A standardizable approach for Tooth Flank Fracture

Stephan André Böhme · Joni Keski-Rahkonen · Tami Komssi

Received: 16 March 2022 / Accepted: 30 August 2022 / Published online: 27 September 2022 © The Author(s) 2022, corrected publication 2022

Abstract

Ultrasonic scanning of large spiral bevel gears and the subsequent fracture analyses of early-stage tooth flank fractures imply a mode II crack initiation under surface-parallel, orthogonal shear stresses. An iteration of the Dang Van criterion is presented that focuses exclusively on surface-parallel material planes, yielding a standardizable approach for the study of tooth flank fracture in spiral bevel gears. It is shown that Dang Van’s shear stress amplitude and maximum hydrostatic stress along a surface-perpendicular path can adequately be described by the Hertzian orthogonal shear stress amplitude and Lang’s compressive residual stress. The fatigue properties along the same path are estimated from Thomas’ hardness model and augmented by ISO-inspired lifetime, size and conversion factors. The article closes with a comparison between the commonly used Witzig-Boiadjiev, Hertter-Wirth, DNV subsurface fatigue and the herein proposed, rationalized Dang Van criterion.

Keywords Tooth flank fracture · Bevel gears · Rolling contact fatigue · Case hardened steel · Standardizable approach

Nomenclature

| Symbol | Description |
|--------|-------------|
| CHD | Case hardening depth |
| \(d_e\) | Outer pitch diameter |
| \(E\) | Young’s modulus |
| \(f_{1,1}\) | Uniaxial and shear fatigue strength |
| \(f_{xK}\) | Conversion factor |
| \(HV_{s,c}\) | Surface and core hardness |
| \(HV(x)\) | Local Vickers hardness |
| \(J_{2,a}\) | Second invariant of deviatoric stress |
| \(k_{hap}, k_{hfp}\) | Addendum and dedendum factors |
| \(K_{1,2}\) | Adjustment factors |
| \(K_{NT}\) | TFF-specific lifetime factor |
| \(K_X\) | TFF-specific statistical size factor |
| \(MB\) | Gear’s design point |
| \(m_{nn}\) | Mean normal module |
| \(n\) | RPM |
| \(N_f\) | Number of load cycles |
| \(P\) | Power |
| \(s_{rel}\) | Relative standard deviation |
| \(T450, T550\) | 450 and 550 HV hardness depths |
| \(x_{hm}\) | Profile shift |
| \(x_{HV,max}\) | Depth of the maximum hardness |
| \(Y_{NT}, Z_{NT}\) | Tooth root and pitting lifetime factors |
| \(Y_X\) | Tooth root-specific size factor |
| \(z\) | Number of teeth |
| \(Z_K\) | Bevel gear factor |
| \(Z_{MB}\) | Midzone factor |
| \(Z_S\) | Slip factor |

Stephan André Böhme
stephan.a.bohme@ntnu.no

1 Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), 7491 Trondheim, Norway
2 Kongsberg Maritime Finland Oy, FI-26100 Rauma, Finland
3 ATA Gears Oy, FI-33101 Tampere, Finland
The herein presented work is part of a wider research program, aimed at the prediction and prevention of subsurface-initiated fatigue failures in large maritime bevel gears. It continues the efforts of the ‘Improved reliability of thrusters’ joint industry project—a cooperation between thruster suppliers, gear manufacturers, steel suppliers, forging companies and the maritime classification society Det Norske Veritas (DNV). The article combines the previously published findings on the material properties of carburized CrNiMo steels [1], the stress and fatigue predictions in the mean cross-section of spiral bevel gear teeth [3], the test results from back-to-back gear testing [4] and the observations from early-stage tooth flank fracture on large maritime bevel gears [2, 5] to a standardizable approach for the prediction and ultimately prevention of tooth flank fracture (TFF).

The failure mode TFF is characterized by crack initiation in the interface between the hard case and soft core of case hardened, but also nitrided or induction hardened gear teeth [20]. Cracks initiate typically in mode II (in-plane shear) underneath the load-carrying tooth flank and propagate through the gear tooth in mode I (opening) due to the acting bending stresses in a 40 to 50° angle to the loaded flank. The primary crack front can be accompanied by secondary and tertiary cracks. When final rupture occurs, a large part of the weakened gear tooth breaks off, leaving traces of macroscopic fatigue features on the fracture surfaces (see Fig. 3). Those features are partially removed due to the relative motion of the two fracture surfaces and the resulting plastic deformation.

TFF is caused by a non-proportional, multiaxial stress state in a material depth characterized by locally changing constitutive parameters. The acting stress tensor is complex, consisting of dynamic Hertzian, out-of-phase bending, transverse shear, transverse normal and frictional shear stresses, and imposed by static residual stresses [5]. With failures occurring in the high to very high cycle fatigue regime, stress-based multiaxial fatigue criteria can suitably be applied. The basis for the herein proposed TFF approach forms the Dang Van (DV) criterion due to its widespread use for the analysis of rolling contact fatigue [10, 13, 23] and its ability to yield a standardizable approach. The DV criterion assumes local plasticity in the mesoscale, even if the material is only loaded elastically in the macroscale. Governing the conversion between the meso- and macroscale is a time-invariant residual stress according to Melan’s shake-down theorem [12]. The DV’s statistically most accurate definition is outlined in Eq. 1 [25], relying on the shear stress amplitude \( \tau_a \) as a function of the Euler angles \( \theta, \phi \) and the maximum hydrostatic stress \( \sigma_{hyd,max} \). It weighs \( \tau_a \) and \( \sigma_{hyd,max} \) through the model constants \( a_{DV} \) and \( b_{DV} \) and compares the maximum equivalent stress against the uniaxial fatigue strength under alternating load \( f_{-1} \).

