While this technique has been used for RNT monitoring and research, the overall procedure is time- and labor-intensive and destructive, provides information only at the location that the strain gauges are mounted, and requires a known zero-stress state (e.g., through rail cutting) be established during the monitoring period. Thus, there is a need for a practical technique which provides accurate measurement without these undesirable characteristics.

Nondestructive evaluation (NDE) approaches offer a potential measurement solution for rail stress management. However, existing NDE approaches for estimating rail axial thermal stress/RNT typically require comparisons between the in situ measurement and a ‘reference’ measurement (e.g., measurements carried out at different stress/RNT/temperature state) or a reference model (e.g., numerical simulation results). A reference measurement carried out at a known zero-stress state is straightforward to use, but this state is rarely available in practice because it only occurs at initial rail installation or after rail destressing, and both cases are neither feasible nor economical to carry out frequently. Even if available, the user should understand that the RNT of the system is expected to change over time; this is seen in Figure 2 where the track RNT reduces over 10°F (5.6°C) over a four-month period. Reference data provided by numerical simulation and/or analytical models can also suffer uncertainty because of material property variability and the complex interaction between a rail and the supporting track structure and environment (e.g. tie-to-tie spans, fastener and support conditions, and rail head wear). Such
interaction over time is difficult to predict. A reference-free NDE approach that can accurately quantify the rail axial thermal stress/RNT estimation is needed but has yet to be developed.

1.3 Scope of this Paper

This review paper summarizes selected significant research on NDE techniques for rail axial force, axial stress, and/or RNT estimation in CWR. In section 2, established methods are presented in terms of physical phenomena that the individual techniques are based on; key assumptions, limitations, and performance for each technique are presented in detail. More recently developed and emerging methods are also reported. In section 3, the methods are generally evaluated and discussed in terms of criteria that represent an ideal stress/RNT measurement method with regard to accuracy and application in the field. Here we present a list of those pertinent ideal criteria derived from references [18] and [19]:

- Nondestructive to the rail track structure
- Able to measure absolute axial stress or force
- No need to unfasten a rail or disturb the rail structure
- Easily portable and field deployable
- Applicable for all types of rail models/profiles and wear profiles
- Not affected by rail curvature or configurations
- Insensitive to the microstructure variations and residual stresses in rail steel caused during manufacture
- Adequately accurate: RNT accuracy of ± 10°F (±5.6°C)
- Able to measure from point to point fast enough to map out force/RNT along the rail, as the neutral temperature may be variable from location to location.

While other existing reviews of rail stress measurement methods are available [20], [21], this paper focuses on the full range of technologies developed from the late 1970s to 2020, with an emphasis on physical phenomena and performance assessment. Furthermore, this review recognizes the importance of the type of reference measurement needed within each NDE technique. To the best of the authors’ knowledge, such perspectives are not discussed in previous studies or reviews.

2 Review of NDE for Rail Thermal Stress and RNT Measurement

This section reviews substantial works related to the development of NDE techniques for rail thermal stress and RNT measurement. The techniques are grouped by the adopted physical phenomena, while the emerging techniques are reported in a separate sub-section.
2.1 Beam-column deflection response

An approach to quantify RNT, originally proposed and validated by Kish and Samavedam [9], relates vertically applied uplift load and rail deflection with axial rail force through the beam-column deflection response model [22]. In this approach, the rail is restrained at two points at a set distance, a concentrated vertical load is applied at the center of the span, and the vertically applied load and resulting deflection are measured. This load-deflection relationship is given by the expression:

\[
\delta = \left(\frac{\lambda Q L^3}{EI}\right) \times \left(\frac{1}{1 - P/P_c}\right),
\]

where \(\delta\) is the vertical deflection at mid-span in response to the applied load \(Q\) at mid-span, \(E\) the Young’s modulus, \(I\) the moment of inertia of the rail, \(\lambda\) a coefficient depending on the assumed end constraints, \(P\) the axial load (where positive \(P\) indicates compression), and \(P_c\) the critical buckling load for the beam-column of an unfastened length \(2L\) for the assumed end condition. The rail deflection response is directly affected by the axial load in the rail. A compressive axial load will increase the deflection (\(\delta\)), while tension will reduce it (Figure 3) for a given set of end conditions, applied vertical load (\(Q\)), and rail properties. This approach determines the axial force by comparing in situ load-deflection (\(Q\)-\(\delta\)) relationships with the beam-column model-enabled reference at known stress states. Following the proposed method, Kish et al. [19] implemented this idea by a measurement system mounted on a track loading vehicle developed by the Volpe National Transportation Systems Center (VNTSC). The system test reported high consistency and accuracy in predicting the axial load-deflection response. On tangent (straight) tracks, the system achieves an axial force estimation accuracy of \(\pm 55.6\) kN (\(\pm 12.5\) kips), with an equivalent RNT accuracy of \(\pm 2.5^\circ\text{C} (\pm 4.5^\circ\text{F})\), for AREMA 136RE rail, but lower accuracy on curved tracks.

![Figure 3: Theoretical relationships between the rail response deflection (\(\delta\)) and the applied load (\(Q\)). \(T\) and \(P\) are the axial tensile and compressive load, respectively. Figure source: reference [23].](image-url)
Subsequent advancement of this technique led to the commercial development of the Vertical Rail Stiffness Equipment, known as VERSE®, which has been validated and adopted by multiple railroads [24] and engineering services providers. Despite being more convenient to operate than rail cutting, the method has limitations and is still time-consuming to execute. It requires unclipping the rail to a length of 30 to 40 ft (roughly 9 to 12 meters) on both sides of the lifting point [24], although an unclipped length of 15 meters is suggested in the manufacturer manual [25]. Furthermore, VERSE® is only applicable when the rail is in tension. Finally, practitioners have reported that the accuracy of VERSE® is limited in tight curves, where lateral loads are typically the highest, especially for those with curvature less than 700 meters [25].

2.2 Mechanical Vibrations

Several researchers have investigated the use of mechanical vibrations as a measurement approach for rail axial force and RNT. During the 1980s through the early 1990s, extensive works were dedicated to modeling the dynamic behaviors of rails in a track system. These studies were initially motivated by understanding wheel-rail interactions in search of impact and noise mitigation solutions. Early efforts include the work by Grassie et al., who studied the dynamic behavior of a track system for vibration frequencies up to 1.5 kHz in responses to vertical, longitudinal, and lateral excitations [26]–[28]. Their model incorporated the supporting structure (e.g. rail pads, rail crossties, and ballast) in addition to the rail. Grassie’s model demonstrated the advantage of dynamic analysis of track systems by representing the system as a Timoshenko beam with discrete support conditions at the elastic rail pads. Another important pioneering study by Thompson [29] modeled the vibration behavior of infinite-length rail through finite element analysis. Thompson built a rail model with finite length that was composed of Timoshenko beam elements for the rail head and thin plate elements for the remaining parts. He extended the model to infinite length using periodic structure theory. Thompson’s models predicted dispersions and dynamic responses of free and supported rails that generally agreed with experimental observations. These foundational efforts contributed to improved understanding of the dynamic behavior of rails. In the following sub-sections, we review the studies focusing on the use of resonant frequencies and dynamic torsional rigidity for rail axial force/stress and RNT measurement.

