Neighborhood and surface effects on polycrystal stress field extreme values: An analysis in linear elastic range by means of cellular automaton

R. Bretin *, P. Bocher

Mechanical Engineering Department, Ecole de technologie superieure (ETS), 1100 Notre-Dame Street, West, Montreal, QC H3C 1K3, Canada

ARTICLE INFO

ABSTRACT

Within polycrystals, significant stress concentrations can arise due to their heterogeneous nature. These stress intensities strongly influence the onset of nonlinear behaviors, such as plasticity and fatigue damage. One often overlooked source of heterogeneity is the crystal anisotropy and its resulting neighborhood effect. Previous research introduced a data-driven analytical model based on a cellular automaton (CA) to account for the neighborhood effect on a grain's stress level within an infinite aggregate under elastic conditions. It was demonstrated that, in some rare specific cases, grains could experience stress levels twice as high as the applied load. The current work extends the CA model by incorporating the effects of a free surface. Randomly oriented polycrystals under uniaxial loading were studied using a regular aggregate structure (Kelvin structure), where all grains are considered spherical and of identical size. Compared to full-field simulations, the extended CA model demonstrated an excellent capability to capture heterogeneities, even in cases where high stress concentrations are generated by the neighborhood. By leveraging the model's speed, a distribution function for grain stress levels was optimized to accurately capture the probability of extreme values. This allows for the estimation of the most likely highest stress within randomly oriented aggregates composed of billions of grains, along with its most probable localization relative to a free surface and the specific crystallographic configurations leading to it.

1. Introduction

Predicting and understanding material fatigue life has been a subject matter for several decades. The fatigue life of metallic parts can roughly be divided into three stages: crack initiation, short, and then long crack propagations [1]. For the latter, engineers often disregard the microstructural characteristics as the propagation is governed mainly by the crack geometry. When it comes to the first two stages, the material cannot be considered homogeneous anymore to accurately assess the mechanisms at stake. Due to the random nature of the heterogeneities, the fatigue life of the material may significantly vary for parts originating from the same bulk material, requiring a probabilistic approach to study the fatigue life of the material. As the crack initiation stage can represent 5 to 20% of the total life in the case of high cycle fatigue (HCF) [2–4] and up to 80% or more in the case of very high cycle fatigue (VHCF) [5–7], it is crucial to account for the material heterogeneities and their distribution probabilities.

HCF and VHCF regimes correspond to stress loads that remain below the macroscopic elastic limit of the material, resulting in small plastic deformation [8,9]. The plasticity is localized and depends on the local microstructure. It was shown experimentally [10,11] and numerically [12–14], in the case of high crystallin-elastic anisotropy, such as nickel or iron base alloys, that the onset of plasticity is strongly driven by the elastic regime. In the case of face-centered cubic polycrystals (fcc), full field simulations have revealed a correlation between grains' normalized resolved shear stress (NRSS) during the linear elastic regime and grains' plastic rate in HCF [15–17]. These studies indicate that grains with the highest NRSS values tend to plastify the fastest, making them more susceptible to crack initiation [18]. Thus, if one wants to predict material fatigue life at the microscale, the elastic stress fields and its heterogeneities must be accurately predicted in the first place.

Full-field simulations have shown that a grain's stress level is influenced as much by its own crystallographic orientation as by its neighboring grain orientations [16,19,20]. Polycrystals exhibit numerous sources of heterogeneities, including variations in grain morphology, distinct phases, crystallographic texture, surface effects, etc. However, all of these heterogeneities ultimately stem from a single trait: crystal anisotropy. The stiffness of the crystal varies depending on its orientation relative to the loading direction. Each grain generates stress-field variations specific to its orientation in its surroundings. Consequently,

---

* Corresponding author.
E-mail addresses: remy.bretin.1@ens.etsmtl.ca (R. Bretin), philippe.bocher@etsmtl.ca (P. Bocher).

the environment of a grain can strongly influence the mechanical state of the grain, a phenomenon known as the neighborhood effect. According to full-field simulations, the stress-field scattering can exhibit high amplitudes, with the highest observed grain mean stress being double or even higher than the lowest, depending on the material. Due to the random nature of crystallographic orientation distributions, a probabilistic approach must be employed to study polycrystal stress fields in the elastic regime.

