Review on the Prediction of Residual Stress in Welded Steel Components

Junyan Ni¹, Xincun Zhuang²,³ and Magd Abdel Wahab⁴,⁵,*

Abstract: Residual stress after welding has negative effects on the service life of welded steel components or structures. This work reviews three most commonly used methods for predicting residual stress, namely, empirical, semi-empirical and process simulation methods. Basic principles adopted by these methods are introduced. The features and limitations of each method are discussed as well. The empirical method is the most practical but its accuracy relies heavily on experiments. Mechanical theories are employed in the semi-empirical method, while other aspects, such as temperature variation and phase transformation, are simply ignored. The process simulation method has been widely used due to its capability of handling with large and complex components. To improve its accuracy and efficiency, several improvements need to be done for each simulation aspect of this method.

Keywords: Computational simulation, steel, welding, residual stress.

1 Introduction

Residual stress residing in welded components is often caused by structural deformation, external thermal load or alteration of phase constituents. It is inevitably introduced during welding process and remains in the components or structures after the removal of external load or constraint [Hemmesi, Farajian and Boin (2017)]. In most situations, residual stress affects negatively the assembly accuracy and service performance [Rong, Xu, Huang et al. (2018)]. Therefore, study of residual stress has been focused by researchers and engineers for long time [Mahur, Bhardwaj and Bansal (2017); Rohde and Jeppsson (2000); Rong, Xu, Huang et al. (2018)]. To predict residual stresses, several approaches have been developed, such as interpolating measurements at selected points [Sharples, Gill, Wei et al. (2011)], analytically solving equilibrium equations [Dong...
numerically computing eigen-strain field [Kartal, Lijedahl, Gungor et al. (2008)], and modelling the complete welding process [Derakhshan, Yazdian, Craft et al. (2018)]. The interpolation approach can be considered as empirical method, in which no mechanical theories are used. The analytical solution and eigen-strain approach try to reconstruct the whole stress and strain fields with the help of mechanics, and is classified as semi-empirical method in this work. The modelling of complete process falls into the third group, which is named process simulation method. These three methods are discussed separately in the following sections.

2 Empirical method

With this method, residual stresses are empirically estimated by a series of equations. The form of and the parameters in equations are determined through several experiments. Mohr et al. [Mohr, Michaleris and Kirk (1997)] analyzed the factors that influence the residual stress distribution in girth butt welds. It was found that the most residual stresses in circumferential direction varied between 20 and 100 percentages of the material yield stress. Moreover, those values depend weakly on the thickness of pipe or the heat input. At the internal surface, the axial residual stress depend strongly on the pipe thickness as the heat input stays between 787.4 J/m and 1574.8 J/m per pass. After extensive investigation and analysis of experimental data, an improved set of equations describing the axial residual stresses $\sigma_{\text{intern}}$ of internal surface in girth butt welds was proposed as [Mohr, Michaleris and Kirk (1997)]:

$$
\sigma_{\text{intern}} = \begin{cases} 
\sigma_{\text{yield}} & n_p \leq 5 \\
\sigma_{\text{yield}} \left(1.33 - \frac{n_p}{15}\right) & 5 < n_p \leq 20 \\
0.0 & n_p > 20
\end{cases}
$$

(1)

where $n_p$ is the number of welding passes, and $\sigma_{\text{yield}}$ is the material yield stress. Bjørhovde et al. [Bjørhovde, Brozzetti, Alpsten et al. (1971)] conducted a similar parametric study on the welding of heavy plates, in which the influence of the plate dimensions was included with a width-factor $\beta$ as [Bjørhovde, Brozzetti, Alpsten et al. (1971)]:

$$
\beta = \frac{2(b_w + t_i)}{b_w^2 t_i}
$$

(2)

in which $b_w$ and $t_i$ are the plate width and the thickness, respectively. However, the interpolated empirical function for the residual stress distribution was not given. Nowadays, those empirical functions are summarized in standards for practical use, e.g., the British Standards Institute provides a guidance to assess the residual stress profiles, fracture resistance and fatigue of various welded joint [BSI (2005)]. Totally six types of weldments are considered in the standard [BSI (2005)]:

1) plate butt welds;
2) pipe seam welds;
Review on the Prediction of Residual Stress in Welded

3) pipe butt welds;
4) T-shape plate butt welds;
5) T-shape tubular/pipe welds;
6) repair welds.

Regarding the first geometry, the longitudinal residual stresses $\sigma_1$ at position $x$ for the austenitic steel is given as [BSI (2005)]:

$$\frac{\sigma_1}{\sigma_{\text{yield}}} = 0.95 + 1.505 \frac{x}{l} - 8.287 \left( \frac{x}{l} \right)^2 + 10.571 \left( \frac{x}{l} \right)^3 - 4.08 \left( \frac{x}{l} \right)^4$$  \hspace{1cm} (3)

where $l$ is the length of the butt welded plate shown in Fig. 1.

![Figure 1: Schematic of plate butt welds [BSI (2005)]](image)

For ferritic steels, the value is simply given as [BSI (2005)]:

$$\frac{\sigma_1}{\sigma_{\text{yield}}} = 1$$  \hspace{1cm} (4)

The transverse residual stresses $\sigma_t$ at position $y$ is interpolated as [BSI (2005)]:

$$\frac{\sigma_t}{\sigma_{\text{yield}}} = 0.9415 - 0.0319 \left( \frac{y}{b_w} \right) - 8.3394 \left( \frac{y}{b_w} \right)^2 + 8.660 \left( \frac{y}{b_w} \right)^3$$  \hspace{1cm} (5)

in which $b_w$ is the plate width as well. This value is not differentiated between austenitic and ferritic steels. The predictive functions for other weldments can be also found in the British Standards Institute standard similarly and are not stated in lengthy words here. A more general set of formulae including the effect of heat input can be found in Dong et al. [Dong, Song, Zhang et al. (2014b)]. As the welding condition changes, those formulae most probably need to be calibrated again. Moreover, they are only valid for width, yield strength and heat input listed in Tab. 1.
Table 1: Valid ranges for predicting the weld profile [BSI (2005)]