\[
A_{DV} = \max_{\theta,\phi} \left( a_{DV} \tau_a(\theta, \phi) + b_{DV} \sigma_{hyd,max} \right) / f_{-1}
\]

\[
a_{DV} = f_{-1} / L_1 = \kappa \quad b_{DV} = 3 - 3/2 f_{-1} / L_1 = 3 - 3/2 \kappa
\]

With TFF occurring in carburized gear teeth in a depth of 1 to 2 times the case hardening depth (CHD), the stresses, fatigue properties and model constants across this specific hardness range need to be quantified. In the case of the quoted DV criterion, \( a_{DV} \) and \( b_{DV} \) are derived from the material constant \( f_{-1} \) and \( L_1 \) (i.e. the uniaxial and shear fatigue strengths under alternating load). The ratio \( f_{-1} / L_1 \) is commonly referred to as the fatigue ratio \( \kappa \). For case hardened gear teeth, \( \kappa \) ensures a brittle material behavior in the hard case and a ductile behavior in the tooth interior or core. As \( \kappa \) increases from case to core, so do \( a_{DV} \) and the effect of the shear stress amplitude \( \tau_a \) on \( A_{DV} \). Contrary, \( b_{DV} \) and the effect of \( \sigma_{hyd,max} \) on \( A_{DV} \) decrease.

Throughout this manuscript, a rationalized version of the DV criterion will be compared against other TFF criteria, specifically the Witzig criterion [34]. Witzig proposed his TFF-criterion in the FVA556 I project, which has since been adapted in the ISO/TS 6336-4 for the prediction of TFF in cylindrical gears [20]. The subsequent FVA556 II and FVA556 III projects verified its applicability to large cylindrical and spiral bevel gears [9, 32]. As the bevel gear-focused FVA556 III project was led by Boiadjiev, the cri-
The WI-BO criterion does not rely on model constants to weigh its stresses and differs thereby from conventional multiaxial fatigue criteria like DV, Findley, Liu & Zenner or Crossland [11, 12, 15, 35]. While these criteria combine the planar shear and normal stresses $\tau_{a}, \tau_{m}, \sigma_{n_{a}}, \sigma_{n_{m}}$, the second invariant of the deviatoric stress tensor $J_{2,a}$ or the hydrostatic stress $\sigma_{h_{yd}}$ to an equivalent stress, the WI-BO criterion utilizes gear-specific stresses. Both, the lack of model constants and the use of gear-specific stresses make it impossible to assess the WI-BO criterion’s accuracy against other stress-based multiaxial fatigue criteria. The accuracy of the herein quoted DV criterion was quantified in a study by Papuga et al. [26] against 284 multiaxial fatigue tests under proportional and non-proportional stresses. It achieved an average fatigue error index ($A_{DV} - 1$) of -2.2% and a standard deviation of the fatigue error index of 15.9%.

The WI-BO criterion has been defined for a specific range of input parameters. It will be shown that the 5 mm lower limit for the equivalent radius of curvature $\rho_{eq}$ is non-conservative and leads to inaccurate results for one of the studied gears. The herein outlined, rationalized version of the DV criterion allows for a simple stress prediction and works for all studied input parameters.

### 2 Material and Methods

Similar to the WI-BO criterion, the herein outlined TFF criterion can either be applied to the gear’s

- design point $MB$ with loads defined according to the ISO10300-2 [18]
- mean cross-section with loads from either the FVA516, FVA519, or Becal [7, 17, 21]
- tooth flank with loads from a loaded tooth contact analysis (LTCA) with Becal [7]

The analytical equations outlined in this manuscript study the material strengths and stresses underneath the gear’s design point $MB$, but can similarly be applied to the entire tooth flank. Assuming constant hardness and residual stress profiles across the tooth, only knowledge of the local Hertzian contact stress $\sigma_{H,i}$ and the local, equivalent radius of curvature $\rho_{eq,i}$ is required. Figure 1 outlines the model in its entirety.

Of the calculations that were run on maritime gears [2, 3, 5], FVA gears [33] and medium-sized test gears [4], three cases are highlighted here. With them, the rationalized DV criterion is compared against the WI-BO [9, 34], the Hertter-Wirth (HE-WI) [16, 33] and the DNV subsurface fatigue criterion [14]. Table 1 specifies the macrogeometries of the
two studied maritime bevel gears (G1, G2) with a history of TFFs and of the medium-sized test gear (B2-1) with 6 documented TFFs on 10 performed tests [4]. Listed are the gears’ pressure angle $\alpha$, the number of teeth $z$, the outer pitch diameter $d_e$, the tooth width $b$, the mean spiral angle $\beta_m$, the profile shift $x_{hm}$ and the addendum/dedendum factors $k_{hap}/k_{hf}$. The produced hardness profile is described through the case hardening depth $CHD$, the surface/core hardness $HV_s/HV_c$ and the depth of the hardness peak $x_{HV, max}$. Its use to describe not the depth of the hardness peak but the hardness gradient in the $CHD$ will be outlined later in the manuscript. The listed load parameters are the power $P$, the RPM $n$ and the cycles to failure $N_f$. All gears were manufactured from case hardened 18CrNiMo7-6 steel.

Gear set G1 will be used to visualize and verify the proposed analytical equations against an accurate, 2D plane strain, numerical stress prediction in the gear’s mean cross-section [3]. Figure 2 plots the LTCA-derived contact stresses $\sigma_{H,i}$ and radii of curvature $\rho_{eq,i}$ across a G1 pinion tooth and its mean cross-section.