Many numerical models exist to predict polycrystal stress fields, which can be roughly divided into two categories. The first category consists of full field models, such as the finite element method (FEM) [21–23] and models based on the fast Fourier transformation (FFT) [12,16,24]. These models can account for as many heterogeneities as needed and be as precise as the problem is described. The FEM does not have a description limit but is computationally intensive, whereas the FFT model incurs lower computational costs. However, FFT models are often restricted when it comes to geometry. They are usually limited to the study of periodic aggregates, excluding surface effect studies. Nonetheless, more recent works [25,26] made the study of heterogenous environment with a free-surface possible. On the other hand, there are analytical models, a.k.a. mean-field homogenization models, such as the widely known self-consistent approach and its derivatives [12,27–30] or the Maximum Entropy Method [31] which trades off the detail level for a much lower computational cost. Most of these models are capable of capturing the first and second moments of a grain's mean stress level. However, they may not accurately determine a grain mean stress level within a specific environment leading to significant stress concentrations and are unable to identify these particular configurations. In order to accurately predict the highest mean stress level within a polycrystalline mechanical part, a model should have the ability to capture a wide range of heterogeneities while keeping a low computational cost. This is essential to effectively evaluate an exceptionally large number of grain-neighborhood configurations. Unfortunately, none of the above models allow such study, one category being computationally expensive and the other considering limited microstructure specificities.

Alternatively, Bretin et al. [32] developed a data-driven cellular automaton (CA) model to predict grains' mean stress tensors within an infinite single phase polycrystalline aggregate. The model's distinct feature is its ability to consider the orientations of both, the grain in question and its neighboring grains, to predict its stress level. The model parameters are calibrated for a given material using the outcomes of a limited number of specific FEM simulations. Once fitted, the model becomes capable of predicting the mean stress tensors of thousands of grains within a fraction of a second, all the while considering the specific distribution of aggregate orientations. The model exhibits excellent accuracy in predicting the grains NRSS during the linear elastic regime when compared to converged FE predictions, even those with very high stress concentrations. The model's simplicity also enables the user to swiftly predict millions of grain-neighborhood configurations, generating ample data to accurately capture the most probable highest stress within an aggregate and the corresponding configurations leading to it.

However, one critical source of heterogeneity missing in the CA model is the impact of a free surface on the aggregate stress field. Experimental observations reveal distinct behaviors in crack initiation location between the High Cycle Fatigue (HCF) and Very High Cycle Fatigue (VHCF) regimes, with surface-initiated cracks prevalent in the former [33] and in-depth initiation more common in the latter [34,35]. Numerical simulations by various researchers have shown that the perturbation caused by a free surface typically extends to a depth of only 3 or 4 grains [36–38]. Additionally, surface grains tend to exhibit fewer active slip systems compared to in-depth grains, attributed to the more constraining neighborhood effect. Despite these surface-induced complexities, studies by Barbe et al. [37] underscore that the surface effect, while noteworthy, has a comparatively lesser impact

on grain stress levels than the neighborhood effect. Assuming ideal surface conditions (clean and polished) the sources of heterogeneities contributing to microscopic crack initiation and propagation can be narrowed down to persistent slip bands, plastic incompatibilities at boundaries, and elastic anisotropy. All of these sources can also be tied to grain crystallographic orientation, making the implementation of a free surface effect in the CA model straightforward, which would allow studying the location of stress concentrations within aggregates relatively to a free surface.

The present work aims to predict, within a standard-sized aggregate composed of millions of grains, the depth of the grain with the highest NRSS as well as the crystallographic configuration leading to it. According to VHCF experimental observations and assuming localized plasticity, grains with the highest NRSS are expected to be located in-depth. Other criteria, such as various critical resolved shear stresses (CRSS) or slip transmission and blockage due to grain boundary misorientation, which have been shown to be relevant factors in crack initiation for HCP crystal structures [39], will be disregarded here. All grains are considered to have the same CRSS level, and plasticity is assumed to be strictly localized within the grains. To address this problem, the CA model developed by Bretin et al. [32] has been extended to include the effect of a free surface on the stress field within the aggregate. Following a methodology similar to that in [32], the Kelvin structure is used to represent the polycrystals, avoiding any morphological or size effects. For this study, the crystal parameters of FCC stainless steel 316L were used as an example.

The article's outline unfolds as follows. In Section 2, the assumptions underlying the CA model are revisited, and the modifications made to account for the grains' depth are presented. Section 3 describes the FEM simulations of polycrystalline aggregates conducted to fit the model's parameters and evaluate its accuracy. A substantial volume of FEM data was generated and employed to precisely ascertain the accuracy of the CA model. In Section 4, this FEM data is utilized to provide an initial insight into the probability distribution of grains' stress as a function of their depth. Given that FEM alone is insufficient to generate enough data to accurately predict the highest stress within an aggregate comprising millions of grains, the CA model's speed was leveraged to generate a larger database. Section 5 details the methodology used to generate this database and how it was utilized to determine a grain's NRSS distribution function based on its depth. Using these functions, the most probable value of the highest NRSS within an aggregate composed of millions of grains was determined, along with its most probable localization relative to the free surface. From this data, grain-neighborhood crystallographic configurations leading to probable high-stress concentration were identified and simulated using FEM with an added viscoplastic behavior. Section 4.2 presents the results of these simulations to study the plastic behavior of these specific configurations and reestablish the correlation between a grain's NRSS and fatigue damage, while including the effects of a free surface. Finally, closing remarks and conclusions can be found in Section 6.