<table>
<thead>
<tr>
<th>Weldments</th>
<th>( b_w ) (mm)</th>
<th>( \sigma_{\text{yield}} ) (MPa)</th>
<th>Heat input (kJ/mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>24~300</td>
<td>310~740</td>
<td>1.6~4.9</td>
</tr>
<tr>
<td>(2)</td>
<td>50~85</td>
<td>345~780</td>
<td>-</td>
</tr>
<tr>
<td>(3)</td>
<td>9~84</td>
<td>225~780</td>
<td>0.35~1.9</td>
</tr>
<tr>
<td>(4)</td>
<td>25~100</td>
<td>375~420</td>
<td>1.4</td>
</tr>
<tr>
<td>(5)</td>
<td>22~50</td>
<td>360~490</td>
<td>0.6~2.0</td>
</tr>
<tr>
<td>(6)</td>
<td>75~152</td>
<td>500~590</td>
<td>1.2~1.6</td>
</tr>
</tbody>
</table>

Therefore, the empirical method is suitable for a fast and rough estimation of the residual stress profiles in similar welding conditions described in the standards or handbooks.

3 Semi-empirical method

3.1 Analytical solution

This method differs from the empirical approach in that residual stresses are obtained with analytical calculations. Yang et al. [Yang and Ziao (1995)] proposed an ideal elasto-plastic model to predict stress distribution across the welded panels. The panels were considered as cantilever bars presented in Fig. 2(a), where the stress possessed a parabolic distribution along the thickness.

Moreover, this stress was equilibrated to a combination of a uniform tension and bending load shown in Figs. 2(b) and 2(c). Using a similar approach, the stress profiles were decomposed into membrane portion \( \sigma_m \), bending portion \( \sigma_b \) and self-equilibrating portion \( \sigma_{\text{se}} \) [Sharples, Gill, Wei et al. (2011)]. Song et al. [Song, Dong and Pei (2015a, b)] solved these three portions by using the shell theory and coupling two key factors determined from parametric analysis namely, the thickness ratio \( r_p / t \) and the characteristic heat input \( \hat{Q} \).
\( r_p \) and \( t_t \) are the radius and thickness of the pipe shown in Fig. 3.

**Figure 3:** A representative drawing of the pipe girth weld [Dong, Song and Zhang (2014a)]

Finally, the distribution of through-pipe-thickness residual stress \( \sigma_{\text{thr}} \) is given as [Song, Dong and Pei (2015a)]:

\[
\frac{\sigma_{\text{thr}}}{\sigma_{\text{yield}}} = \bar{\sigma}_m + \bar{\sigma}_b + \bar{\sigma}_{\text{se}}
\]

where \( \bar{\sigma}_m \), \( \bar{\sigma}_b \) and \( \bar{\sigma}_{\text{se}} \) are the normalized terms of the three parts. The normalized self-equilibrating portion \( \bar{\sigma}_{\text{se}} \) is a complex term and can be expressed approximately in a polynomial form [Song, Dong and Pei (2015a)]:

\[
\bar{\sigma}_{\text{se}} = C_1 \bar{r}_t^2 + C_2 \bar{r}_t^3 + C_3 \bar{r}_t^4 + C_4 \bar{r}_t^5
\]

where \( \bar{r}_t = 2x/t_t - 1 \) is a dimensionless coordinate parameter with \( x \) measured from the internal surface of the pipe. \( C_1 \) to \( C_4 \) are coefficients to be determined from two self-equilibrating and two consistency conditions [Song, Dong and Pei (2015a)]. The normalized membrane and bending stresses are solved as [Song, Dong and Pei (2015a)]:

\[
\bar{\sigma}_m = C_{m1} \left( \dot{Q} \right) \ln \frac{r_p}{t_t} + C_{m2} \left( \dot{Q} \right)
\]

\[
\bar{\sigma}_b = C_{b1} \left( \dot{Q} \right) \ln \frac{r_p}{t_t} + C_{b2} \left( \dot{Q} \right)
\]

where \( C_{m1} \), \( C_{m2} \), \( C_{b1} \) and \( C_{b2} \) are four coefficients determined from the characteristic heat input \( \dot{Q} \).

### 3.2 Eigen-strain reconstruction

The main concern of this method is to construct a stress field inherited with eigen-strain. Eigen-strain was introduced by Mura [Mura (2012)] to describe the permanent deformation resulting from inelastic phenomena, such as phase transformation, plastic deformation and mismatch between assembled parts. Using the assumption of small
strain, the total strain $\varepsilon$ was additively decomposed as:

$$
\varepsilon = \varepsilon^e + \varepsilon^* 
$$

(10)

where $\varepsilon^e$ is the elastic strain. $\varepsilon^*$ is the eigen-strain accounting for residual stress and equals to the summation of all inelastic strains. Castles et al. [Castles and Mura (1985)] expressed the relationship between a known eigen-strain distribution and the corresponding elastic strain through the Green’s function. However, this function is only explicitly known in the case of special geometries, such as infinite and semi-infinite spaces [Qin, Fan and Mura (1991)]. As a result, it is not practical to solve the problem directly for complex geometries.