2.2.1 Vibrational resonance frequency

The influences of axial load on flexural resonant modes of vibration in a beam structure have been investigated using Euler-Bernoulli (E-B) theories. Pioneering work by Lusignea et al. [30] analytically and experimentally investigated the effect of axial compressive load on the flexural vibration frequencies of rail of finite length over a frequency range from 1 to 10 kHz. They aimed to quantify the axial load by comparing measured resonant frequencies and phase velocities with the ones collected at the reference zero-stress state. This study reported that resonant frequencies shift with changing axial loads but acknowledged the significant influence of changing end conditions during the loading process. Flexural resonances of finite-length rail
were investigated by Boggs [31] and Béliveau et al. [32], [33], and effects from rail supports were taken into account in the derivation of analytical models and in their experiments. Multiple analytical rail track models, which yield resonant frequencies as a function of axial force and vertical stiffness of the support, were developed. Those analytical models were used to determine a relationship among support stiffness, rail axial force, and measured frequency of a given resonance, which forms a curved surface in parameter space, an example of which is shown in Figure 4 for a simulated beam. The investigated resonances, including the first eight modes, demonstrate a range of sensitivities to support stiffness and/or axial force. Given the established relationship for each mode, one can determine a contour line which identifies possible combinations of support stiffness and axial loads from the measured frequency of that mode. The intersection of all the developed contour lines from multiple measured mode frequencies defines, in theory, the estimated support stiffness and axial load. This estimation approach was similar to the work by Livingston et al. [34]. Boggs concluded that the modified E-B beam models, which account for rotatory inertia and shear deformation, provide better prediction for high-frequency vibrational modes with a wavelength shorter than twice a span between ties. The reported average estimation error of Boggs’ approach tested on a steel I-beam is 13.6 kips (60 kN) or 6 ksi (41 MPa) calculated based on its cross-sectional area. A possible source of error was attributed to the unknown lateral stiffness of the end supports.

Figure 4: The effects of supporting stiffness and compressive axial load on the frequency of one flexural resonance mode. Figure source: reference [31]

These studies generally measure resonant frequencies up to a few hundred Hertz, which may limit capability of RNT estimation in a practical railroad environment. In realistic railway track structures, tie-to-tie variations in terms of lateral, longitudinal, and vertical support conditions
which may disrupt and influence low-frequency responses and thereby introduce errors. CWR typically exhibits much fewer excitable resonant modes compared to short lengths of rail used in laboratory tests [35]. Laboratory tests on a rail of finite length illustrate that numerous factors may affect vibration behavior. For example, Damljanović and Weaver [36] reported a resonance near 200 Hz increased in frequency with increasing uniaxial compressive load, which contradicts the implications of the E-B beam theory. They found the considered resonant frequencies are highly sensitive to changes in the rail system, including variations in rigidity due to changes in load, support conditions, and locations of the vibrational shaker and sensor. Variations among rail spans are largely responsible for the difficulty of reproducing a generic damping mechanism from track structures, which are not sufficiently captured in a laboratory setup or in numerical models especially for frequencies below 1 kHz.

2.2.2 Dynamic torsional rigidity

The use of vibrational torsional rigidity has been investigated for RNT estimation [37]. The technique is widely known by the commercial name D’stresen. The Transportation Technology Center (TTC) evaluated its performance for RNT estimation and reported the accuracy, limitations, and operational guidelines [38]. The test configuration, partially shown in Figure 5 (a), includes a 0-90 Hz variable speed shaker, tune bar (TB) with an accelerometer, magnetic rail temperature probe, and data acquisition system. The technique does not require labor-intensive preparation such as rail cutting or clip unfastening across a certain length. During the test, a speed-variable shaker clamped at a rail head induces rotational rail deflections with low amplitude. It consequently excites lateral rail head vibration, the vertical component of which is measured by the attached tune bar and accelerometer. The peak amplitude of TB vibration along with the rail temperature, measured by the temperature probe, are recorded and tracked across a range of rail temperatures that aims to capture the RNT within that temperature span.

![Figure 5: (a) D’stresen measurement system and (b) hypothesized relationship between TB amplitude and rail temperature. Figure source: reference [38]](image-url)
This technique is based on the hypothesis that rail thermal loads (either tensile or compressive) induce lateral and torsional stiffness in the rail and add resistance to the rotation generated by the shaker. The higher the thermal load (either tension or compression), the more resistant the rail is to torsional motion. At the zero-stress state or RNT, the rail will demonstrate the minimum torsional stiffness and thereby maximum rail rotational deflection, which induces the highest TB deflection amplitude as illustrated in Figure 5(b). This stiffness change is postulated to stem from two mechanisms: the contact pressure between rail base and lateral constraints in the rail seat as temperature changes, and increased bending stiffness when subjected to increasing tensile load. The latter mechanism does not contribute to the behavior when the rail is subjected to a compressive load, so in that case only the rail seat contact mechanism contributes to bending stiffness. By monitoring rail temperature and TB vibration amplitude over time with changing rail temperature, the rail temperature corresponding to the maximum amplitude will be identified as the RNT.

The reported RNT values measured at TTC generally fall within the range of ± 8.3°C (± 15°F) as compared with the reference values provided by strain gauges and the VERSE® uplift system. Although this range is generally acceptable by TTC’s standards, the performance of D’stresen may be inherently limited by its strong reliance on the unverified rail temperature-TB amplitude model shown in Figure 5(b). Engineering judgment is also required to determine whether the rail is in tension or compression. Furthermore, the system performance has been reported to be notably affected by the conditions on or around the rail (e.g. spiked rails, rails with elastic fasteners, rail seats, tie types and conditions, etc.), and it cannot effectively sense the change in compression at tangent rails when the base is highly confined to rail seats. Although not specifically noted by Read [37], [38], the fact that D’stresen operates at a low-frequency range and promotes a mixture of torsional and bending modes, these vibrational motions may be disrupted by constraints at the rail foot.

2.3 Mechanical waves

Mechanical waves provide a basis for a wide range of NDE techniques, including stress or load measurements in different materials. Here we review mechanical wave-based methods that have been applied to assess rail axial force/stress and RNT.