2. Definition of the CA model

2.1. Definition and approximation of the neighborhood effect proposed in [20]

The neighborhood effect of a specific neighborhood N on a grain g immersed in a crystalline aggregate under an uniform loading $\boldsymbol{E}$, noted $\Delta\boldsymbol{\varepsilon}_N^g$ and illustrated in Fig. 1(a), was defined in [20] as the difference between the mean strain tensor $\boldsymbol{\varepsilon}^g$ of the grain g immersed in the specific environment N and the mean strain tensor $\boldsymbol{\varepsilon}_0^g$ of the same grain g (same size, shape, orientation, and mechanical properties) immersed in a homogeneous environment with the material effective properties. Therefore, with this definition of the neighborhood effect, the mean strain and stress tensors of a grain g within a polycrystal can be defined as follows:
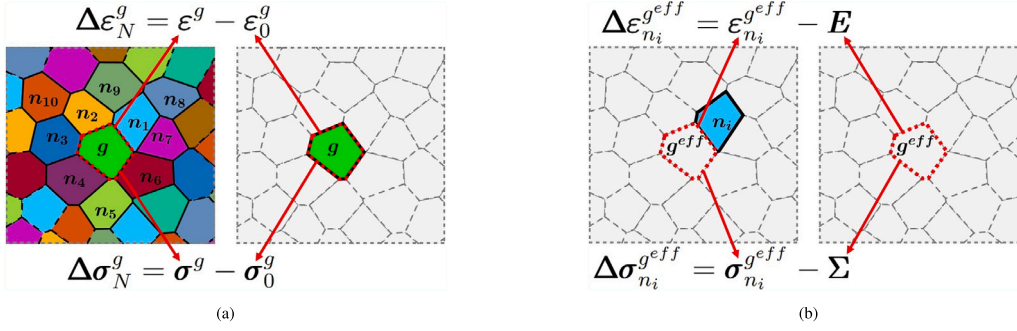
Fig. 1. 2D schematic representations of the neighborhood effect in aggregates under uniform loading (E/Σ). (a) A grain g's strain/stress deviation due to the neighborhood N ($\Delta\varepsilon_N^g/\Delta\sigma_N^g$) is quantified by the difference between grain g's strain/stress tensor in the polycrystalline aggregate (ε^g/σ^g) and grain g's strain/stress tensor in an infinite medium having the aggregate effective properties ($\varepsilon_0^g/\sigma_0^g$). (b) Within a homogeneous medium, a subset g^{eff}'s strain/stress deviation due to a neighboring grain n_i ($\Delta\varepsilon_{n_i}^{g^{eff}}/\Delta\sigma_{n_i}^{g^{eff}}$) is quantified by the difference between region g^{eff}'s strain/stress tensor in the homogeneous medium with the presence of the neighboring grain n_i ($\varepsilon_{n_i}^{g^{eff}}/\sigma_{n_i}^{g^{eff}}$) and the applied loading (E/Σ) [20].

$$\varepsilon^g = \varepsilon_0^g + \Delta\varepsilon_N^g \tag{1a}$$

$$\sigma^g = \sigma_0^g + \Delta\sigma_N^g = \mathbb{C}^g : \varepsilon^g \tag{1b}$$

where ':' denotes the double tensor contraction, $\mathbb{C}^g$ is the grain g elastic stiffness tensor and $\sigma_0^g = \mathbb{C}^g : \varepsilon_0^g$. Bretin et al. [32] studied the neighborhood effect within single phase polycrystals randomly oriented (meaning that the grains orientations were chosen randomly leading to isotropic effective properties of the material) where all grains were of identical shape and size. It was observed that for crystals with high elastic anisotropic, such as iron or nickel crystals, the neighborhood effect can increase or decrease the grain average strain but is in average null. It was also shown that for some specific configurations of environment, the amount of strain induced by the neighborhood effect can at least account for a half or more of the total strain ($\|\Delta\varepsilon_N^g\|_{eq} \geq \|\varepsilon_0^g\|_{eq}$).