To retrieve the stress distribution practically, residual strains at selected points are measured by diffraction techniques at first. Based on the measurements, the complete eigen-strain field can be obtained inversely through a repeated process until a target function is minimized [Luckhoo, Jun and Korsunsky (2009)]. This inverse algorithm expresses the unknown eigen-strain in a series of basis function [Luckhoo, Jun and Korsunsky (2009)]:

$$
\varepsilon^* = \sum_{i=1}^{N_b} c_i \psi_i 
$$

(11)

where $N_b$ is the basis function number and $\psi_i$ is the corresponding function. $c_i$ is the unknown coefficient. Then, the elastic strain is written in terms of eigen-strain with certain initial guess of coefficients, and is put into the linear elastic constitutive law to figure out the total strain [Korsunsky (2006)]. Normally, the calculated total strain deviates from the measurement. The coefficients are optimized repeatedly through the least squares method until the differences between the calculated and measured total strain are acceptable. With those optimized values, the residual stress is obtained by solving the elastic constitutive law eventually. The choice of the basis function varies from authors to authors [Cao, Hu, Lu et al. (2002); Korsunsky, Regino and Nowell (2007); Qian, Yao, Cao et al. (2004, 2005)]. Cao et al. [Cao, Hu, Lu et al. (2002)] adopted a series of two-term polynomials and trigonometric polynomials. Qian et al. [Qian, Yao, Cao et al. (2004)] chosen the same forms but employed boundary element method to solve the strain rather than finite element method. Kartal et al. [Kartal, Lijedahl, Gungor et al. (2008)] adopted the Legendre polynomials, while Korsunsky et al. [Korsunsky, Regino and Nowell (2007)] used the Chebyshev polynomials as bases functions. The contour plot of the eigen-strain field in the $y$ direction of butt welded plate calculated in a symmetric model is presented in Fig. 4.
A simplified version of this approach that simply treats the eigen-strain as thermal expansion is developed by Luo et al. [Luo, Ishiyama and Murakawa (1999)]. Implementing this kind of approach requires writing subroutines in commercial or open source software. The choice of basis function also depends on the experience of the code developer [Deng, Murakawa and Liang (2007)].

As the welded structure becomes complicated or the order of function increases, the calculation of those unknown coefficients $c_i$ may not converge [Kartal, Lijedahl, Gungor et al. (2008)]. Moreover, except the stress and strain fields, the other information, such as the temperature and the microstructure fields, are simply ignored in this method.

### 4 Process simulation method

In this method, the whole welding process is simulated. Reliable process simulation requires extensive efforts to track material properties and states. A common concern in welding simulation includes three aspects, i.e., temperature, microstructure and deformation [Ueda, Ronda, Murakawa et al. (1994)]. The temperature field interacts with the stress/strain field all the time. The microstructure varies as the temperature and stress change. Inversely, the modification of microstructure affects the development of temperature and stress. To increase the simulation efficiency, many authors [Deng and Murakawa (2008); Kang and Im (2007); Lee and Chang (2009); Lee, Chiew and Jiang (2013); Lindgren (2001a, b); Ma, Cai, Huang et al. (2015)] adopted the sequential coupling approach, in which the temperature field is solved first and is used as external load for the subsequent analyses. The thermal history is then used as an external load leading to the evolution of microstructure and deformation. In such a coupling sequence, a metallo-thermo-mechanical model including all the three aspects is often built. Different specific models for different physical phenomena are discussed in separate sections below.
4.1 Heat source models

The thermal model is the approximation of weld heat source. It reduces the physical complexity of input heat, and allows the temperature field to be solved in an isolated system. The first attempt that implements the heat source model using analytical solutions was made by Rosenthal [Rosenthal (1946)], in which the energy was supposed to concentrate in an artificial point and simple geometric domains were assumed. The point model is quite attractive when simulating the welding process of thick plates [Nunes (1983)]. Because the heat source is treated as monopole, the analytical solutions of power density and temperature tend to be infinite as it comes close to that point. As it is sufficiently far from the weld pool, an accurate prediction was obtained. Besides the point model, the heat source can be specified as a line segment, along which the energy was uniformly distributed. The line model was suitable for situations where laser or electron beam fully penetrates sheets or plates, but led to infinite temperature inside the applied position either [Rosenthal (1946)]. Ashby et al. [Ashby and Easterling (1982)] placed the line source above the plate so that the infinite temperature inside the specimen was avoided. By combining the point and line models, Steen et al. [Steen, Dowden, Davis et al. (1988)] developed a new heat source that is able to regard the keyhole mode for welding specimens of infinite thickness. An improvement was made by Akhter et al. [Akhter, Davis, Dowden et al. (1989)] to extend the combined model suitable for plates of finite thickness. Dowden et al. [Dowden, Ducarme and Kapadia (1998)] proposed a new model by generalising the ones proposed by Steen et al. [Steen, Dowden, Davis et al. (1988)] and Akhter et al. [Akhter, Davis, Dowden et al. (1989)], and provided analytical solutions for periodic moving point and line sources. A summary of the analytical solutions was made by Lindgren [Lindgren (1986)], in which three different types of solutions were studied in five simulations. The temperature solution of the point model in a thin and infinite plate at \( x - y \) plane was recalculated as [Lindgren (1986)]:

\[
T(x, y, t) = \frac{Q_{in}v}{2\pi k t_i} B_0(\omega r) \exp(-\omega y + \omega vt)
\]  

(12)

with

\[
r = \sqrt{x^2 + (y - vt)^2}
\]  

(13)

\[
\omega = \frac{v \rho C}{2k}
\]  

(14)

where \( Q_{in} \) is the effective input energy. \( v \) is the velocity of heat source, and \( k \) is the thermal conductivity. Similarly, \( t_i \) denotes the thickness of the thin plate. \( \rho \) and \( C \) are the heat capacity and density of the material. \( B_0 \) is the modified zero-order Bessel function of second type. The initial time at which the point source is applied is set to be zero. Along the welding direction, the heat flow could be ignored for low welding speed, and the temperature distribution could be obtained using this approach if the residual stress is of interest [Lindgren (1986)].