2.3.1 Acoustoelasticity of Ultrasonic bulk and Rayleigh surface waves

Since the early 1960s, researchers have explored the utility of acoustoelasticity and ultrasonic birefringence [39] to measure stress states in metals [20]. Acoustoelasticity was first considered for rail steel in the late 1970s [40] and was once deemed a solution for rail force measurement [41]. The term acoustoelasticity generally refers to the phenomenon that the velocity of propagating mechanical waves changes when the medium is subjected to static elastic deformation [40]. Thus, by measuring changes in wave velocity one can estimate changes in applied stresses within a material. Assuming uniaxial stresses and small wave velocity variations, Egle and Bray [40] established theoretical acoustoelastic and third-order elastic constants by
modifying a model from Hughes and Kelly [42]. Steel bar samples machined from rail head and web sections were tested under applied tensile and compressive uniaxial loads. The relative velocity changes of longitudinal and shear waves, propagating in directions both parallel to and perpendicular to the load path, were measured to compute the acoustoelastic constants. As predicted by the modified model, the measured relative velocity changes are linear functions of strain. They reported that among all the considered modes, longitudinal wave velocity propagating along the load path \((V_{11})\) provided the highest acoustoelastic constant and thus highest sensitivity to load, as shown in Figure 6. Subsequently, Egle and Bray extended their study to rail axial stress measurement using stress-sensitive longitudinal wave velocity [43]. A relationship between longitudinal and shear wave velocities at the zero-stress state was first hypothesized. Because shear waves exhibit relatively low sensitivity to axial stress change, the \textit{in situ} shear wave velocity is assumed to remain constant with changing stress state and thus can be used as a reference to determine the longitudinal wave velocity at zero-stress state, given the assumption that the bulk modulus of the rail is constant. They designed a prototype to collect both longitudinal and shear waves propagating along the axial load direction. Their approach assumes that one can estimate the absolute rail stress by the difference between the \textit{in situ} longitudinal wave velocity and the calculated longitudinal wave velocity at zero-stress state from the measured shear wave velocity, given the acoustoelastic constant (stress sensitivity) of the longitudinal wave estimated by Egle and Bray [40]. However, their laboratory test results suggested the hypothesized longitudinal-shear wave velocities relationship is invalid. They sampled longitudinal and shear velocities at the web section of unstressed new rail samples, where no clear agreement between measurements and hypothesized velocity relationship was identified. Furthermore, they found that the material processing treatment of the rail steel has a significant effect on stress wave velocities. They concluded that residual stress and microstructure variations could influence the velocity measurement, and thereby, the acoustoelastic technique cannot reliably measure absolute stresses. Researchers also reported that wave velocities in rail are temperature-sensitive [44], [45], but it is unclear whether the temperature influence on ultrasonic bulk wave velocities was considered in Egle and Bray’s field work.
More recently, the acoustoelastic effect on Rayleigh surface wave polarization, defined by the ratio of maximum in-plane and out-of-plane displacement components, was investigated. Junge et al. [46] developed an analytical model to predict the influence of applied uniaxial stress on Rayleigh wave velocity and polarization. Their model results demonstrated that the relative Rayleigh wave polarization is more sensitive to applied stress than the relative Rayleigh wave speed in mild steel. Gokhale and Hurlebaus [47] measured Rayleigh wave polarization of an unstressed rail by extracting in-plane and out-of-plane components with a laser Doppler vibrometer and experimentally observed the acoustoelastic effect on a rail sample [48]. A linear correlation between the polarization extent and uniaxial tensile load was observed. Its performance was evaluated by scanning a 40-ft long track testbed and extracting Rayleigh wave polarization along the rail web using two laser Doppler vibrometers mounted on a cart [49]. The measurements were conducted while the rails of the testbed were subjected to varying axial compression levels. The potential influence of rail temperature was not reported. The measurements showed inconsistent Rayleigh wave polarization results at the zero-stress state, and there was no clear correlation with the applied stresses. Moreover, researchers have shown that residual stresses in rails would also affect Rayleigh wave polarization [50], and thereby the influence of residual stress must be appropriately compensated when using Rayleigh wave polarization for rail axial stress measurement.

Overall, ultrasonic bulk wave and Rayleigh wave measurements operating at MegaHertz frequencies can support relative stress measurement by detecting changes in propagating wave propagation as the medium is subjected to a static axial load. However, both bulk wave velocities and Rayleigh wave polarization were reported to be sensitive to variations in material textures, composition, and residual stresses as a result of the manufacturing and heat treatment.
processes. Subsequent attempts considered the acoustoelasticity of surface skimming horizontal shear waves [51] and ultrasonic backscattering wave phenomena [52], [53]. Estimating the absolute stress state by acoustoelasticity has so far been effective only in laboratory tests under controlled environments and materials.

2.3.2 Acoustoelasticity of guided waves

Researchers have also studied the sensitivities of guided waves (other than Rayleigh waves) to axial load in rails in terms of the change in the relations between wave velocities and frequency; such behavior is also known as dispersion. An infinite number of guided wave modes theoretically exist for each waveguide, and each has its unique dispersive relation. Initially, researchers implemented an acoustoelasticity approach by investigating the effects of axial load on phase/group velocity or wavenumber of guided waves in rail. A 1979 report on rail dynamic response by Luisignea et al. [30] pointed out that the vibration behavior of a finite-length piece of rail differs from CWRs, and thus measurements on CWRs will “probably involve measuring the phase velocity of flexural waves rather than resonant frequencies.” They experimentally investigated the influence of axial loads on guided wave modes in a 10-ft rail sample subjected to a maximum compressive axial load of 110 MPa. Guided modes with vertical and lateral bending motion within the frequency range of 1-10 kHz were considered, but no significant or consistent changes in phase velocity were observed with increased axial loads. Their observations can be attributed to relatively small applied compressive load, limited stress sensitivity supported by the identified modes within the frequency range, and lack of an advanced data acquisition system. More recently, wave dispersion phenomena in rails were more elaborately modeled and experimentally verified by Thompson [54] and Lanza di Scalea [55], [56], providing additional understanding about guided waves in rails.

In addition to experimental studies, numerical simulations for guided wave dispersions in rails, and their sensitivity to stress, have been considered. Early models by Thompson were developed to predict the frequency-wavenumber relationship in rails using finite element simulation [29]. Specifically, Thompson expanded his investigation to track dynamics, where he compared three existing two-layer track models accounting for rail pad, ballast, and tie structures in addition to the rail [57], and evaluated those models with experiments [58] providing a broader understanding of the effects of the elements in a track system on its dynamic responses. Similarly, Knothe et al. [59] built finite ‘stripe’ models and compared the predicted dynamic responses and corresponding dispersion relations among those models. Gavrić [60] developed the basis for the semi-analytical finite element (SAFE) simulation approach that accounts for axially infinite geometry of CWR rail by assuming harmonic wave propagation in the longitudinal direction. This approach reduces the computational cost by only discretizing the waveguide’s cross-section. Damljanović and Weaver [61], [62] expanded the SAFE method for both propagating and evanescent modes in rails. Considering acoustoelasticity of guided waves in rails, Chen and Wilcox [63] developed a finite element model to investigate changes in dispersion relations in terms of the phase and group velocities on a rail subjected to different tensile load levels. A two-step finite element model was proposed to estimate the effect of load on dispersion...
relations in rails, where a statically loaded short piece of rail with an updated global stiffness matrix was calculated for eigenfrequency analysis. Dispersion curves were established by populating frequency-wavenumber pairs using rail models with different lengths. Loveday [64] developed an automated modeling approach by extending the SAFE method to consider rails subjected to axial load. The axial load effect on wave propagation was modeled by adding a longitudinal initial stress term to the linear strain energy, which modified the stiffness matrix with input from the axial load. Bartoli et al. [65] adopted the same strategy and predicted dynamic responses when a rail is subjected to static axial load and transient excitation. Their simulation results showed that the acoustoelastic effect measured by phase velocity is predominant in the low-frequency range but sensitivity decreases with increasing frequencies. As shown in Figure 7, among the considered modes, the dispersion curve of the fundamental horizontal/lateral flexural mode demonstrated the highest sensitivity in response to applied axial load within frequencies ranging from 0.1 to 1 kHz. The study also suggested that stress-sensitive modes with minimal rail foot motion are preferred for rail stress measurement to avoid influences from fastener and tie boundary conditions.