In order to predict such strain concentration, Bretin et al. [20] proposed an approximation of $\Delta\varepsilon_N^g$ for the case of an infinite elastic polycrystalline aggregate: first $\Delta\varepsilon_N^g$ is approximated by the sum of each neighboring grain n_i's individual influence on the grain g, noted $\overline{gn_i}$, which, according to the definition of neighborhood effect proposed earlier, is the strain difference within the grain g immersed in a homogenized environment with and without the neighboring grain n_i. Then, the second approximation $\overline{gn_i} \approx \overline{g^{eff}n_i}$ is made, which consists of replacing the properties within the space occupied by the grain g with the effective ones. By combining these two approximations, we derive the subsequent approximation of the neighborhood effect, and consequently, the mean stress tensor for a grain g is obtained as follows:

$$\Delta\varepsilon_N^g \approx \sum_{n_i} \Delta\varepsilon_{n_i}^{g^{eff}} \quad \Rightarrow \quad \Delta\sigma_N^g \approx \mathbb{C}^g : \mathbb{C}^{eff^{-1}} : \sum_{n_i} \Delta\sigma_{n_i}^{g^{eff}} \tag{2a}$$

$$\sigma^g \approx \sigma_0^g + \mathbb{C}^g : \mathbb{C}^{eff^{-1}} : \sum_{n_i} \Delta\sigma_{n_i}^{g^{eff}} \tag{2b}$$

with $\quad \sigma^g = \mathbb{C}^g : \varepsilon^g \quad, \quad \sigma_0^g = \mathbb{C}^g : \varepsilon_0^g \quad$ and $\quad \Delta\sigma_{n_i}^{g^{eff}} = \mathbb{C}^{eff} : \Delta\varepsilon_{n_i}^{g^{eff}}$ (2c)

where $\mathbb{C}^g$ is grain g stiffness tensor, $\mathbb{C}^{eff}$ is the homogenized aggregate stiffness tensor attributed to the medium, and $\Delta\varepsilon_{n_i}^{g^{eff}}$ is the mean strain deviation within the volume of grain g^{eff} generated by the presence of the neighboring grain n_i within the homogeneous medium as illustrated in Fig. 1(b).

This approximation was tested using the FE method on single-phase aggregates with periodic boundary conditions where all grains were of identical size and shape using the Kelvin's structure as illustrated Fig. 2, with high elastic anisotropy (nickel, iron, and titanium crystals were tested) and accounting for the influence of all the neighboring grains within 3 grain's layers (258 neighboring grains n_i). The approximation has shown excellent accuracy, for various applied loadings, and even
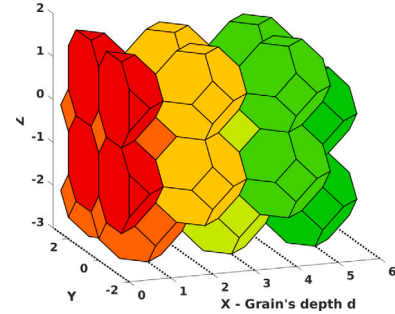

Fig. 2. Schematic illustration of the Kelvin's structure with the presence of a free surface.

for the specific configurations leading to high strain concentrations. It was also shown that the more grain g elastic properties differ from the effective properties, the less accurate this approximation is. This observation was negligible in the case of a single-phase material, but it might be significant in the case of a multiphase material with significant difference in elastic properties. Concerning the neighborhood influence, it was shown that the neighboring grains' influence decreases exponentially the farther they are from the central grain. A neighboring grain at three grain's radii from grain g will generate a $\Delta\varepsilon_{n_i}^{g^{eff}}$ much lower than a close grain, but all the neighboring grains located at 3 grains' radius put together, representing a bigger volume than the close grains, can have an influence on grain g as significant as the grains in close contact if they all have a crystallographic orientation favoring a stress concentration on it.

2.2. Definition of the new CA model accounting for the free surface

Two alterations from Eq. (2) were made to approximate a grain's mechanical state while accounting for the neighboring effect and the surface effect. The new governing equations of the CA model are now defined as follows:

$$\Delta\sigma_N^g(d) \approx \frac{1}{2}\left(\sum_{n_i} \Delta\sigma_{n_i}^{g^{eff}}(d) + \mathbb{C}^g : \mathbb{C}^{eff^{-1}} : \sum_{n_i} \Delta\sigma_{n_i}^{g^{eff}}(d)\right) \tag{3a}$$

$$\sigma^g(d) \approx \sigma_0^g(d) + \frac{1}{2}\left(\mathbb{I} + \mathbb{C}^g : \mathbb{C}^{eff^{-1}}\right) : \sum_{n_i} \Delta\sigma_{n_i}^{g^{eff}}(d) \tag{3b}$$

where $\mathbb{I}$ is the fourth order identity tensor and (d) refers to the dependency of the stress tensors on the depth of the grain g.