The intensity of heat source can be also described by distribution functions instead of Dirac delta function. Pavelic [Pavelic (1969)] first proposed the distributed heat source
model, in which the heat flux was deposited on the surface of the sample with a Gaussian distribution. This approach was appropriate for modelling heat source of low power density. Krutz et al. [Krutz and Segerlind (1978)] suggested a modified form of this model. They implemented this model in finite element analysis, and close approximations of fusion zone (FZ) and heat affected zone (HAZ) were achieved [Krutz and Segerlind (1978)]. As the density increases, the penetration of heat source needs to be considered. Therefore, Goldak et al. [Goldak, Chakravarti and Bibby (1984)] developed a non-axisymmetric three-dimensional model that could simulate somewhat complex weld pool. His model is widely known as double ellipsoid model and frequently used for simulating welding processes [Joshi, Hildebrand, Aloraier et al. (2013)]. Assuming that the heat source moves at \( x-z \) plane and along \( x \) direction, the heat flux \( q_{in} \) in the front quadrant is expressed as [Flint, Francis, Smith et al. (2017)]:

\[
q_{in} = \frac{6\sqrt{3}Q_{m}f_{t}}{a_{t}b_{c}\pi\sqrt{\pi}}\exp\left(-\frac{3\left[x+v(t-t_{0})\right]^{2}}{a_{t}^{2}}-\frac{3y^{2}}{b^{2}}-\frac{3z^{2}}{c^{2}}\right)
\]  

(15)

where \( f_{t} \) is the input energy fraction of front quadrant. Parameters \( a_{t}, b \) and \( c \) are the lengths of ellipsoid semi-axes shown in Fig. 5.

\[ \text{Figure 5: Heat source distribution: a) double ellipsoid source and b) conical source} \]

The expression in the rear quadrant has the same form but with different semi-axis length \( a_{r} \) in its moving direction, and different power fraction \( f_{r} \). The two fractions satisfy the condition that \( f_{t} + f_{r} = 2 \). \( t_{0} \) is the time at which the source is applied, and \( v \) is again the welding speed. In fact, this model was established by extending the distributed model of Pavelic [Pavelic (1969)] into the spherical and later ellipsoidal configurations. The most general form of this kind was created by distinguishing ellipsoid semi-axes in left and right quadrants or even using four independent semi-axes to consider the welding process of dissimilar materials [Goldak, Chakravarti and Bibby (1984)].

Another example for the deep penetration of the heat source is conical model as presented in Fig. 5(b). The power density of this model is described with a Gaussian distribution on
the truncated plane (e.g., $x-z$) as [Wu, Wang and Zhang (2006)]:

$$q_{in} = q_0 \exp\left( -\frac{3r^2}{c^2} \right)$$  \hspace{1cm} (16)

where $r$ has the same expression as in Eq. (13). $q_0$ is the maximum heat flux. $c$ is characteristic radius at the current position $y$ and linearly decreases from the top surface at $y_i$ to the bottom $y_b$ as [Wu, Wang and Zhang (2006)]:

$$c = r_i - \frac{(r_i - r_b)(y_i - y)}{y_i - y_b}$$  \hspace{1cm} (17)

where $r_i$ and $c$ are the corresponding characteristic radii at top and bottom. Those mentioned models can be also combined to form other shapes of volumetric heat sources [Dal and Fabbro (2016)].

Paley et al. [Paley and Hibbert (1975)] simplified the application of those models by assigning heat power directly to the areas belonging to FZ. Brickstad et al. [Brickstad and Josefson (1998)] simply applied the averaged power per welding pass for multi-pass welding simulation. With an impulse equation, Hibbitt et al. [Hibbitt and Marcal (1973)] regarded the approaching and leaving arcs as ramps of linearly increasing and decreasing surface heat input, and kept the value constant as the material is heated above the melting point.

The prescribed temperature serves as the third kind of approaches for approximating the heat source. An artificial temperature field is directly applied to the subsequent analyses as external load. With a prescribed temperature, Goldak et al. [Goldak, Zhou, Breiguine et al. (1996)] conducted simulations that are able to switch constitutive equations from rate independent to rate dependent. Recent functional extension for the process simulation in commercial software ABAQUS has been devoted to create such a temperature field as well [Shubert and Pandheeradi (2014)]. The approach of prescribed temperature that has been discussed so far needs a beforehand thermal analysis. Börjesson et al. [Börjesson and Lindgren (2001)] simply assigned the material melting temperature to FZ and kept the rest at room or pre-heated temperature.

By including fluid dynamics of the melt in simulation, the heat flow can be described by the Navier-Stokes equations as well. The merit of this approach resides in the capability to regard the effect of advection on bead geometry and molten pool [Cheon and Na (2016)]. The impinging effect of molten droplets on the molten pool was investigated by Cao et al. [Cao, Yang and Chen (2004)] in a Gas Metal Arc Welding process. Cho et al. [Cho, Lim and Farson (2006)] conducted a parametric study on current density, arc force and pressure radius using computational fluid dynamics technique. To implement the thermal results for metallurgical analysis, a combined framework of computational fluid dynamics and finite element analysis was established by Cheon et al. [Cheon, Kiran and Na (2016)]. Moreover, they applied the same temperature result to predict residual stresses during a bead-on-plate welding process [Cheon and Na (2017)].
4.2 Solid-to-solid phase transformation models

For fusion welding, the input energy is always high enough to cause the growth of grain and the change of microstructure in the HAZ. Due to the fact that HAZ connects to FZ on one side and to the base material on the other side, the modification of microstructure, especially the dominance of hard products such as martensite and bainite, renders the welded structures vulnerable to fracture and corrosion. Fig. 6 presents the characteristic microstructure in different zones of a weld bead.