![Dispersion curves](image)

Figure 7: (a) Comparison of the SAFE model results in terms of phase velocity-frequency between unloaded and loaded rails, and (b) the simulated time histories for unloaded and loaded rails showing stress sensitivity of wave velocity. Figure source: reference [65]

### 2.3.3 Model-based rail force estimation using guided waves

Understanding both the stress and temperature sensitivities of guided waves, one can in theory estimate the rail stress state by taking in situ measurements and reference measurements at its zero-stress state or RNT. Because zero-stress state reference measurements are rarely available, the desire for reference-free measurements led to the development of model-based force estimation utilizing load-sensitive guided modes. Reference measurements are replaced by predictions of the zero-stress state based on analytical or finite element models. A notable example of model-based approach for guided waves in slender aluminum bars is provided by Albakri et al. [66], [67].
A model-based rail axial force measurement framework, developed by Damljanović and Weaver [36], is summarized in Figure 8 (a). In this approach, the wavenumber, denoted by \( k \), of the low-frequency lateral flexural mode in a rail is tracked. Exploiting the sensitivity of lateral flexural rigidity to axial load, they extracted the wavenumber of the lateral flexural mode in rail by spatial sampling along the rail axis over an unsupported span, which is in principle independent from varying boundary conditions. This framework is rooted in wave dispersion obtained by the SAFE model from their earlier work [61], [62]. While the data collection scheme used a steady-state excitation similar to a vibrational test, this approach is categorized as a wave propagation approach because it relies on guided wave dispersions and signal processing techniques where both propagating and non-propagating wave modes were considered.

The underlying idea is to use the change in wavenumber (\( \delta k \)) of the lateral flexural mode and its force sensitivity \( \frac{\delta k}{k} \) for rail axial force estimation, as shown in Figure 8 (a). The term \( \delta k \) is defined as the difference between the in situ wavenumber, \( k(P) \), and the reference wavenumber at the zero-stress state, \( k(P = 0) \). Force estimation relies on an equation derived using E-B beam theory:
\[ P = 4EA r_g^2 k^2 \left( \frac{\delta k}{k} \right), \]  

where \( P \) is the axial load, \( E \) the Young’s Modulus, \( A \) the area of cross section, \( r_g \) the radius of gyration of cross section, \( k \) the lateral bending wavenumber, and \( \delta k \) the change in wavenumber. The \( P \) vs. \( \frac{\delta k}{k} \) relation in theory behaves linearly below 100 kips for an AREMA 136RE rail and intersects at the origin. By measuring the in situ wavenumber \( k(P) \) and modeling the reference wavenumber \( k(P = 0) \), the axial force \( P \) could be estimated with a precision of ±26 kips that is equivalent to and RNT accuracy of ±10°F. However, the approach is challenging in two aspects. First, conventional E-B beam theory was adopted to quantify the sensitivity of wavenumber to axial force, but is inadequate to predict \( k(P = 0) \) for rails with complex cross-sectional geometries. Second, it is difficult to accurately extract the \( k(P) \) of the lateral flexural mode experimentally owing to the simultaneous presence of multiple modes in a measured response. Damljanović and Weaver adopted the SAFE method to overcome both challenges. As shown in Figure 8(b), they obtained the complete dispersion relations for 136RE AREMA rail with propagating and evanescent modes, and used the dispersion relations to determine the wavenumber of the lateral flexural mode at the zero-stress state, \( k(P = 0) \). Furthermore, they decoupled the lateral flexural mode from other undesirable modes using a least-square-based fitting algorithm. This process relies on SAFE-predicted wavenumbers and eigenmodes of multiple propagating and evanescent modes [61], [62].

Based on their force estimation framework, Damljanović and Weaver designed a deployable sensing and data acquisition system, which is shown in Figure 8 (c). The technique was evaluated by experiments on rail segments in the laboratory and at field sites [23], [35], [36], [68]. The laboratory study reported a discrepancy between the predicted and estimated load: although a linear relationship between the axial load and the change in wavenumber (\( P \) vs. \( \frac{\delta k}{k} \)) was observed, the value at zero-load exhibited a significant \( \frac{\delta k}{k} \) offset that decreases the estimation precision from ±26 kips to ±36 kips. The performance of the framework depends on the accuracy of lateral flexural wavenumber at the zero-stress state, the precision of measured/extracted wavenumbers and mode shapes for the considered modes, and the validity of E-B beam assumption on short thick rail segments. Their lab test results also showed the importance of accurate geometry and material properties. Excellent performance was reported on a circular rod and an unworn 136RE rail, within an error of ±25kips (±111.2 N) or ±9.5°F (±5.3°C) in RNT. In contrast, a worn 136RE rail exhibited poor accuracy [35], [68]. A follow-up work by Kjell and Johnson [69] identified the possible error sources of this approach using an 11-meter long rail with a full track configuration in a laboratory; their results demonstrated a general RNT estimation accuracy of ±9°F (±5°C). They confirmed that this approach relies strongly on a realistic and accurate representation of the rail, including the actual cross-sectional geometry and material properties. Note that the potential temperature effect on wave dispersions in rails was not considered in these studies. Nonetheless, previous research demonstrated that the temperature influence on guided wave propagation is relatively insignificant in the low-frequency range [70], [71].
2.3.4 Nonlinear ultrasonic guided waves

Nucera and Lanza di Scalea exploited nonlinear ultrasonic guided waves for RNT estimation [72], [73]. Their approach utilizes the nonlinear second harmonic that is promoted at a frequency of $2\omega$ in response to a fundamental ultrasonic excitation at $\omega$. They postulated the nonlinear phenomenon is induced by the potential energy stored in the material owing to constrained thermal expansion, rather than a finite strain as in acoustoelasticity. Considering the lack of global deformation in CWR due to constraints, they proposed a new theoretical framework to model the nonlinearity of wave propagation in a medium subjected to constrained thermal expansion. They derived a theoretical model based on interatomic potential and verified the existence of nonlinearity arising from the strain energy caused by constrained thermal expansion [72]. They then expanded the investigation to RNT measurement [73]. The magnitude of the second harmonic is proportional to the theoretical nonlinear material parameter. They found that the minimum of the nonlinear parameter when measured across different stress states corresponds to the zero-stress state of CWR, where the strain energy induced by the prevented thermal expansion vanishes. A nonlinear SAFE approach was formulated to model the fundamental and second harmonic of guided waves in CWR. Based on the SAFE model results, a propagating guided wave mode at 200 kHz with energy mainly confined in the rail web was selected for nonlinear resonance analysis. A resonant second harmonic mode at 400 kHz was identified, which also exhibits a similar energy concentration in the rail web area. These selected modes can inherently avoid the influences from conditions of rail head (worn geometry and residual stress) and rail foot (supports and fasteners). A prototype, the Rail-NT system, was developed to promote the fundamental and second harmonic modes of interest. Measurements are carried out on rail at one location across a range of temperatures and the system identifies the RNT at the minimum nonlinear parameter, as shown in Figure 9. Its performance was evaluated by laboratory and field tests and reportedly achieved RNT accuracy within ±2°F (±1.1°C) on CWRS with concrete ties and within ±5°F (±2.8°C) on CWR with timber ties. However, effective measurement using the Rail-NT system requires continuous monitoring over time and changing temperature, during which the rail temperature must cross the RNT point. Moreover, Nucera and Lanza di Scalea observed fluctuation of nonlinear parameters owing to the passage of a train. The frequent and heavy dynamic loads by train traffic compromise nonlinear parameter measurement by inducing changes in rail temperature, stress distribution in rails, and transducer coupling conditions.
Figure 9: Field test results for nonlinear guided waves on CWR on concrete ties. The test was done on a test track at TTC. The force (upper plot) calculated from the calibrated strain readings, and the nonlinear parameter (lower plot) was extracted from nonlinear ultrasonic guided waves at the corresponding time; the minima of nonlinear parameter correspond to RNT. Figure source: reference [73]

2.4 Magneto-elasticity and magnetic Barkhausen noise

Researchers have investigated the influence of mechanical loads on magnetic parameters in ferromagnetic rail steel. All the approaches described in this sub-section are based on magnetic hysteresis (magnetic flux density-magnetic field strength or B-H) behavior. Magneto-elastic methods generally monitor the stress-sensitive magnetic permeability and coercivity, which are the slope and points of intersection with zero flux or magnetization in a B-H curve, respectively. On the other hand, magnetic Barkhausen noise (MBN) quantifies the discrete noise events (sudden jumps) in the B-H curve caused by domain wall movements. The theoretical basis and experimental observations of stress sensitivities for both phenomena in steel can be found in [74]–[78]. Based on these phenomena, several magnetic measurement systems have been developed for rail stress and RNT measurements, including the MAPS-SFT (magnetic anisotropy and permeability system - stress free temperature) [78]–[80] and the RailScan system.