The first alteration concerns the neighborhood effect approximation presented in Eq. (2a). Bretin et al. [20] chose to use the strain variations

$\Delta \boldsymbol{\varepsilon}_{n_i}^{g^{eff}}$ to approximate the neighborhood effect, but another possible solution was instead to use the stress variations $\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}$ as follows:

$$\Delta \boldsymbol{\sigma}_N^g \approx \sum_{n_i} \Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}} \tag{4a}$$

$$\boldsymbol{\sigma}^g \approx \boldsymbol{\sigma}_0^g + \sum_{n_i} \Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}} \tag{4b}$$

It will be proven in Section 3.3.2, that the average of the two approximations Eqs. (2) and (4), resulting in Eq. (3) without the depth dependency, is a more accurate approximation of the grain mechanical state than Eq. (2) alone.

The second alteration is the insertion of the grain's depth dependency. The influence of the free surface and the grain g's depth, noted d, are taken into account by making $\boldsymbol{\sigma}_0^g$ and $\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}$ depth dependent. By introducing depth dependency to the stress variables, we assume that the neighborhood effect approximation derived from Eqs. (2) and (4) remains accurate, even in the presence of a free surface. This assumption is put to test in Section 3.3.2.

From Eq. (3), the prediction of $\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}(d)$ and $\boldsymbol{\sigma}_0^g(d)$ remain to be defined.

2.2.1. Prediction of $\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}(d)$

A method for predicting $\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}$ was introduced by Bretin et al. [32]: for a given applied loading, $\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}$ can be estimated through a linear regression involving the components of the grain n_i's stiffness tensor expressed in the global axis system. This relationship, expressed in Voigt tensor notation, can be noted as:

$$\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}} = \boldsymbol{U}^{\overrightarrow{gn_i}} \times \boldsymbol{X}^{n_i} \tag{5}$$

where $\boldsymbol{X}^{n_i}$ is a N_{cst}-list of independent components of the neighboring grain n_i stiffness tensor expressed in the global axis system, and $\boldsymbol{U}^{\overrightarrow{gn_i}}$ is a 6 by N_{cst} matrix of fitted parameters. These parameters are fitted by means of a multilinear regression using a set of values obtained through FEM for an adequately large number of random crystallographic orientations of grain n_i. The fitted parameters are specific to the relative position $\overrightarrow{gn_i}$, the submitted loading $\boldsymbol{E}$ and the medium properties $\mathbb{C}^{eff}$. There is one tensor $\boldsymbol{U}^{\overrightarrow{gn_i}}$ for each accounted neighboring grain n_i. $\boldsymbol{U}^{\overrightarrow{gn_i}}$ relies solely on geometric factors, such as the shapes of grains g and n_i, as well as their relative positions, while $\boldsymbol{X}^{n_i}$ is influenced by the orientation of grain n_i and its stiffness tensor. Due to the structure periodicity of the Kelvin structure, the same set of parameters $\boldsymbol{U}^{\overrightarrow{gn_i}}$ can be used for all grains and their neighbors.

The addition of a free surface affects the Kelvin's structure periodicity: only grains sharing the same depth have an identical environment geometrically. Therefore, the choice to keep the same equations to compute $\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}(d)$ was made with a small change: a set of parameters $\boldsymbol{U}^{\overrightarrow{gn_i}}$ need to be fitted for each grain g depth d up to a certain depth where the surface effect can be neglected. Eq. (5) can then be modified as:

$$\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}(d) = \boldsymbol{U}^{\overrightarrow{gn_i},d} \times \boldsymbol{X}^{n_i} \tag{6}$$

This formulation has the inconvenience to require much more FEM data to fit the parameters as one tensor $\boldsymbol{U}^{\overrightarrow{gn_i},d}$ has to be fitted for each grain-neighbor's relative position accounted and for each depth accounted. But, as it will be shown in Section 3.3, the use of an in-house FEM solver accelerates the process.