Figure 6: Representative of microstructures in different zones of a single-pass weld [Francis, Bhadeshia and Withers (2007)]

The material in FZ transforms into a full austenite structure as the temperature falls a little below the solidification point, and then experiences a solid-to-solid phase transformation [Francis, Bhadeshia and Withers (2007)]. HAZ is partially heated, but is subjected to the same kind of phase transformation. The insufficient heating leads to an even more complex situation in the HAZ. As depicted in Fig. 6, the HAZ can be subdivided into coarse grain HAZ, fine grain HAZ and inter-critical HAZ based on the extent to which the material is austenitized and austenite grain grows. Besides, an over-tempered area in the base material can be observed.

Therefore, tracking the microstructure during simulation plays an importance role of predicting residual stresses. A robust metallurgical algorithm should be able to determine the transformation temperatures, and to describe kinetics of austenitization, grain growth and austenite decomposition. Leslie [Leslie (1981)] proposed an equation to estimate the austenite/austenite+ferrite boundary \( A_1 \) empirically. The eutectoid temperature \( A_t \) for pearlite was calculated similarly [Andrews (1965)]. The bainite and martensite transformation temperatures, \( B_s \) and \( M_s \), were predicted by Kirkaldy et al. [Kirkaldy and Venugopalan (1984)], and Andrews [Andrews (1965)] respectively. Those equations were
summarized and adopted by Watt et al. [Watt, Coon, Bibby et al. (1988)] to constitute the overall metallurgical model. The shift of those transformation starts temperatures due to non-isothermal condition was ignored. With modified parameters, the transformation temperatures were estimated with the same methodology by other authors for microstructure computation as well [Bok, Choi, Barlat et al. (2014); Hamelin, Muransky, Smith et al. (2014); Jiang, Chen, Woo et al. (2018); Lee, Kim, Han et al. (2009a, b)]. Instead of empirical estimation, computational thermodynamics offers an alternative way to predict those temperatures. Shiflet et al. [Shiflet, Bradley and Aaronson (1978)] laid a solid foundation for calculating $A_3$ in iron-carbon alloys, which was extended to an algorithm capable of dealing with additions of other elements [Kirkaldy and Bagans (1978)]. By adding extra equilibrium condition between austenite and cementite, $A_i$ could be also theoretically calculated [Takahashi (1992)]. Besides, Bhadeshia [Bhadeshia (1981, 1997)] proposed the approaches for calculating $B_s$ and $M_s$ based on the equilibrium of free energy potentials. It should be mentioned that $A_3$ and $A_i$ are the critical temperatures observed at the equilibrium state of phase transformation. In most situations, the material is subjected to continuous variation of temperature so that the transformation temperatures may shift. Therefore, the additive rule of Scheil [Scheil (1935)] is often coupled to predict the transformation temperatures in non-isothermal conditions based on the values at equilibrium [Hamelin, Muransky, Smith et al. (2014); Kang and Im (2007)]. The idea is to discretize the continuous heating or cooling history into small isothermal steps as presented in Fig. 7. As long as the sum of the fraction between the discretized time and the isothermal incubation time reaches one, the transformation occurs.

![Figure 7](image-url)

**Figure 7:** (a) Time Temperature Transformation (TTT) and (b) Continuous Cooling Transformation (CCT) diagrams [Chen, Xiao, Li et al. (2014)]

The austenitization kinetics was often assumed to be proportional to the temperature increment between $A_i$ and $A_3$ as Deng et al. [Deng and Murakawa (2008); Ma, Cai,
Huang et al. (2015)):

\[ f_\gamma = \frac{T - A_i}{A_3 - A_i} \]  

(18)

where \( f_\gamma \) is the volume fraction of austenite. \( T \) is the temperature at which the volume fraction is calculated.

Andres et al. [de Andres, Caballero, Capdevila et al. (1998)] presented a model that is able to modelling the pearlite-to-austenite transition upon heating. This model was extended to describe the transformation kinetics from ferrite to austenite as well [Mi, Xiong, Wang et al. (2016)]. The extent of austenitization is given as Mi et al. [Mi, Xiong, Wang et al. (2016)]:

\[ f_\gamma = \left[ 1 - \exp \left( -\frac{\pi}{3} N_\gamma G_\gamma t^3 \right) \right] f_{\gamma,\text{eq}} \]  

(19)

where \( f_{\gamma,\text{eq}} \) is the equilibrium austenite volume fraction. \( N_\gamma \) and \( G_\gamma \) are the nucleation and grain growth rates of austenite, respectively.

A similar work was presented by Oddy et al. [Oddy, McDill and Karlsson (1996)] in which two austenitization models, one for instantaneous and homogeneous formation and one for transient and heterogeneous formation, were included. Watt et al. [Watt, Coon, Bibby et al. (1988)] expected that austenitization occurs under near equilibrium conditions upon. However, the kinetic function of austenite growth was not given. A more general model that describes the austenite increment was proposed by Leblond et al. [Leblond and Devaux (1984)]:

\[ \frac{df_\gamma}{\tau_c} = \frac{f_{\gamma,\text{eq}} - f_\gamma}{\tau_c} \]  

(20)

where \( \tau_c \) denotes the characteristic time for transformation and is a temperature-dependent value. It needs to be calculated by interactively comparing \( A_i \) and the shifted temperature \( A_{i1} \) due to continuous heating [Leblond and Devaux (1984)]. By assuming that the isothermal transformation time of austenitization was equal to the time spent in austenite decomposition and that the continuous heating can be discretized into small isothermal steps [Scheil (1935)], the extent of austenitization can be also calculated in equality to the fraction between the step time and the transformation time [Ni and Wahab (2017)]:

\[ \frac{df_\gamma}{dt} = \frac{1 - f_\gamma}{\tau_f - \tau_i} \]  

(21)

where \( \tau_i \) is the initial time for the isothermal transformation. \( \tau_f \) represents the corresponding finish time.