### 2.1 Stress model

This chapter details the rationalization of the DV criterion when predicting TFF in spiral bevel gears. In its adapted form, it requires the definition of the orthogonal shear stress amplitude $\tau_{xy,a}$ for Hertzian line contacts, the residual stress $\sigma_{res}$, the model constants $a$ and $b$, which rely on

![Fig. 2](image)

**Fig. 2** Stress model for G1 pinion tooth and mean cross-section [5]. a) LCTA pressure distribution $\sigma_{H,i}$. b) LCTA radius of curvature $\rho_{eq,i}$. c) Mean cross-section
the local fatigue ratio $\kappa$ and the definition of the uniaxial fatigue strength under alternating load $f_{-1}$. The derivation of $f_{-1,K}$ for case hardened bevel gears and the considered lifetime, size and conversion factors are outlined in the next chapter.

$$A(x) = \frac{a(x) \tau_{xy,a}(x) + 2/3 b(x) \sigma_{res}(x)}{f_{-1,K}(x)}$$

$$a(x) = \kappa(x) \quad b(x) = 3 - 3/2\kappa(x)$$ \hspace{1cm} (3)

The original DV criterion maximizes the sum of the shear stress amplitude $\tau_{a}(\theta, \phi)$ and the maximum hydrostatic stress $\sigma_{hyd,max}$ across all material planes. Whereas $\sigma_{hyd,max}$ remains constant under coordinate transformation, $\tau_{a}(\theta, \phi)$ needs to be derived for each material plane. The Minimum Circumscribing Circle (MCM) or the Maximum Rectangular Hull methods (MRH) are typically employed for that purpose [6, 8, 12]. Due to their computational complexity, the derivation of $\tau_{a}(\theta, \phi)$ from the MCM or MRH is not suitable for a standardizable TFF criterion.

Gear teeth inspections by means of phased array ultrasonic scanning during service intervals have become a standard procedure for many suppliers of azimuthing thrusters. If crack-like indications are detected, the gears are oftentimes scraped to prevent subsequent TFFs. Failure investigations of the affected gear teeth reveal predominantly surface-parallel cracks of elliptical shape underneath the load-carrying flank. Over a 3-year period, 174 and 32 crack-like indications were detected on 16 G1 and 6 G2 gear sets respectively [3]. Two examples of G1 gear teeth are given in Fig. 3, showing two early-stage TFF cracks 0.45 and 0.65 mm in size in a respective depth of 2.6 and 3.3 mm, orientated 9 and 14° to the load-carrying flank. A terminal stage TFF is similarly plotted in Fig. 3, highlighting the macroscopic fatigue features found on TFF fracture surfaces.

These observations suggest that the DV criterion’s iteration over all material planes could be foregone in favor of the exclusive study of surface-parallel material planes. To verify this observation, the numerically-derived, true maximum of the shear stress amplitude $\tau_{MRH,max}$ is compared against the analytically-predicted orthogonal shear stress amplitude $\tau_{xy,a}$. Under Hertzian line contact, $\tau_{xy,a}$ can readily be approximated according to Eq. 4 [14]. The required inputs are the ISO10300-2’s Hertzian contact pressure in the gear’s design point $\sigma_{H}$ [18], the bevel gear factor $Z_{K}$ [18], the mid-zone factor $Z_{MB}$, the tooth depth $x$ and the half Hertzian contact width $b_{H}$ that depends in turn on the equivalent radius of curvature $\rho_{eq}$, the Young’s modulus $E$ and the Poisson ratio $\nu$. $Z_{K}$ and $Z_{MB}$ are to be removed from Eqs. 4 and 5 if LTCA inputs are considered. The WI-BO’s definition of the local equivalent stress without consideration of residual stresses $\tau_{eff,L}(x)$ is outlined in Eq. 6.

$$\tau_{xy,a}(x) = 0.25 \sigma_{H,K} \cos \left( \frac{x}{b_{H}} - 0.5 \frac{\pi}{2} \right)$$

with $\sigma_{H,K} = \frac{\sigma_{H}}{Z_{K}}$ \hspace{1cm} (4)
do therefore not affect teeth. These Hertzian stresses are compressive in nature and a result of the Hertzian contact between the meshing gear. The dynamic stresses inside a gear tooth are by and large according to Lang [22], relying on the local hardness $HV(x)$ and the core hardness $HV_c$. The local hardness $HV(x)$ is based on the in Table 1 specified surface and core hardenesses, the case hardening depth and Thomas’ hardness model [31] that is detailed in the next chapter.

$$\sigma_{res}(x) = \begin{cases} -5/4(HV(x) - HV_c) & \text{for } HV(x) - HV_c \leq 300 \\ 2/7(HV(x) - HV_c) - 460 & \text{else} \end{cases}$$

Figure 4 plots $\tau_{MRH,max}$ for the mean cross-section of G1 and compares it along the surface-perpendicular path P1 starting in $MB$ to $\tau_{xy,a}$ and $\tau_{eff,L}$. The surface stresses in $MB$ between the numerical and the analytical calculations were matched through the ISO’s mounting factor $K_{MB}$. The window in Fig. 4b specifies the surface parameters in $MB$. All three stresses track each other closely along P1 and deviate only in the surface-near region, where the numerical model considers the occurring frictional shear stresses. As Fig. 4 shows, the difference between $\tau_{MRH,max}$ and $\tau_{xy,a}$ is in the TFF-critical material depth small and typically less than 3%.