In the case of a cubic stiffness tensor, $N_{cst} = 10$ and thus the list $\boldsymbol{X}^{n_i}$ is composed of the 10 following components:

$$\boldsymbol{X}^{n_i} = \left[\mathbb{C}_{1111}^{n_i}, \mathbb{C}_{1131}^{n_i}, \mathbb{C}_{2212}^{n_i}, \mathbb{C}_{3323}^{n_i}, \mathbb{C}_{2323}^{n_i}, \mathbb{C}_{2331}^{n_i}, \mathbb{C}_{2312}^{n_i}, \mathbb{C}_{3131}^{n_i}, \mathbb{C}_{3112}^{n_i}, \mathbb{C}_{1212}^{n_i} \right] \tag{7}$$

where $\mathbb{C}^{n_i}$ is the grain n_i stiffness tensor expressed in the global axis system and thus depending on the grain crystallographic orientation. All the other components of $\mathbb{C}^{n_i}$ can be expressed as a multilinear function

of these 10 components, independently of the grain orientation. The dependency of $\mathbb{C}^{n_i}$ components was found using the QR decomposition on a matrix 81 by 100, where each column is the 81 components of $\mathbb{C}^{n_i}$ for 100 random crystallographic orientations.

2.2.2. Prediction of $\boldsymbol{\sigma}_0^g(d)$

The Eshelby's equivalent inclusion method was used by Bretin et al. [32] to predict $\boldsymbol{\sigma}_0^g$. A modified version of this method accounting for a free surface could have been used to predict $\boldsymbol{\sigma}_0^g(d)$, but a simpler solution is proposed in the present work. In the same way as $\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}(d)$, $\boldsymbol{\sigma}_0^g(d)$ can be predicted as a multi-linear regression of $\mathbb{C}^g$ components as:

$$\boldsymbol{\sigma}_0^g(d) = \boldsymbol{U}^{\overrightarrow{0},d} \times \boldsymbol{X}^g \tag{8}$$

where is $\boldsymbol{X}^g$ is a list of the same 10 specific components from Eq. (7) of $\mathbb{C}^g$. The FEM simulations generating the set of values used to fit the tensors $\boldsymbol{U}^{\overrightarrow{gn_i},d}$ also provide a set of values of $\boldsymbol{\sigma}_0^g(d)$ that can be used to fit the tensors $\boldsymbol{U}^{\overrightarrow{0},d}$. The accuracy of this model will be discussed in Section 3.3.1.

2.2.3. Accounted neighboring grains

The influence of a neighboring grain n_i was shown by Bretin et al. [20] to exponentially decrease with the distance from the central grain g. Therefore, only the stress deviation $\Delta \boldsymbol{\sigma}_{n_i}^{g^{eff}}(d)$ from the three layers of neighboring grains were accounted in the summation from Eq. (3). In the present work, four layers were considered to capture the influence of a wider neighborhood. In the Kelvin's structure, a grain has 14 grains in its first neighboring grains layer, 50 in the second, 194 in the third and 278 in the fourth, making a total of 536 grains accounted in the neighborhood effect. Due to the presence of the free surface, grains located at the surface ($d = 0$) see these numbers shortened to 9, 29, 105, 149 and 292, respectively.

3. Validation of the CA model accuracy by means of the FE

3.1. Variables used to estimate model's accuracy

The CA model accuracy will be measured on how well it can predict a grain mean stress tensor $\boldsymbol{\sigma}^g$ and the resulting resolved shear stress τ_s^g in comparison to a reference model, the FE method:

- To quantify the distance between the stress tensors predicted by the FE and CA models, the equivalent von Mises stress, denoted as $\| \bullet \|_{eq}$, serves as a tensorial norm.
- The normalized resolved shear stress (NRSS), noted τ_s^{*g}, is a variable that is used to gauge a grain's chances to undergo plastic deformation, i.e., to potentially initiate a crack. In an homogeneous environment (meaning that $\boldsymbol{\sigma}^g = \boldsymbol{\Sigma}$), the NRSS would be equal to the Schmid's factor, but due to the stress concentration that can occur, the NRSS can reach much higher values than the Schmid's factor. It was shown by Bretin et al. [17] that a grain cyclic plastic strain is proportional to its highest NRSS, making the NRSS a good indicator to where damage will occur. A grain g's NRSS for its slip system s is defined as the slip system RSS normalized by the applied load in the linear elastic range (for which no visco-plasticity has occurred yet):

$$\tau_s^{*g} = \frac{\tau_s^g}{\| \boldsymbol{\Sigma} \|_{eq}} \tag{9}$$

where τ_s^g is the resolved shear stress (RSS) of grain g slip system s, computed as:

$$\tau_s^g = \boldsymbol{\sigma}^g : \boldsymbol{m}_s^g = (\mathbb{C}^g : \boldsymbol{\varepsilon}^g) : \boldsymbol{m}_s^g \tag{10}$$

were $\boldsymbol{m}_s^g$ is the orientation tensor of the grain g's slip system s defined as:

$$\boldsymbol{m}_s^g = \frac{1}{2} \left(\vec{\boldsymbol{u}}_s^g \otimes \vec{\boldsymbol{v}}_s^g + \vec{\boldsymbol{v}}_s^g \otimes \vec{\boldsymbol{u}}_s^g \right) \tag{11}$$

where $\otimes$ denotes the tensor product, $\vec{v}_s^g$ is the normal to the slip plane and $\vec{u}_s^g$ is the slip direction.