Austenite decomposition is important in the metallurgical analysis of welding process as well. Depending on the transformation mechanism, two types, i.e., reconstructive and displacive transformation, are classified [Thewlis, Whiteman and Senogles (1997)].
Reconstructive transformation is also known as diffusional transformation, and is featured by new grain growth and thermally activated jumps of individual atoms at the boundary between the new and parent phases [Novikov (2002)]. Instead, a migration of the atoms via a cooperative movement smaller than the inter-atomic spacing in the parent phase leads to the displacive transformation, which is identically recognized as non-diffusional transformation. In other words, all bonds in the original phase are broken to form new ones in reconstructive transformation, while in displacive transformation, the new phase is formed by coordinated displacement [Caballero (2014)].

Johnson-Mehl-Avrami-Kolmogorov (JMAK) equation is the most widely used equation that describes the isothermal kinetics of reconstructive transformation during austenite decomposition, and is expressed as [Deng and Murakawa (2006); Kang and Im (2007); Lee, Kim, Han et al. (2009b); Ma, Cai, Huang et al. (2015)]:

\[
f_i = 1 - \exp\left(-k_i t^n_i\right)
\]

where \( f_i \) is the volume fraction of phase \( i \), \( k_i \) is the parameter related to the growth and nucleation rates mentioned in Eq. (21). Similarly, \( n_i \) is the other parameter related to the change of temperature, size of austenite grain, chemical composition, etc. [Xu, Lu, Yu et al. (2013)]. A representative curve of JMAK is shown on the top of Fig. 7(a). Similarly, this equation needs to be combined with the additive rule so that it is suitable for continuous cooling condition. Several authors [Bok, Choi, Suh et al. (2015); Chen, Xiao, Li et al. (2014); Kamamoto, Nishimori and Kinoshita (1985)] modified the original equation in order to provide better predictions of microstructure evolution. Before usage, the kinetic parameters of JMAK equation need to be derived through the TTT diagram shown in Fig. 7(a) or other tests, such as dilatometry, electrical resistivity method and salt bath technique combined with metallography and hardness measurement. The general model proposed by Leblond et al. [Leblond and Devaux (1984)] mentioned above is applicable to austenite decomposition as well. The equilibrium volume fraction and the characteristic time need to be deduced from the phase and CCT diagrams alike. Their model was capable of reproducing satisfying metallurgical results over a wide range of cooling rates and was adopted by other researchers [Bergheau and Leblond (1991); Schenk, Richardson, Kraska et al. (2009); Zain-ul-abdein, Nelias, Jullien et al. (2011)]. Moreover, this model has already been integrated in commercial software SYSWELD as the metallurgical algorithm [Ferro, Porzner, Tiziani et al. (2006)].

The development of phase field method makes it possible to handle with grain coarsening, impingement phenomenon, diffusion, and interface mobility [Loginova, Odqvist, Amberg et al. (2003)]. The microstructure state is represented by the order parameter \( \phi_i \), which indicates the region covered by the corresponding phase [Qin and Bhadeshia (2010)]. In consideration of a polycrystalline system, the governing equation of this method is given as [Cho, Kim, Cho et al. (2012)]:

\[
\frac{\partial \phi_i}{\partial t} = -\frac{2}{s} \sum_{\alpha \beta} s_{i\beta} M_{ij} \left[ \frac{\delta F}{\delta \phi_i} - \frac{\delta F}{\delta \phi_j} + F' - F' - \left(x_i - x_j\right) F_c \right]
\]
where $s_i$ is a step function that [Cho, Kim, Cho et al. (2012)]:

$$ s_i = \begin{cases} 
1 & \text{if } \phi_i > 0 \\
0 & \text{otherwise}
\end{cases} \tag{24} $$

satisfies:

$$ s = \sum_{i=1}^{n_{\text{ph}}} s_i \tag{25} $$

where $s$ is the number of grains coexisting at a given position. $M_{ij}$ is the mobility of phase $i$ to $j$. $n_{\text{ph}}$ is the phase number. $F_{i}$ is the total free energy including both grain and inter-phase boundary contributions. $F_i^i$ and $x_i$ are the free energy and the carbon concentration in $i$th phase. $F_{C}^{i}$ is the partial free energy of carbon in phase $i$. The order parameter $\phi_i$ satisfies that [Cho, Kim, Cho et al. (2012)]:

$$ \sum_{i=1}^{n_{\text{ph}}} \phi_i(x,y,z) = 1 \tag{26} $$

In the case of $n_{\text{ph}}$-phase system, $\phi_i = 1$ represents that phase $i$ exists at the given position [Qin and Bhadeshia (2010)]. $\phi_i = 0$ indicates its absence and $0 < \phi_i < 1$ means the bounding interface. With the metallurgical results solved by this method, Cho et al. [Cho, Kim, Cho et al. (2012)] included the effect of phase transformation when predicting the evolution of residual stress.

The models that have been described so far needs calibration with experimental data before implementation, that is to say, the transformation parameters inside should be adjusted if the material changes. Comparably, Kirkaldy et al. [Kirkaldy and Venugopal (1984)] developed a series of incremental functions which only require chemical composition and thermal history as inputs. Watt et al. [Watt, Coon, Bibby et al. (1988)] sorted out these equations to simulate simultaneous decomposition of austenite, which were implemented in a finite element analysis of their subsequent work [Henwood, Bibby, Goldak et al. (1988)]. Later, Dai et al. [Dai, Francis and Withers (2010)] employed the CCT diagram predicted by this algorithm in a commercial software for stress analysis. The general expression of this algorithm is expressed as [Dai, Francis and Withers (2010)]:

$$ \frac{df_i}{dt} = \text{Func}_i (D_{\text{ASTM}}, T, \text{chemical composition}) f_i^{0.4(1-f_i)} (1-f_i)^{0.4f_i} \tag{27} $$

where $D_{\text{ASTM}}$ is the grain size number of American Society for Testing and Materials (ASTM). $\text{Func}_i$ is a specific function depending on ASTM grain size number, temperature and chemical composition. For each transformation, $\text{Func}_i$ has no physical meanings and is simply empirically evaluated [Dai (2012)].