The MAPS-SFT system relies on the relationship between stress and magnetic properties to measure the RNT and rail stress [78], [79]. The technique measures a combination of thermal and residual stresses, where the latter components must be removed or compensated for accurate RNT estimation. Thus, it is typical to calibrate the residual stress for a specific rail type
and material. The technique presumes a linear relationship between the vertical and axial residual stress components, which provide a calibration relationship for a specific rail profile and steel grade. The thermal stress is estimated in three steps. First, the data acquisition system measures magnetic field properties at several angles with respect to the rail to predict bi-axial (vertical and axial) stress components. Assuming free thermal expansion along the vertical direction, and therefore no thermally induced vertical stress, the measured vertical stress component reflects only residual stress. Second, the axial residual stress is inferred using the established calibrated relationship between the vertical and axial residual stresses. Finally, the axial thermal stress value is estimated by removing the inferred axial residual stress from the total axial stress component. The TTC evaluated this technique and confirmed its RNT accuracy as $\pm 8\degree F$ ($\pm 4.4\degree C$) compared with RNT calculated directly from strain gauge readings. Although the accuracy is within the acceptable error range of $\pm 10\degree F$ ($\pm 5.6\degree C$) established by TTC, the technique depends on the accuracy of measurement of the vertical-axial residual stress relationship and calibration for baseline residual stress, which varies among rail types, grades, and manufacturers [79].

Another commercial product based on MBN, the RailScan system, was tested on the Darwin-Alice Springs line in Australia [81]. A calibration curve that relates MBN amplitude and uniaxial stress for the specific carbon steel rail was obtained in the laboratory. The rail axial stress, and thereby the RNT, was determined using the measured MBN signals and a lab-established calibration curve. The system performance was evaluated by comparing its results with the ones from the ‘cutting rail and gap measurement method’, which demonstrated a discrepancy of 9.8$\degree$C in RNT. Shu et al. [82] combined metal magnetic memory (MMM) [83] and MBN techniques for thermal stress and RNT estimation. They estimated thermal stress based on a similar approach using the laboratory-established relationship between magnitudes of MBN and given applied stress levels. Additionally, they calculated RNT based on the estimated thermal stress and in situ rail temperature. The estimated RNT was compared with the RNT at installation. Despite the reported good agreement between the measured and the initial RNT, the RNT at installation likely did not maintain for a long time, as mentioned in the introduction section, and thus may not remain valid for future comparisons. On the other hand, additional studies have shown both the residual stresses and rail temperature will affect Barkhausen noise [74], [77]. The MBN signatures-temperature curves must be calibrated accordingly to facilitate accurate thermal stress estimation. It is unclear whether these studies [81], [82] considered the influences of residual stresses and temperature on MBN in the field tests.

2.5 Resonance fluorescence

Photoluminescence piezospectroscopy (PLPS) is a type of fluorescence spectroscopy and has been used for measuring the residual stress in thermally grown oxides. Kim and Yun [84] adopted this technique for rail axial stress estimation. They measured the stress-induced changes in fluorescence responses of aluminum oxide in the alpha polymorphemic phase ($\alpha$-Al$_2$O$_3$), which exists in thermite welds of CWR. The $\alpha$-Al$_2$O$_3$ phase is known for its high fluorescence-reactivity, thus it is favorable for the application of PLPS. In their laboratory test, two ‘fingerprint’ peaks of
$\alpha$-Al$_2$O$_3$ in the fluorescence spectrum were identified, as shown in Figure 10(a). Both peaks are sensitive to uniaxial stress, demonstrating a lower wavenumber with increasing compression, as shown in Figure 10(b). The stress-wavenumber relation is approximately linear within the elastic range. Thus, they assumed a linear relationship between the wavenumber of the fingerprint peaks and absolute axial stress:

$$w = a \cdot \sigma_{abs} + b,$$

(4)

where $a$ is the piezospectroscopy coefficient, $b$ the wavenumber of fingerprints of aluminum oxide at the zero-stress state, and $w$ the in situ wavenumber when subjected to an absolute axial stress level $\sigma_{abs}$. The slope $a$ is determined by the relationship between absolute wavenumber and relative stress (inferred by strain measurement). The intercept $b$ is obtained through spectral analysis on pulverized thermite weld samples, where it is assumed that no stress is present. Once the parameters were determined for each peak, one can estimate the absolute uniaxial stress with in situ wavenumber measurements.

![Figure 10](image)

Figure 10: (a) The characteristic wavenumbers in a fluorescence spectrum from pulverized thermite weld samples and (b) their individual sensitivity to axial compressive stress. Figure source: reference [85]

Based on the success with laboratory measurements, Yun et al. further developed a portable PLPS system with line scanning capability and evaluated the system performance on an instrumented full-size high-speed rail testbed [85]. A section of the CWR was cut, replaced by new rails, and reconnected using thermite welding, where strain gauges and temperature probes were installed to provide ground truth stress values. The measured fingerprint peaks demonstrated a general trend of decreased wavenumbers with increasing compressive stresses, which agrees with the laboratory study. However, the piezospectroscopy coefficients obtained were approximately five times as large as the ones obtained in the lab. Moreover, the uncertainties of the sampled wavenumbers at each fixed stress level led to an axial stress estimation uncertainty of at least ±25 MPa. There are several limitations that prevent PLPS from
high-accuracy rail axial stress or RNT measurements. First, $\alpha$-Al$_2$O$_3$ is randomly distributed in the thermite welds such that the stress transfer mechanism between iron matrix and aluminum oxide particle is unclear; stress concentrations at the microstructural level could affect the measurement accuracy. Second, the adopted reference measurement at the ‘zero-stress state’ from the pulverized thermite weld sample does not represent the actual field state of applied or residual stresses. The presence of excessive residual stresses from other sources in rails will introduce significant measurement errors. Studies [86]–[88] reported a wide range of residual stresses in rails and welds, with extremes of approximately 300 MPa in compression. Residual stresses need to be compensated for to ensure accurate rail axial stress estimation. Furthermore, the piezospectroscopy coefficients obtained from the field test were determined using strain gauges that were not shunt calibrated for temperature compensation. Thermal outputs of the strain gauges could lead to erroneous strain measurements and piezospectroscopy coefficients.