The highest NRSS among all the grain g's slip systems will be noted τ^{*g}:

$$\tau^{*g} = max(\ |\tau_s^{*g}| \ )_s \tag{12}$$

With such a definition, within the elastic range, a grain with the highest NRSS within an aggregate would start to plastify for an applied load of $\Sigma = min(crss_s^g \ / \ \tau_s^{*g})_s$, where $crss_s^g$ is the grain g critical resolved shear stress at which the slip system s would start yielding.

3.2. Finite element model definitions

Following the same methodology as Bretin et al. [20], static elastic finite element simulations were performed on polycrystalline aggregates with the addition of a free surface. These simulations' results are later used to fit the CA model's tensors $U^{0,d}$ and $U^{\overline{gn_i},d}$, evaluate its accuracy and make a preliminary study of the surface effect.

Two kinds of FEM aggregates were generated:

- **S-1G**: A Kelvin structured aggregate constituted of one central grain and its 4 layers of neighboring grains are immersed in an homogeneous cubic matrix located at a depth d from a free surface as illustrated in Fig. 3(a). The crystallographic properties and a specific orientation are attributed only to the central grain (in red on Fig. 3(a)), and the effective properties are attributed to the rest of the mesh. The S-1G aggregates serve two purposes: either the central grain is considered as the grain g, providing the value $\sigma_0^g(d)$ (Fig. 1(a)), which is later used to fit the tensors $U^{0,d}$. Or, the central grain is considered as a neighboring grain n_i, and the values $\sigma_{n_i}^{g^{eff}}(d)$ can be extracted from all the neighboring grains considered as g^{eff} (Fig. 1(b)). This leads to $\Delta\sigma_{n_i}^{g^{eff}}(d) = \sigma_{n_i}^{g^{eff}}(d) - \Sigma$, which is later used to fit the tensors $U^{\overline{gn},d}$. Various central grain depths d ranging from 0 to 9, as well as $d = C/2 \approx \infty$, were meshed, where a Kelvin cell radius serves as the unit of length. For each depth, 100 random crystallographic orientations for the central grain were generated, making a total of 1100 different aggregates studied.
- **S-Agg**: A Kelvin structured aggregate constituted of 1395 grains is immersed in a homogeneous cubic matrix as illustrated in Fig. 3(b). The crystallographic properties and a specific orientation are attributed to all the grains, and the effective properties are attributed to the remaining mesh. In these 1395 grains, only 65 have their full 4 layers of neighboring grains, and only 10 have their full 5 layers of neighboring grains (one for each depth from 0 to 9). The crystallographic orientations were distributed within the aggregate in three different ways:

  - **S-Agg-1**: orientations are distributed randomly. A total of 400 aggregates with random distributions were generated.
  - **S-Agg-2**: Using the CA model, the NRSS of 10^8 grain-neighborhood randomly oriented located at a given depth d were evaluated (more details on the methodology in Section 5.2). The configuration of orientations leading to the highest NRSS predicted was saved and distributed to the aggregate central grain located at the given depth d and its neighboring grains, as illustrated in Fig. 3(b). All the remaining grains from the aggregate not being accounted for in the CA prediction of neighborhood effect are randomly oriented. One hundred of such configurations were identified by repeating this process for the depths $d = \{0, 1, 2, 9\}$, making a total of 400 S-Agg-2 aggregates.

  - **S-Agg-3**: Using the CA model, the grain-neighborhood orientations configurations leading to the highest possible NRSS value for a given depth d were identified. Taking advantage of the CA model decomposition of the neighborhood effect, for each neighboring grain n_i starting from the closest to the farthest, the crystallographic orientation leading to the highest increase of the grain NRSS is attributed to that neighboring grain. These orientations are then distributed accordingly to the corresponding grains from the aggregate mesh. All the grains not being accounted for in the CA prediction of neighborhood effect (namely, all the grains farther than 4 grain layers from the central grain) are randomly oriented. Four S-Agg-3 aggregates were generated, one for each depth $d = \{0, 1, 2, 9\}$.