The second stress component that is considered in the DV criterion is the maximum hydrostatic stress $\sigma_{hyd,max}$. The dynamic stresses inside a gear tooth are by and large a result of the Hertzian contact between the meshing gear teeth. These Hertzian stresses are compressive in nature and do therefore not affect $\sigma_{hyd,max}$. Disregarding all other dynamic stresses inside a gear tooth, Eq. 7 outlines a first approximation of $\sigma_{hyd,max}$. It relies solely on the static residual stress $\sigma_{res}$ and assumes equal tangential and longitudinal residual stresses and no normal residual stresses.

$$\max_i(\sigma_{hyd}(x,t)) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \sigma_{res}(x) & 0 \\ 0 & 0 & \sigma_{res}(x) \end{bmatrix} = 2/3\sigma_{res}(x)$$

In line with the ISO/TS 6336-4 [20] and WI-BO criterion [9, 34], only the compressive residual stresses along P1 are considered. They are estimated as for the WI-BO criterion according to Lang [22], relying on the local hardness $HV(x)$ and the core hardness $HV_c$. The local hardness $HV(x)$ is based on the in Table 1 specified surface and core hardenesses, the case hardening depth and Thomas’ hardness model [31] that is detailed in the next chapter.

$$\sigma_{res}(x) = \begin{cases} -5/4(HV(x) - HV_c) & \text{for } HV(x) - HV_c \leq 300 \\ 2/7(HV(x) - HV_c) - 460 & \text{else} \end{cases}$$

Figure 5 plots the stress difference between the numerically calculated, maximum hydrostatic stress and the hydrostatic stress according to Eq. 7 for the mean cross-section and path P1 of G1.

The observable linear stress difference along P1 proves that the Hertzian stresses have no effect on $\sigma_{hyd,max}$ but highlights the missing bending stress component in Eq. 7. They can adequately be approximated by adapting the ISO10300-3 [19] to the gear’s design point $MB$ as outlined in [5]. Here, the bending stresses are disregarded in favor of a simpler definition of the DV criterion that can directly be applied to all LTCA’s contact points, without detailed
knowledge of the local bending force and tooth width. Across the studied gear sets, the omission of the bending stress reduces the DV criterion’s equivalent stress in the TFF-critical material depth by 10 to 15%. This discrepancy was accounted for through a reduction of the uniaxial fatigue strength \( f_{-1} \).

### 2.2 Strength model

The measurements conducted in the context of this research support the approximation of a bevel gear’s hardness profile according to Thomas [31]. Thomas describes the hardness profile as two second-order polynomials up to and beyond the \( CHD \). In contrast to the often-quoted Lang model [22], it can represent a gear size-dependent hardness gradient through the use of its parameter \( x_{HV,max} \).

The inputs to Thomas’ hardness model are the surface hardness \( HV_s \), the core hardness \( HV_c \), the case hardening depth \( CHD \) and the depth of the maximum hardness \( x_{HV,max} \).

\[
egin{align*}
HV(HV > HV_s) &= HV_s \quad \text{hardness limit} \\
HV(x > x_c) &= HV_c \quad \text{core hardness} \\
HV(x \leq CHD) &= a_a x^2 + b_a x + c_a \\
HV(x_c \geq x > CHD) &= a_b x^2 + b_b x + c_b
\end{align*}
\]  

\[
a_a = \frac{550 - HV_s - CHD^2 - 2x_{HV,max}CHD}{HV_dC}
\]

\[
a_b = \frac{HV_dC}{2(CHD - x_c)}
\]

\[
b_a = -2a_a x_{HV,max}
\]

\[
b_b = -2a_b x_c
\]

\[
c_a = HV_s
\]

\[
c_b = 550 - a_b CHD^2 - b_b CHD
\]

\[
HV_{dC} = 2a_a CHD + b_b
\]

\[
a_1 = -HV_{dC}
\]

\[
b_1 = 2CHDHV_{dC} + 2(HV_c - 550)
\]

\[
c_1 = -CHD^2HV_{dC} - 2CHD(HV_c - 550)
\]

\[
x_c = \frac{-b_1 + (b_1^2 - 4a_1c_1)^{0.5}}{2a_1} \quad \text{core hardness depth}
\]

Rather than describing the depth of the maximum hardness, \( x_{HV,max} \) is used here to specify the hardness gradient expressed through the hardness depth ratio \( T450/T550 \) (i.e. the depths where \( HV(x) \) equals 450 and 550HV).

Extensive hardness measurements were carried out during the production of module 5 to 35 bevel gears made from 18CrNiMo7-6 steels of high, medium and low hardenability that were quenched at different severities. Figure 6a plots the 10, 50 and 90 percentile curves for the obtained \( T450/T550 \) ratios, suggesting a significantly smoother hardness transition for small than large gears. On average, a module 35 gear will transition 30% earlier from its 550HV to its 450HV hardness than a module 5 gear, significantly weakening the TFF critical material depth due
to a reduction in local fatigue strength and compressive residual stress. An equation for \( x_{HV,max} \) for the average \( T450/T550 \) ratio is given in Fig. 6a and applied in Fig. 6b to a module 5 and 35 gear.

\( T450/T550 \) or \( x_{HV,max} \) can be understood as a technological size factor, which should only be considered if no detailed knowledge of the manufactured hardness profile exists. It outlines the necessity of a higher relative CHD on larger gears due to a steeper hardness gradient. The increasing hardness gradient can be attributed to the, for large gears typical, carburizing at multiple carbon potentials, the amount of martensitic transformation (i.e. the cooling rate) and the gears’ autotempering response.

From the hardness profile, the necessary uniaxial fatigue strength under alternating load \( f_{-1} \) and fatigue ratio \( \kappa \) are estimated. Whereas \( f_{-1} \) was calibrated to match the material utilizations of the herein proposed TFF criterion against the BO criterion detailed in [3–5], \( \kappa \) relies on the fatigue tests performed on 18CrNiMo7-6, 34CrNiMo6 and 18NiCrMo14-6 steels at multiple forging ratios and material cleanlinesses [1]. The steels were tested under uniaxial stresses at \( R = -1 \) and \( R = 0 \) in the high and very high cycle fatigue regime and under shear at \( R = -1 \) in the high cycle fatigue regime. The results suggest that within a 350 to 700 HV hardness range, \( \kappa \) transitions from 1.56 to 1.38, capturing the change from a ductile material behavior in the tooth core to a more brittle behavior in the case.