2.6 Emerging technologies

2.6.1 Rail deformation sensing by digital image correction

Several researchers have estimated RNT by monitoring static deformations when subjected to thermal load [17], [88], which typically involves rail cutting or material removal for the reference at the zero-stress state. A recent effort by Knopf et al. [89] proposed a contactless NDE method for RNT estimation that uses a visual data acquisition system to monitor high-contrast patterns on the rail surface to capture rail deformation. The researchers hypothesized that the vertical deformation on the top of a rail head caused by thermally-induced axial stress is nonuniform along length owing to the constraints of discrete supports, thus the resultant shape profile along rail length is related to the extent of developed thermal stress. This hypothesis was verified through a finite element model of a rail segment rigidly attached to the ties. The simulation confirmed the curvature of the rail head shape profile over the tie support changes proportionally with the rail temperature, and reversed curvature occurs as the thermal load shifts from compression to tension (see Figure 11(a)). An experimental study was performed on partially constrained 16-inch (41 cm) 132RE rail segments with elevated temperature to mimic deformations in CWR when subjected to temperature increases. A 3-D stereo-digital image correlation (StereoDIC) system was employed to perform contactless deflection measurements focusing on the regions of interest at the rail web and head. The experimental setup does not accurately represent CWR field conditions regarding rail length, constraints, and access to rail head over supports. Despite the limitations, the experimental study still showed reasonable agreement with the numerical model. The deformed rail head shape profile between supports and its quadratic polynomial fit demonstrated a negative or positive curvature when subjected to compressive or tensile loads, respectively. More importantly, a linear relationship between the curvature and rail temperature was observed. Upon the verification of hypotheses through these numerical and experimental studies, a RNT prediction procedure was proposed, as illustrated in Figure 11(b). First, the rail head shape profile curvatures at two distinct temperatures are measured and used to determine a linear relationship between rail temperature and the curvature. Once the linear function is established by these two points, the
temperature corresponding to zero curvature, or equivalently zero-stress state, is determined by extrapolating to the y-intercept. Furthermore, the rail axial stress can be estimated using vertical strains with a plane stress simplification. The proposed RNT estimation approach was later evaluated with a more realistic finite element model. While the approach demonstrated excellent accuracy for RNT and thermal stress estimation based on the model results, it requires more system evaluation in a field environment. Specifically, it is critical to evaluate the validity of the linear relationship between curvature at rail head over the tie across a wide range of rail temperatures. Influences from support variations, including tie condition, tie space, and fastener types, are yet to be established. Moreover, the adoption of StereoDIC requires rail surface preparation, whose performance could be impacted by unfavorable illumination conditions in the field.

![Image](a) Illustration showing hypothesized curvatures of a rail head top over a tie, and (b) the proposed procedure for RNT estimation. Figure source: reference [89]

2.6.2 Video-based rail vibration measurement

Researchers have developed a contactless approach to infer RNT by monitoring rail vibration through a video-based measurement technology [90]. The approach considers the pixels of video images as virtual accelerometers to extract resonant frequencies and mode shapes from the dynamic motion. A laboratory test was conducted to verify its feasibility on rails. A uniaxial load frame applied stepwise incremental axial loads to a 2.4-meter rail sample. At each load step, an impulse vibration test was performed by impacting the side of rail head to promote the lateral bending mode. Vibrational responses were collected by two accelerometers and a high-speed camera where QR-code targets were attached to the rail head for contrast enhancement. Rail resonant frequencies were identified by performing a fast Fourier transform of the lateral displacement data extracted from video images. A phase-based displacement extraction method [91] was adopted to sense small motions; however, only pixels that exhibit sufficient contrast within the QR code image capture the rail motion, and in this experiment the pixels along the edges (rail-target or rail-background boundary) show the best contrast. The displacements at the pixels close to the edges of the QR code image were extracted. As a result,
two real resonances and an artificial resonance were identified. The two real resonances were confirmed to be lateral and coupled lateral-torsional modes using the motion magnification technique [92]. The frequency of these resonances identified from the contactless video-based method and those measured by accelerometers were in excellent agreement. This image processing method shows the capability of contactless rail vibration measurement and mode shape extraction, which has the potential to be used in vibration-based techniques for RNT measurement. Possible challenges for field implementation are brought by illumination conditions and high-contrast boundary requirements. Furthermore, CWR is typically not as resonant compared with finite lengths of rail that are used in laboratory tests [35], thus CWR exhibits lower vibration amplitude values. The laboratory data in this case would not well represent those from CWR in the field. More research and field tests are needed to verify the system performance in capturing more subtle vibration signals with a higher level of measurement noise.

2.6.3 Electro-mechanical impedance

The electro-mechanical impedance (EMI) technique utilizes the interaction between a piezoelectric element, usually made of lead zirconate titanate (PZT), and a host structure to which the PZT is attached for structural condition assessment. The first analytical model developed by Liang et al. [93] integrated piezoelectric actuators with a spring-mass-damper system to relate the structural stiffness/mechanical impedance with the measured electrical admittance of the piezoelectric element. Giurgiuţiu and Zagrai [94] extended this model by deriving electrical admittance of a bonded PZT element based for the dynamic stiffness of the host structure and the quasi-static stiffness of the PZT. The EMI technique, thereby, shows potential for determining the axial load in a structure considering the relation between the load level and the combined PZT-structure response.

Extensive laboratory efforts have been carried out to study the feasibility of determining the axial load using EMI. Ong et al. [95] investigated the axial loading effect on the dynamic structural stiffness and impedance signature using an E-B beam model. They modified Giurgiuţiu and Zagrai’s model by introducing axial load to an E-B beam with damping. This PZT-structure interaction model demonstrated the capability of the conductance (real part of admittance) of the bonded PZT to characterize axial loads in the host structure. Lim and Soh [96] reported similar findings through a series of experimental and analytical studies considering both flexural and longitudinal modes of vibration. Furthermore, they reported the boundary conditions (clamping force from fixtures) would induce variations in the conductance spectrum. Susceptance (imaginary part of admittance) of the bonded PZT was also investigated for stress measurement. Annamdas et al. [97] demonstrated that externally applied loads could affect both the conductance and susceptance. Their analysis shows susceptance serves as a better stress indicator, especially in the case of simultaneous presence of axial, torsional, and bending loads. After Phillips et al. [98] experimentally studied the stress effect on EMI responses on a steel sample sectioned from rail, Zhu and Lanza di Scalea [99] derived an analytical model for EMI considering the strain transfer mechanism and stress stiffening effect on PZT element. Specifically, the stress effect on the PZT is described using modified piezoelectric constitutive equations,
which incorporate nonlinearity into piezoelectric and dielectric constants. The model predictions were compared with the results from experiments on an aluminum beam and a rail steel sample. The conductance, susceptance, and PZT-circuit electrical resonances responses all demonstrate clear dependence on the applied axial stress, although with different levels of sensitivity. It is notable that the PZT-circuit electrical resonances demonstrate higher stress sensitivity compared to the conductance resonances associated with flexural modes. Although the EMI technique has shown its potential for axial stress measurement, all the associated investigations so far are still at the stage of model development and laboratory tests on general metallic beams. Furthermore, influences from varying temperature and support conditions on the EMI technique still need to be investigated, and uncertainties of material and geometric properties need to be accounted for considering operating and environmental variabilities in a field setup.

2.6.4 Highly nonlinear solitary waves

Nasrollahi and Rizzo explored highly nonlinear solitary waves (HNSW) as a tool for rail stress determination [100]. The HNSWs are compact non-dispersive waves that travel in periodic arrays, such as a chain of steel spheres. The mechanical interaction between two adjacent spheres is primarily governed by Hertzian contact law when specific geometric and mechanical conditions are satisfied [101]. Solitary waves offer unique properties, for example the phase velocity of a solitary wave depends on its amplitude, and the reflected wave speed and amplitude are also affected by the stiffness of the material in contact with the end of the sphere chain. These properties are highly nonlinear and have been utilized for applications in various NDT areas for characterizing structural materials and stress measurements [102]–[105].