Instead, Jones et al. [Jones and Bhadeshia (1997)] proposed a phenomenological model for
predicting the microstructure during steel welding. The derivation of their model started with the extension of the JMAK equation to the multi-phase transformation case. Two stages, nucleation and growth, were handled separately in their model. The nucleation rate was considered temperature-dependent, and the growing phases were treated as the accumulation of deposited new phase layers. The deposition is controlled by carbon diffusion. The example of ferrite growth is shown in Fig. 8, where the newly generated ferrite is regarded as disc plotted in black shadow.

\[ f_{\alpha} = 1 - \exp\left[-k_{\alpha}\left(M_s - T\right)\right] \]  

(28)

where \( f_{\alpha} \) is martensite volume fraction and \( M_s \) the start temperature for transformation. \( k_{\alpha} \) is the corresponding kinetic parameter. The model of Leblond et al. [Leblond and Devaux (1984)] can be also easily applied to the KM relationship by setting a small value for characteristic time of transformation. Besides, Khan et al. [Khan and Bhadeshia (1990)] proposed a new relationship for the austenite-to-martensite transformation based on a series of quenching tests. Comparably, there also exist several models for the transition to bainite during welding. Xu et al. [Xu, Lu, Yu et al. (2013)] simply employed the JMAK equation for the bainite transformation though it is intended for reconstructive transformation. The model developed by Leblond et al. [Leblond and Devaux (1984)] can be used for the same transformation as well. Watt et al. [Watt, Coon, Bibby et al. (1988)] and Dai [Dai (2012)] simply followed the same type of incremental function proposed by Kirkaldy et al. [Kirkaldy and Venugopalan (1984)] for the bainite reaction. Rees et al. [Rees and Bhadeshia (1992)] proposed another phenomenological model in which bainite was assumed to grow as attaching one platelet to another. A description of this mechanism is presented in Fig. 9.

**Figure 8:** A description of ferrite growth in austenite grain [Ni, Wang, Gong et al. (2018)]

Moreover, this model has the same advantage that only the chemical composition is needed as the input for calculation.

The displacive transformation of steel mainly refers to the generation of bainite and martensite. Almost all the research works [Chen, Xiao, Li et al. (2014); Deng and Murakawa (2008); Kang and Im (2007); Lee and Chang (2009); Lee, Kim, Han et al. (2009a)] adopted the Koistinen-Marburger (KM) relationship for martensite transformation, which is given as:
The austenite grain size (AGS) before austenite decomposition, which is also known as prior AGS, is another important factor for the transformation kinetics although only a few researchers [Dai (2012); Ferro, Porzner, Tiziani et al. (2006)] considered in the complete thermo-metallo-mechanical analysis of welding. Ikawa et al. [Ikawa, Shin, Oshige et al. (1977)] used a plausible equation in which the increment of AGS was proportional to the exponential function of the energy ratio. Later, Ashby et al. [Ashby and Easterling (1982)] extended the original function to a more general form, in which the austenite grain was assumed to grow as [Watt, Coon, Bibby et al. (1988)]:

$$\frac{d\bar{D}}{dt} = \frac{1}{2D} k_{\tau} \exp\left(-\frac{Q_{app}}{RT}\right)$$  \hspace{1cm} (29)

where $\bar{D}$ is the measure of AGS. $k_{\tau}$ and $Q_{app}$ are two constants related to the AGS growth. The influences of carbides and nitrides on the grain boundary mobility were discussed in the research work by Andersen et al. [Andersen and Grong (1995); Andersen, Grong and Ryum (1995)]. Leblond et al. [Leblond and Devaux (1984)] extended the original function in another aspect so that the evolution equation of AGS also depends on transformation kinetics.

### 4.3 Plastic constitutive models

The constitutive models are needed for solving residual stresses in the process simulation method. Using that depend on temperature is the earliest model for simulating welding process. Tall [Tall (1961)] adopted the elasto-plastic property of no work-hardening for simulation. A representative curve of this material behaviour is plotted in dash line of Fig. 10(a).

![Figure 9: A schematic diagram of bainite transformation](image)

![Figure 10: Schematics of (a) elasto-plastic, (b) elasto-viscoplastic and (c) transformation plastic behaviours](image)
A later work by Ueda et al. [Ueda and Yamakawa (1971)] assumed no hardening in their research either. To improve the model, piecewise linearly interpolated isotropic hardening and kinematic hardening behaviours were used in their subsequent researches [Shimizu (1986); Ueda (1990)]. Rybicki et al. [Rybicki, Schmueser, Stonesifer et al. (1978)] and Troive et al. [Troive, Karlsson, Näsström et al. (1990)] simply adopted the isotropic hardening constitutive equations. Börjesson et al. [Börjesson and Lindgren (2001)] calculated the isotropic hardening properties as the multiplications of single phase properties and corresponding volume fractions. The residual stresses profiles after the multi-pass welding process are presented in Fig. 11.

**Figure 11:** Transverse residual stress and longitudinal residual stress for welds in upper groove [Börjesson and Lindgren (2001)]. (Distance in meter)

**Figure 12:** Distribution of the predicted residual stress using different constitutive equations compared with the measured data. (a) Longitudinal and (b) Transverse [Wang, Liu, Wang et al. (2017)]. (IH: isotropic hardening, KH: kinematic hardening, PP: perfect plastic, CH: combined hardening)

Kinematic hardening model was also employed [Michaleris (1996); Tekriwal and Mazumder (1991)]. Recently, Wang et al. [Wang, Liu, Wang et al. (2017)] investigated the influences of different hardening models on the prediction of residual stresses. By adopting various constitutive models, the isotropic hardening model was found to provide
the prediction of residual stress with highest accuracy when simulating the 10Ni5CrMoV high strength steel welding process. The comparison between measurements and various predictive results are shown in Fig. 12.