Based on the performed tests and the weakest link theory [29], a TFF-specific, statistical size factor \( K_X \) was derived. It relies on the obtained relative standard deviation \( (s_{rel} = 4\%) \) and compares the highly stressed volumes of the fatigue specimens and gear teeth of various sizes. The obtained logarithmic relationship between the highly stressed volume and the size factor can adequately be simplified to a linear relationship between \( K_X \) and the gear’s mean normal module \( m_{mn} \) [5]. The effect of the applied load on the size factor is small in comparison to the size of the gear tooth, deeming the \( K_X(m_{mn}) \) simplification adequate. \( K_X \) is in magnitude equivalent to the ISO10300-3’s tooth root-specific size factor \( Y_X \) [19], differing only in the chosen lower boundary. \( s_{rel} \) was also used to define the conversion factor \( f_{xK} \). It translates the mean or average fatigue strength to the ISO-typical 1% fatigue strength and corresponds well to the tooth root and pitting specific conversion factors [30]. \( f_{xK} \) should be used if ISO-comparable failure probabilities are targeted on gears used in industrial applications and excluded to identify the necessary loads to promote TFF on test gears. For the study of the outlined three gear sets, \( f_{xK} \) is initially disregarded to plot the material utilization for the average fatigue strength. Furthermore, a lifetime factor was derived from the TFF-specific gear tests performed on \( m_{mn} = 9.45 \) mm, case-hardened gears loaded at 750, 875 and 1000 kW with failures and run-outs in the range from \( 10^6 \) to \( 10^8 \) load cycles [4]. By relying on the maximum likelihood method [27], a continuous rather than multi-staged SN-curve or lifetime factor was obtained.

\[
\begin{align*}
\kappa(x) &= \frac{\ln(N_f)}{-2.73} + 0.89 \quad \text{with} \quad K_{NT} \leq 1.6 \\
f_{xK} &= 0.91
\end{align*}
\]

\[
\begin{align*}
f_{-1}(x) &= 0.7HV(x) \\
\kappa(x) &= \sqrt{3} - 5 \times 10^{-4}HV(x) \\
f_{-1,K}(x) &= f_{-1}(x)K_XK_NTf_{xK}
\end{align*}
\]
Figure 7 plots the fatigue strength $f_{-1, K}$ along path P1 for G1 alongside the model constants $a$ and $b$. The increase of $a$ and decrease of $b$ during the case/core transition visualize the changing effect of the shear and hydrostatic stresses on the local material utilization throughout the hardened layer. Plotted is also the derived lifetime factor $K_{NT}$ in comparison to the ISO10300’s pitting and tooth root breakage lifetime factors $Z_{NT}$ and $Y_{NT}$ for case hardened steel [18, 19].

2.3 Case study

With the equivalent stress and fatigue strength established, the local material utilization $A(x)$ can be calculated. It will be shown that the study of the stresses and strengths along P1 is adequate for industrial applications and gear teeth with a conventional tooth profile. For test gears and gears with an adapted microgeometry, the gear’s mean cross-section or entire tooth flank need to be analyzed. To better highlight the differences between the studied criteria, a second surface-perpendicular path P2 is defined that intersects the predicted maximum material utilization for each criterion. Gear sets G1, G2 and B2-1 are studied according to the modified DV, the DNV subsurface fatigue [14], the Hertter-Wirth [16, 33] and Witzig-Boiadjiev criteria [9, 28, 34]. Hertter modified the Liu & Zenner shear stress intensity criterion [35] for the study of pitting and tooth root breakage in cylindrical gears. Wirth modified in turn the Hertter criterion through the implementation of tensile residual stresses in the tooth interior and the slip factor $Z_S$ to enable the study of TFF [33]. The Witzig criterion was developed out of the Oster criterion [24] specifically for the study of TFF in cylindrical gears and verified for its applicability to spiral bevel gears by Boiadjiev in the FVA follow-up project FVA556 III [9]. A standardizable approach of the WI-BO criterion for the study of TFF in the mean cross-section of spiral bevel gears is outlined in FVA556 III and the works by Pellkofer et. al [9, 28]. Figures 8, 9 and 10 plot the respective LTCA-derived material utilizations of gear sets G1, G2 and B2-1 without the consideration $f_{-1, K}$ in case of the DV criterion and compare the utilizations along path P2 with one another. The specific P2 subplots state the relevant Hertzian pressures and curvatures in the surface nodes or P2 starting points.

The risk for TFF in G1 is adequately captured by the WI-BO and the rationalized DV criteria as large utilizations are predicted in the TFF-critical material depth of 1 to 2 times $CHD$. The vicinity of P1 and P2 suggests that the study $A(x)$ underneath $MB$ is sufficient for gear set G1. The HE-WI and DV criteria predict very comparable surface utilizations but deviate significantly in the TFF-critical depth. The late transition from compressive to tensile residual stresses according to Wirth’s residual stress model [33] means that the tensile stresses do not affect the maximum utilization of 0.63 [5]. The DNV criterion lies between the HE-WI, WI-BO and modified DV criteria in the TFF-critical material depth but predicts relatively large surface-near utilizations for the studied moderate contact stresses.

Figure 9 supports the observations and statements made for the different criteria in Fig. 8. Comparing gear sets G1 and G2, the DV’s utilization maxima of 1.08 and 0.95 suggest a considerable risk for subsurface fatigue as the average and not 1% fatigue strength was considered.