A series of studies led by Rizzo [100], [105], [106] investigated stress effects in structural materials using the properties of nonlinear solitary waves. The study developed a measurement approach using a carefully designed L-shape transducer, which accommodates a chain of metal spheres instrumented with piezoelectric elements. The transducer introduces solitary pulses to the test material and measures the reflections from it. The stiffness of the test material can change the properties of the solitary pulses reflected at the interface. During tests, a striker was dropped freely to hit the top sphere, generating a single incident solitary wave (ISW) propagating in the chain. When the ISW arrives at the interface between the last sphere and the testing material, the pulse is partially reflected, forming the primary reflected solitary wave (PSW). If the material is much softer than the spheres, one or more secondary reflected solitary waves (SSW) would form and travel within the chain [106]. Rizzo and co-authors first verified the feasibility of mechanical and thermal stress measurement using HNSW features (the time-of-flight between the arrival of the PSW and ISW, the amplitude of the ISW, and the ratio of the PSW amplitude to the ISW amplitude) on slender steel beams [105], [106]. They then extended the investigation to consider HNSW features on a prestressed rail sample using an L-shaped transducer in contact with the web section. In addition, a numerical model was developed to predict the influence of axial stress on the HNSW features with an assumption about the equivalent length of the rail specimen, which is closely related to rail structural stiffness. Based on model results, the HNSW features are sensitive to uniaxial load for rails with an equivalent length of 3.6 to 4.8 meters. However, due to technical difficulties of uniaxial loading test on long samples, only rail samples
with lengths of 0.9 and 2.4 meters were tested in the laboratory, which were too stiff to show noticeable sensitivity to axial loads. More research and field test efforts are needed to evaluate its performance, especially verifying the assumption about the equivalent length and applying the approach to long-length CWR structures.

3 Discussion

Table 1 summarizes the results and highlights the key assumptions, limitations, and performance of each technique to provide a comprehensive comparison. Because reference measurements or reference models are important in defining the overall performance of the NDE approaches for rail thermal stress and RNT estimation, here we discuss the specific reference measurement type associated with each measurement technique: reference measurement at zero-stress state, model-based reference, or reference measurement at known rail temperature.

The first type of reference is the measurement at the zero-stress state, which typically involves measurements carried out on laboratory samples where the zero-stress state (no resulting axial load) of the sample can be controlled and defined. Nondestructive measurement methods that utilize this type of reference include the magneto-elastic method (section 2.4); the magnetic Barkhausen noise technique (section 2.4), which requires establishing an MBN signature-uniaxial stress curve through laboratory tests on rail steel samples; and the photoluminescence piezospectroscopy technique (section 2.5), which uses laboratory measurements on pulverized thermite weld samples. In these cases, reference measurements are normally obtained on surrogate samples at known zero-stress state, although the variation of microstructure and in situ residual stress in rails could still significantly affect system performance.

The second type of reference relies on numerical or analytical models, where the performance is limited by how precisely the model represents the actual physical state of the rail and the method’s prediction accuracy. Nondestructive measurement methods that utilize this type of reference include the beam-column response method (section 2.1), which estimates axial force using load-deflection relationships assuming the unclipped rail section span behavior is well represented by a slender beam; the acousto-elasticity method for ultrasonic bulk waves (section 2.3.1), which relies on a hypothesized constant bulk velocity and low stress-sensitivity of shear wave velocity assuming the influences of residual stress and microstructure variations are negligible; and the model-based method using guided waves (section 2.3.3), which can provide RNT estimation with acceptable accuracy (±9°F, or ±5°C) when geometric and material uncertainties are well addressed.

For methods that use the first two types of reference, the validity of the reference is important for the performance of the method. However, in some cases maintaining the validity of the reference can be a challenge for rails in a real-world environment. As discussed in the introduction, the RNT of the rail system will likely change over time. That means that RNT reference values measured once on-site may gradually lose validity unless the strain values and
temperature of the track at that location are continuously monitored. References provided by laboratory measurement or numerical/analytical models also need to consider limitations of surrogate samples/models and to accommodate for uncertain track structure interactions and material property variations in the field, unless the employed physical phenomenon is inherently immune from, or minimally affected by, them.

The third type of reference are measurements taken at known rail temperatures, but not necessarily at the zero-stress state. NDE methods that utilize this type of reference include the nonlinear guided wave method (section 2.3.4), where the RNT is determined by identifying the occurrence of the minimum in nonlinear behavior among multiple measurements carried out across a range of rail temperatures; and the static deformation sensing technique using DIC (section 2.6.1), which requires multiple measurements at different rail temperatures to predict RNT based on the assumption of the changing sense of rail head curvature when shifting from tension to compression states, or vice versa. Reference measurements taken at different rail temperatures must be obtained, but for these particular test methods surrogate samples or model-established references are not required. In theory, methods based on these types of measurements are less likely to be affected by variations in residual stress and uncertainties in model parameters.

4 Conclusions

Although many different nondestructive testing methods have been developed across a wide range phenomenological bases, strictly speaking none of the techniques reported in this review satisfy all the criteria listed in the introduction for an ideal technique for rail axial stress/RNT measurement. The VERSE system has been adopted frequently by practitioners, given its reliability, measurement accuracy, and field-deployability. Several other existing techniques, although only satisfying part of the mentioned criteria, provide acceptable accuracy of estimation in the field. The approach using nonlinear guided waves, for example, provides accurate estimation of RNT with a human-portable testing module that is practical for point to point deployment in a field context. However, the method does have drawbacks as previously reported. Even though no one measurement technique satisfies all the criteria for an ideal method, effective and useful results may be obtained by a given method if the limitations and required reference measurements are well understood.

5 Acknowledgement

This work was supported by the U.S. National Academy of Sciences Rail Safety IDEA program, project RS-41. The authors acknowledge and appreciate the input from IDEA program manager Dr. Basemera-Fitzpatrick. This effort is also funded in part by the Federal Railroad Association through contract 693JI621C000025. The opinion expressed in this paper are solely of the authors, and the U.S. National Academy of Sciences, Engineering, and Medicine, The Federal Railroad Administration, and the U.S. Government do not necessarily concur with,
endorse, or adopt the findings, conclusions, and recommendations either inferred or expressly stated in the paper. Finally, the authors would like to thank Professor Francesco Lanza di Scalea, Dr. Robert Wilson, Wesley Mui, Harold Harrison, and Andrew Kish for thoughtful discussions.
### Table 1 Summary of technology performance

* For 136RE rail; ** error from targeted value on a steel I-beam; *** error from targeted values