The constitutive models discussed until now are rate-independent. Several authors Argyris et al. [Argyris, Szimmat and Willam (1982); Goldak, Zhou, Breiguine et al. (1996); Ronda and Oliver (1998); Wang and Inoue (1985)] used rate-dependent models for welding simulation as presented in Fig. 10(b). Argyris et al. [Argyris, Szimmat and Willam (1982)] computed the distribution of residual stress using a thermo-elasto-viscoplastic model. Wang et al. [Wang and Inoue (1985)] used a viscoplastic constitutive relation to describe the change of material behaviour from solid to liquid. Goldak et al. [Goldak, Zhou, Breiguine et al. (1996)] presented a methodology that changes constitutive equations from elasto-plastic to elasto-viscoplastic at a temperature of 50 percentage of the melting point. As the temperature arises above 80 percentage of the melting point, the model becomes linear viscoplastic. Ronda et al. [Ronda and Oliver (1998)] compared the feasibility of different thermo-viscoplastic constitutive models for welding simulation. Quantitatively different residual stresses and strains were obtained and discussed. To account for recovery and recrystallization during heating period, Geijselaers [Geijselaers (2003)] added a relaxation term to the overall constitutive equations.

The effect of phase transformation on the stress-strain evolutions is another aspect that cannot be neglected. The modification of microstructure during welding causes plastic deformation even if the mechanical load that the material is subjected to is smaller than the yield stress [Leblond, Mottet and Devaux (1986a)]. This behaviour is depicted in Fig. 10(c), and is explained by two proposed mechanisms. Greenwood et al. [Greenwood and Johnson (1965)] considered that the volume difference between two coexistent phases produces internal stresses, and that the stress is large enough to induce plasticity in the phase of lower yield stress. In the second mechanism, Magee et al. [Magee and Paxton (1966)] reasoned that martensite is generated with a preferred orientation in presence of external load, leading to local strain variations that cannot be averaged out to zero. A complete framework of theoretical approach to describe the plastic behaviour of steels during phase transformation was first formulated by Leblond et al. [Leblond, Mottet and Devaux (1986a)]. The transformation plastic behaviour was first investigated in ideally plastic phases [Leblond, Mottet and Devaux (1986b)]. Three years later, Leblond et al. [Leblond (1989)] presented a more general model, in which the term of transformation plastic strain was explicitly derived by homogenization procedure and hardening effects were taken into account. Taleb et al. [Taleb and Sidoroff (2003)] released the assumption of neglecting elasticity in austenite during homogenization so that a cut-off function was derived naturally to avoid singularity. Further, both elasticity and plasticity were regarded in austenite when homogenizing the increment of microscopic equivalent plastic strain with respect to the variation of phase volume fraction [Weisz-Patrault (2017)]. A simplified version of this framework simply regards the transformation plastic deformation as a priori term that is proportional to the applied stress and the transformed volume fraction [Gautier, Denis, Liebaut et al. (1994); Hamelin, Muransky, Smith et al. (2014); Jiang, Chen, Woo et al. (2018)].
5 Conclusion

The current work is devoted to summarize the main methods of predicting residual stress residing in steel structures. Three types are roughly categorized. The empirical and semi-empirical methods attempt to construct the residual stress field using known values. The idea of the first method is simple and straightforward. The semi-empirical method improves the treatment of the measured data, and solves the unknown fields based on certain theories. However, neither of them is able to capture other phenomena, such as phase transformation and temperature variation. The method of process simulation arises as an option to track all the occurrences during welding. The simulation often includes three aspects, i.e., thermal, metallurgical and mechanical analyses. Among various heat source models, the double ellipsoid source is shown to be a flexible approximation of the heat input. For the solid-to-solid phase transformation, most existing models require calibration with extensive experiments before implementation. The serial kinetic functions proposed by Kirkaldy et al. [Kirkaldy and Venugopalan (1984)] form a framework able to work independently of metallurgical measurements. However, most parameters inside are evaluated empirically. A more flexible metallurgical algorithm based on metallurgical principles has been also implemented for describing the solid-to-solid phase transformation [Ni and Wahab (2017)]. Traditional constitutive models, such as elasto-plastic and viscoplastic, are the earliest ones employed for modelling welding processes. A generalized framework of mechanical constitutive equations was derived by Leblond et al. [Leblond, Mottet and Devaux (1986a)], in which the transformation plastic behaviour was reasonably explained while the strain contributed by Magee’s mechanism [Magee and Paxton (1966)] was not explicitly given. Regarding to the aspects above, we consider that the following improvement can be made in future:

- A one-to-one relationship between the parameters of heat source (such as the laser spot diameter and focus) and the ones in heat source model (such as $a_r$, $b$ and $c$ in Eq. (15)) can be established so that the process simulation method becomes more efficient.

- The influence of local inhomogeneity of chemical composition on the evolution of residual stress can be analysed using the metallurgical models that only require composition as input.

- The transformation plastic framework can be further enhanced, such as releasing some restrictions when deriving the macroscopic terms and including the strain contribution of Magee’s mechanism [Magee and Paxton (1966)] in the constitutive equations.

Acknowledgement: The authors would like to acknowledge the supports from DeMoPreCI-MDT SIM SBO project.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.
References


Tall, L. (1961): Residual stresses in welded plates-A theoretical study, No. 235 (64-1). Spring Meeting of the American Welding Society in Dallas.