For gear set B2-1 and the HE-WI criterion, the large $CHD$ in combination with the narrower tooth width result in large tensile residual stresses in the tooth core. They were modeled according to the equations suggested by Wirth and capped at 500 MPa to prevent the HE-WI criterion from
Fig. 8 Material utilization comparison, G1 pinion. a Hertter-Wirth criterion [16, 33]. b WI-BO criterion [9, 34]. c DNV criterion [14]. d Mod. Dang Van criterion. e Comparison along path P2. f Wheel initiated TFF on G1
Fig. 9 Material utilization comparison, G2 wheel. a Hertter-Wirth criterion[16, 33]. b WI-BO criterion[9, 34]. c DNV criterion [14]. d Mod. Dang Van criterion. e Comparison along path P2. f Wheel initiated TFF on G2
Fig. 10  Material utilization comparison, B2-1 wheel [4]. a Hertter-Wirth criterion [16, 33]. b WI-BO criterion [9, 34]. c DNV criterion [14]. d Mod. Dang Van criterion. e Comparison along path P2. f Wheel initiated TFF on B2-1
producing even larger utilizations. When compared against the herein proposed TFF criterion, not the material depth with the largest shear stresses is deemed critical, but the gear tooth’s neutral axis, where the highest tensile residual stresses occur. The failure to predict appropriate utilizations for G1 and G2 and the overestimation of the tensile residual stresses in B2-1 do not support the application of the HE-WI criterion for the prediction of TFF in spiral bevel gears. The DNV criterion fails to reproduce the observed TFF on gear set B2-1, predicting instead a surface failure. Both, the DNV and the herein proposed TFF criterion rely on the same orthogonal shear stress but differ in the consideration of the compressive residual stresses. The cross-sectional plot of the modified DV criterion highlights that the study of the gear’s design point \( MB \) or path P1 no longer yields the largest material utilization.

Apparent is furthermore a large utilization peak in the tooth tip of the analyzed B2-1 wheel tooth according to the WI-BO criterion. The utilization peak can not be attributed to large tensile residual stresses as they are not considered. It stems instead from the WI-BO’s quasi shear mean stress \( \Delta \tau_{eff,L,RS} \) and specifically the adjustment factor \( K_2 \). Both parameters are outlined in Eqs. 14 and 15 [9, 34].

\[
\Delta \tau_{eff,L,RS}(x) = K_1(x) \frac{\sigma_{res}(x)}{100} - 32 \tanh(9x^{1.1}) - K_2(x)
\]

\[
K_2(x) = (-\tanh(0.1(\rho_{eq} - 10)) + 1 \times \left( \frac{CHD^2}{16} \times \frac{\max(|\sigma_{res}(x)|)}{\tanh \left( -2 \frac{CHD^2}{1000} - 200 \right)} + \frac{\max(|\sigma_{res}(x)|)}{10} \right)
\]  

(14)  

(15)

Figure 10e details also the Hertzian pressure and radius of curvature in the surface nodes or P2 starting points. For the WI-BO criterion, those parameters are 1680MPa and 10.44mm. When considered alongside the parameters listed in Table 1, the stress components plotted in Fig. 11a can be calculated. For the studied B2-1 tooth, path P2 and a 5mm material depth, \( \tau_{eff,L} \) has fallen to 38MPa, \( \tau_{res} \) to 0MPa, whilst \( K_2 \) and thereby \( -\Delta \tau_{eff,L,RS} \) have grown to 138MPa, prompting the elevated material utilization in the specific part of the gear tooth. Specifically the term \( -\tanh(0.1(\rho_{eq} - 10)) + 1 \) in Eq. 15 causes the sudden rise in \( K_2 \) along the tooth profile. Figure 11b plots the material utilization for the meshing pinion tooth, yielding an even higher utilization, not due to a change in stresses but an increase in tooth width in the pinion dedendum. \( K_2 \) increases linearly with \( x \). The WI-BO criterion is defined for a specific input parameter range, namely \( \sigma_H = 500 \) to 3000MPa, \( \rho_{eq} = 5 \) to 150mm and \( CHD = 0.3 \) to 4.5mm [9, 34]. The results suggest that the lower limit for \( \rho_{eq} \) should be reevaluated or that the analysis of off-centered contact points should be excluded from the WI-BO criterion. The modified DV criterion yields an adequate utilization in the correct tooth height and depth.

3 Discussion

A standardizable approach for the prediction of TFF in spiral bevel gears has been presented. It relies on the well-established DV criterion and assumes TFF crack initiation parallel to the load-carrying flank. The DV criterion’s iteration over all material planes is thereby gone to enable a standardizable TFF prediction. The fractographic analysis of early-stage TFFs on large maritime bevel gears and the shear stress comparison between \( \tau_{MRH,max} \) and \( \tau_{xy,a} \) substantiate the rationalization. In the relevant material depth, the stress difference between the orthogonal shear stress amplitude and the true maximum is less than 3%. TFF cracks initiate approximately parallel to the load-carrying flank under the acting orthogonal shear stresses and alter their trajectory once the crack has grown to a specific size, adopting the typically observed 45° angle. As shown in the manuscript, the DV criterion’s maximum hydrostatic stress can sufficiently be approximated as a function of Lang’s residual stress [22].

Two of the three presented gear sets suggest a close match in material utilization but not failure probability between the WI-BO [9, 34] and the simplified DV criterion. Gear set B2-1 outlines an inaccuracy in the WI-BO criterion, in particular \( K_2 \) and the lower limit of the \( \rho_{eq} \) range. While the plotted results for the DV criterion study a path at approximately half tooth height, the WI-BO criterion predicts the maximum utilization along a path towards the tooth tip on the wheel and the tooth root on the pinion, where \( \rho_{eq} \) is small and \( K_2 \) exceedingly large. The following section is used to further highlight the differences between the WI-BO and DV criteria.

The presented results document comparable stress levels for the shear stress intensity derived, analytical equations of the WI-BO criterion and the herein promoted orthogonal shear stress amplitude. The integration of the shear stress across all material planes and the calculation of a single shear stress amplitude in surface-parallel material planes yield similar results. But without the need for a stress integration, the standardized DV criterion relies on two comprehensible stress equations and avoids the need for an input parameter range.