<table>
<thead>
<tr>
<th>Category</th>
<th>Phenomena / Example test prototype</th>
<th>Key assumptions / reference type</th>
<th>Limitations</th>
<th>Reported performance (thermal stress/force or RNT)</th>
<th>Paper Section / key references</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Static deformation</strong></td>
<td>Beam-column response / VERSE®</td>
<td>A beam-column model simulates the load-deflection relationship for a prestressed beam / model-based reference (analytical model results)</td>
<td>The process is time-consuming and labor intensive; the technique is only applicable when the rail is in tension, and less accurate on tight curves.</td>
<td>N/A ±6.5 MPa*</td>
<td>2.1 / [9], [19], [24]</td>
</tr>
<tr>
<td></td>
<td>Non-uniform static deformation over rail head / StereoDIC</td>
<td>Assumes deformation patterns of rail head over supports and a linear rail head curvature-temperature relationship; the occurrence of zero rail head curvature corresponds to the zero-stress state / reference measurements at known temperatures</td>
<td>Unfavorable illumination conditions in a field environment may affect its performance.</td>
<td>N/A N/A</td>
<td>2.6.1 / [89]</td>
</tr>
<tr>
<td><strong>Mechanical Vibration</strong></td>
<td>Vibrational resonance frequency of beams</td>
<td>Predicts axial force with E-B beam-based models modified for shear deformation and</td>
<td>The performance is affected by variations in support conditions; modeling the support</td>
<td>41 MPa** N/A</td>
<td>2.2.1 / [31]–[33]</td>
</tr>
<tr>
<td>Method</td>
<td>Description</td>
<td>Conditions</td>
<td>Notes</td>
<td>Reference(s)</td>
<td></td>
</tr>
<tr>
<td>------------------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------</td>
<td>--------------------------------------------------------------------------------------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Dynamic torsional rigidity/ D’stressen</td>
<td>Uses a hypothesized rail temperature-TB amplitude model, whose maximum occurs at the zero-stress state / reference measurement at known temperatures</td>
<td>N/A</td>
<td>The hypothesized rail temperature-TB amplitude relationship has not been verified; the performance is affected by the conditions of ties and fasteners. The sensitivity is limited when rail is subjected to compressive load.</td>
<td>2.2.2 / [37], [38]</td>
<td></td>
</tr>
<tr>
<td>Video-based rail vibrational response measurement</td>
<td>Not yet applied for stress determination in rail / Undefined</td>
<td>N/A</td>
<td>In the field test context, additional illumination and high-contrast boundaries may be required; CWR may not support the modes of interest.</td>
<td>2.6.2 / [90]</td>
<td></td>
</tr>
<tr>
<td>Electro-mechanical impedance</td>
<td>Not yet applied for stress determination in rail / model-based reference (analytical model results)</td>
<td>N/A</td>
<td>Uncertainties in material and geometric properties affect model performance.</td>
<td>2.6.3 / [98], [99]</td>
<td></td>
</tr>
<tr>
<td>Mechanical waves</td>
<td>Acoustoelasticity - ultrasonic bulk wave velocity</td>
<td>Assumes a stress-insensitive shear wave velocity and a longitudinal-shear wave velocity relationship / model-based reference (hypothesized constant bulk wave velocity)</td>
<td>Variations in residual stresses, microstructure, and temperature can affect the wave velocities. This approach, so far, is unable to reliably measure absolute stress.</td>
<td>N/A</td>
<td>±6.9 MPa***</td>
</tr>
<tr>
<td>-----------------</td>
<td>--------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------</td>
<td>-------</td>
<td>----------------</td>
</tr>
<tr>
<td>Acoustoelasticity - ultrasonic Rayleigh wave polarization</td>
<td>Not yet applied for stress determination in rails / Undefined</td>
<td>Inconsistent experimental results from field tests were reported, with no clear agreement between analytical model and experimental results.</td>
<td>N/A</td>
<td>N/A</td>
<td>2.3.1 / [47]–[49]</td>
</tr>
<tr>
<td>Acoustoelasticity - guided waves</td>
<td>Not yet applied for stress determination in rails / Undefined</td>
<td>Only numerical simulations on the effect of the acoustoelastic effect by axial load have been carried out so far.</td>
<td>N/A</td>
<td>N/A</td>
<td>2.3.2 / [30], [64], [65]</td>
</tr>
<tr>
<td>Guided waves - model-based method</td>
<td>Predicts the wavenumber of a given lateral flexural mode at the zero-stress state using a SAFE model; stress sensitivity of the wavenumber is determined based on E-B model / model-based reference</td>
<td>Performance depends on the accuracy of numerical model parameters (e.g., rail geometry and material properties)</td>
<td>±36 kips</td>
<td>Roughly 80 kips (in report [68])***</td>
<td>2.3.3 / [35], [36], [68], [69]</td>
</tr>
<tr>
<td>Method</td>
<td>Description</td>
<td>Performance</td>
<td>Literature</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nonlinear ultrasonic guided waves/Rail-NT</td>
<td>Assumes the zero-stress state occurs at the minimum of nonlinearity / reference measurement at known temperatures</td>
<td>The RNT can be determined only if the rail temperature passes the RNT during the monitoring period. Train traffic may affect the performance.</td>
<td>±1.1°C</td>
<td>±1.1°C (on concrete-tie track) &amp; ±2.8°C (on wooden-tie track)</td>
<td>2.3.4 / [72], [73]</td>
</tr>
<tr>
<td>Highly nonlinear solitary waves</td>
<td>Not yet applied for stress determination in rails / Undefined</td>
<td>No clear relation between the axial stress and the investigated features from a laboratory test on a short rail sample.</td>
<td>N/A</td>
<td>N/A</td>
<td>2.6.4 / [100]</td>
</tr>
<tr>
<td><strong>Magnetoelasticity</strong></td>
<td>Magneto-elastic methods/MAPS-SFT</td>
<td>Calibration is typically required for residual stress compensation; the performance depends on the accuracy of vertical-axial residual stress relationship, which varies among rail type, grades, and manufacturers.</td>
<td>N/A</td>
<td>±5.6°C</td>
<td>2.4 / [79], [80]</td>
</tr>
<tr>
<td>Magnetic Barkhausen noise/RailScan</td>
<td>Utilizes the MBN amplitude-stress curve established from laboratory test samples / reference measurement at the</td>
<td>MBN signature-stress curves need calibration for different rail types and manufacturers; the performance depends on the variations of residual stress and temperature.</td>
<td>N/A</td>
<td>0.26-9.8°C***</td>
<td>2.4 / [81], [82]</td>
</tr>
<tr>
<td><strong>Resonance fluorescence piezospectroscopy</strong></td>
<td>Photoluminescence</td>
<td>Assumes a linear relationship between the wavenumber of the fingerprint peaks in fluorescence spectrum and absolute axial stress; assumes the condition of pulverized thermite weld samples can represent that of a weld joint at zero-stress state in practice / reference measurement at the zero-stress state (on surrogate samples)</td>
<td>Residual stress compensation is necessary; the performance is affected by uncertainties associated with stress concentrations and piezospectroscopy coefficient.</td>
<td>N/A</td>
<td>$\geq \pm 25$ MPa</td>
</tr>
</tbody>
</table>
6 References


D. Fletcher, M. Kraeling, and R. Patton, “Research and development of a new method for wirelessly interrogating the stress free temperature of continuously welded rail,” in *AusRAIL PLUS 2017, Rail’s Digital Revolution*, Nov. 2017. [Online]. Available: http://railknowledgebank.com/Presto/content/Detail.aspx?ctlID=MTk4MTRjNDUtNWQ0My00OTTmLTIyYWUtZWFjM2U2OTE0ZDY3&rID=NTMzMQ==&sID=MzU=&ph=VHJ1ZQ==&qcf=&rrtc=VHJ1ZQ==&bckToL=VHJ1ZQ==


Declaration of interests

☐ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☒ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

John Popovics reports financial support was provided by Federal Railroad Administration. John Popovics reports financial support was provided by National Academy of Sciences.