The herein outlined material model is based on numerous fatigue tests [1] and details the correlation between a hardness decrease and a ductility increase from case to core.
(see Eq. 12 for $\kappa$). The DV criterion considers that transition by increasing the impact of the acting shear stresses on $A$ in the tooth interior and the effect of the compressive hydrostatic stresses in the hard case. As a result, it suggests the application of high hardenability steels that yield a higher core hardness after quenching and tempering to prevent TFF. When recalculating the maximum material utilization $A$ for gear set G1, a core hardness increase from 393 to 450HV decreases $A$ from 1.06 to 1.05. For the same case, the WI-BO criterion predicts an increase in material utilization from 1.08 to 1.11, despite the criterion’s linear correlation between hardness and shear strength. The increase in $A_{FF}$ can be attributed to the definition of Lang’s residual stress model [22], which sets the compressive residual stresses to 0MPa for $x > x_c$ and estimates the local residual stress from the hardness difference ($HV(x) - HV_c$). Despite relying on the same residual stress model, the DV criterion correctly predicts a reduction in material utilization for an increase in core hardness.

According to the FVA556 I and III projects [9, 34], the WI-BO criterion predicts a 50% failure probability for a material utilization of $A_{FF} = 0.8$ under constant load. What utilization is suitable for an ISO-typical 1% failure probability is unclear. The conversion factor $f_xK$ promoted in this manuscript allows for the TFF-free design of industrial gears and the necessary load adjustment to reliably produce TFF on test gears. Also, the technological size factor, represented through the hardness depth ratio $T_{450}/T_{550}$, the statistical size factor $K_X$ and lifetime factor $K_{NT}$ help quantify the differences between small, highly-loaded test gears and large industrial gears subjected to moderate loads. The steep hardness gradient observed on large gears, the large, highly stressed volume, the lower core hardness and the load cycle accumulation over multiple years of opera-

---

**Fig. 11** Detailed study of adjustment factor $K_2$[9, 34]. **a** WI-BO equivalent stresses along P2. **b** WI-BO utilization for B2-1 pinion

---

**Fig. 12** Reanalysis of G1 according to DV criterion for $P = 1760$kW and $f_xK$ based on LTCA [7]. **a** Pressure distribution $\sigma_{H,j}$. **b** Radius of curvature $\rho_{eq,i}$. **c** Material utilization $A_i$
tion elevate the TFF likelihood on maritime gears. Gear set G1 is reanalyzed at a more operationally-typical load of P = 1760 kW or 80% of the nominal load. The modified DV criterion considers now the conversion factor $f_{xK}$ and plots the material utilization for a 1% failure probability, being much closer to the practically observed failure frequencies than the initially plotted 50% probability.

The DV criterion yields its 1.06 utilization peak in the TFF-typical depth of 4.4 mm, indicating a substantial risk for TFF even under the considered moderate loads. The high utilization corresponds well with the large number of observed crack-like indications from ultrasonic scanning on G1 [3]. To demonstrate the criterion’s applicability to the study of all LTCA contact points between meshing pinion and wheel teeth, the material utilization of a G1 pinion tooth is plotted in Fig. 12.

4 Conclusions

The article draws on the results of the ‘Improved reliability of thrusters’ joint industry project and summarizes the conducted research on the material properties of carburized CrNiMo steels, the performed gear tests and the multiaxial stress and fatigue predictions to a, for spiral bevel gears applicable, standardizable TFF approach. The highlights and main findings of this article are:

- the apparent surface-parallel TFF crack initiation under the acting orthogonal shear stresses in case hardened spiral bevel gear teeth
- the compatibility of the analytically calculated orthogonal shear stress amplitude and the true maximum of the shear stress amplitude across all material planes in the TFF-critical depths
- the rationalization and standardization of the Dang Van criterion for the prediction of TFF based on the exclusive study of surface-parallel material planes
- a material model derived from conventional fatigue tests, gear testing and hardness measurements that utilizes ISO-inspired influence factors, namely:
  - a conversion factor $f_{xK}$ to convert the 50% to the 1% fatigue strength, similar in magnitude to tooth root breakage and pitting conversion factors [30]
  - a TFF lifetime factor $K_{NT}$ comparable to $Y_{NT}$ and $Z_{NT}$ of the ISO10300 [18, 19], suggesting a continuous strength decrease in the very high cycle fatigue regime
  - based on the weakest link theory [29], a TFF-specific, statistical size factor $K_{X}$, in magnitude equivalent to the ISO10300-3’s tooth root-specific size factor $Y_{X}$ [19], differing only in the chosen lower boundary
  - a technological size factor expressed through Thomas’ hardness peak factor $X_{HV,max}$ [31], capturing the need for a higher relative case hardening depth on large gears
- the general comparability of the Witzig-Boiadjiev [9, 32] and the herein proposed Dang Van criterion for the prediction of TFF, despite the very different calculation approaches
- the partially inaccurate TFF predictions from the DNV [14] and Hertter-Wirth [16, 33] criteria for the studied gear sets
- an inaccuracy in the Witzig-Boiadjiev’s allowable $\rho_{eq}$ range, suggesting that the criterion is most suitable for the study of centered contact points.

The presented TFF criterion can readily be applied to a single contact point, typically the gear’s design point or the entire gear tooth, if an LTCA is carried out. The model has been set up to yield a failure probability of 1% under constant load at $A = 1$ when considering $f_{xK}$ and 50% at $A = 1$ without $f_{xK}$. Due to the quadratic relationship between $\sigma_{H}$ and $A$, $A_{max}$ should be set conservatively depending on the reliability of the considered load data. While extensive gear calculations were carried out in the context of this research, further studies are necessary to verify the proposed influence factors.

Funding Open access funding provided by NTNU Norwegian University of Science and Technology (incl St. Olavs Hospital - Trondheim University Hospital)

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

1. Böhme SA, Vinogradov A, Biermann H, Weidner A, Schmiedel A, Henkel S (2020) Fatigue of carburised CrNiMo steel: Testing and modelling concept. Fatigue Fract Eng Mater Struct 44:788–804. https://doi.org/10.1111/ffe.13394
2. Böhme SA, Vinogradov A, Papuga J, Berto F (2021) A novel predictive model for multiaxial fatigue in carburized bevel gears. Fatigue Fract Eng Mater Struct 44:2033–2053. https://doi.org/10.1111/ffe.13475
3. Böhme SA, Szanti G, Keski-Rahkonen J, Komssi T, SANTAELLA JG, Vinogradov A (2022) Tooth flank fracture – An applied fatigue
study of case hardened bevel gears. Eng Fail Anal. https://doi.org/10.1016/j.engfailanal.2021.105911
4. Böhme SA, Merson D, Vinogradov A (2020) On subsurface initiated failures in marine bevel gears. Eng Fail Anal. https://doi.org/10.1016/j.engfailanal.2020.104415
5. Böhme SA (2022) Tooth flank fracture in spiral bevel gears – multiaxial fatigue and material properties. Thesis
6. Araújo J, Dantas A, Castro F, Mamiya E, Ferreira J (2011) On the characterization of the critical plane with a simple and fast alternative measure of the shear stress amplitude in multiaxial fatigue. Int J Fatigue 33(8):1092–1100. https://doi.org/10.1016/j.ijfatigue.2011.01.002
7. Baumann V, Bär G, Haase A, Hutschreiter B, Hünecke C (1998) Programm zur Berechnung der Zahnflanken – und Zahnfußbeanspruchung an Kegelrad – und Hypoidgetrieben (Heft 548)
8. Bernasconi A, Papadopoulos I (2005) Efficiency of algorithms for shear stress amplitude calculation in critical plane class fatigue criteria. Comput Mater Sci 34:355–368
9. Boiadjiev I (2014) Normfähiger Berechnungsansatz zum Flankenbruch bei Kegelrad- und Hypoidgetrieben. Forschungsvereinigung Antriebstechnik, München
10. Cerullo M (2014) Application of Dang Van criterion to rolling contact fatigue in wind turbine roller bearings under elasto-hydrodynamic lubrication conditions. Proc Inst Mech Eng C 228(12):2079–2089. https://doi.org/10.1177/0954406213516946
11. Crossland B (1956) Effect of large hydrostatic pressure on the torsional fatigue strength of an alloy steel. In: Proceedings of International Conference on Fatigue of Metals, London (Institution of Mechanical Engineers), pp 138–149
12. Dang Van K (1973) Sur la résistance à la fatigue des métaux. Sciences et Technique de l’Armement, vol 3eme fascicule, pp 647–722
13. Desimone H, Bernasconi A, Beretta S (2006) On the application of dang van criterion to rolling contact fatigue. Wear 260(4-5):567–572. https://doi.org/10.1016/j.wear.2005.03.007
14. DNV GL (2015) DNVGL-CG-0036: Calculation of gear rating for spur and helical gears—Part 4: Calculation of load capacity of spur and helical gears—Part 4: Calculation of tooth flank fracture load capacity. ISO copyright office, Switzerland
15. ISO (2019) ISO/TS 6336-4:2019—Calculation of load capacity of spur and helical gears—Part 4: Calculation of tooth flank fracture load capacity. ISO copyright office, Switzerland
16. Klein M (2013) Bestimmung der Frestragfähigkeit von Kegelrad- und Hypoidverzahnungen, Forschungsvorhaben Nr. 519 vol 1071 (Forschungsvereinigung Antriebstechnik e.V.)
17. Lang O (1989) Berechnung und Auslegung induktiv randschichtgehärteter Bauteile. In: Kloos KH (ed) Induktives Randschichthärten. Berichtsband, Tagung 23 bis 25 März 1988, München: Arbeitsgemeinschaft Wärmebehandlung und Werkstofftechnik (AWT), pp 332–348
18. van Lieshout PS, den Besten JH, Kaminski ML (2016) Validation of the corrected dang van multiaxial fatigue criterion applied to turret bearings of fpso offloading buoys. Ships Offshore Struct 12(4):521–529. https://doi.org/10.1080/17445302.2016.1182461
19. Markkula R, Meek W (2022) Benchmarking newer multiaxial fatigue strength criteria on data sets of various sizes. Metals. https://doi.org/10.3390/met12020289
20. Papuga J, Cízová E, Suchý L (2022) Benchmarking newer multiaxial fatigue strength criteria on data sets of various sizes. Metals. https://doi.org/10.3390/met12020289
21. Witzig J (2012) Entwicklung eines erweiterten Berechnungsverfahrens zur Ermittlung optimaler Zahnflankentragfähigkeit bis in den Bereich großer Werkstofftiefe – Ergänzungsversuche zur Verifizierung an größeren Zahnrad. Forschungsvorhaben Nr. 304 vol 260 (Forschungsvereinigung Antriebstechnik e.V.)
22. Thomas J (1998) Flankentragfähigkeit und Laufradverhalten von hartfeinbearbeiteten Kegelrädern. Thesis
23. Wickborn C (2013) Entwicklung eines erweiterten Berechnungsverfahrens zur Ermittlung optimaler Zahnflankentragfähigkeit in den Bereich großer Werkstofftiefe – Ergänzungsversuch zur Verifizierung an größeren Zahnräder. Forschungsvorhaben Nr. 556 II. Forschungsvereinigung Antriebstechnik, München
24. Witzig J (2012) Entwicklung eines erweiterten Berechnungsverfahrens zur Ermittlung optimaler Zahnflankentragfähigkeit bis in den Bereich großer Werkstofftiefe. Forschungsvorhaben Nr. 556 I vol 1000 (Forschungsvereinigung Antriebstechnik e.V.)
25. Zentner H, Simbürger A, Liu J (2000) On the fatigue limit of ductile metals under complex multiaxial loading. Int J Fatigue 22:137–145