A total of 804 S-Agg aggregates were generated. For each aggregate, only the resulting mean stress tensors of the 65 central grains with their 4 complete layers of neighboring grains (bright-colored in Fig. 3(b)) are retained and will be used later for comparison between the FE and CA predictions.

For both types of aggregates, linear-elastic simulations are conducted with the following boundary conditions applied to the cubic matrix:

- Nodes at the bottom of the cube ($z = -C/2$) have their displacement along Z-axis set to 0.
- Nodes at the top of the cube ($z = C/2$) have their displacement along Z-axis set to $C \times L$, C being the cube's size.
- The node located at the bottom-left ($x = C/2$, $y = C/2$, $z = -C/2$) is pinned, meaning that its displacement is set to 0 in all directions.

Such boundary conditions yields to a uni-axial stress load Σ along Z-axis such that all its components equal 0, except Σ_{33} such that:

$$\Sigma_{33} = E_y^{eff} \times L = 19.64 \quad [\text{MPa}] \tag{13}$$

where E_y^{eff} is the material effective Young's modulus, and L is the strain amplitude equal to 0.01%. A uniaxial stress load was chosen to demonstrate the accuracy and applications of the CA model; however, other types of loading, such as tension-torsion applied to the mesh boundaries while maintaining the free surface, could also be used. Additionally, the CA model parameters are fitted to a specific applied load. Given the linearity of the problem, as demonstrated in [32], any linear combination of the fitted loads would yield the same linear combination of the fitted parameters.

The size of the cubic matrix $C = 100$ (1 being a Kelvin's cell radius) in which the aggregates are immersed was chosen large enough to avoid any border effect (except for the intended free surface) on the grains' stress field. This way, the aggregate's volume fraction represents less than 0.03% of the whole mesh volume. After conducting a convergence study on the grains' mean stress tensors, as displayed in Appendix A, the Kelvin structure is meshed with an average of 1284 tetrahedral elements per cell with quadratic interpolation. The element size in the matrix increases farther from the aggregate, reaching a maximum size of 10 elements per cube's edge. The average number of elements per grain used is nearly twice as fine as the mesh density used by Bretin et al. [20] for similar simulations. This finer mesh is employed to capture not only the grain mean stress tensor but also the standard deviation within a grain, as it will be used in Section 4. For each FEM simulation, a grain mean stress tensor σ^g is obtained by averaging the grain elements' strain tensors at the centroid, weighted by their volume fraction.

An in-house FEM code was used to perform these simulations. The code was made using CUDA-FORTRAN, taking advantage of graphics processing unit (GPU) parallel-computing. The code was validated by comparing it to the results obtained with the commonly known software ABAQUS for one of the S-Agg simulations using the exact
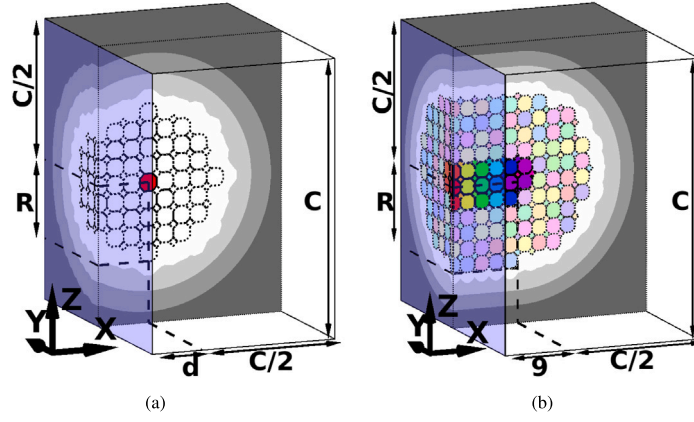
Fig. 3. Schematic illustrations of the different meshes used for the FEM simulations: (a) 2D cut of a grain (in red) and its four layers of neighboring grains (delimited by dotted lines) immersed in a homogeneous matrix at a depth d varying from 0 to 9 or equaled to $C/2$ for the in-depth case (d is set to 6 in this illustration). Only the central grain has the crystallographic properties, while both the matrix and neighboring grains share the effective material properties. $R = 4 \times 2 + 1 = 9$, $C = 100$; (b) 2D cut of 10 grains scattered at different depth from 0 to 9 and their five layers of neighboring grains immersed in a homogeneous matrix. All colored grains have crystallographic properties. All bright-colored grains (total of 65 grains) have their 4 layers of neighboring grains complete. $R = 5 \times 2 + 1 = 11$, $C = 100$; In both illustrations, the free surface is represented by the blue transparent plan and the grayer the matrix, the coarser the mesh.

same mesh, boundary conditions and material properties. By comparing the two codes results at each integration point, the maximum value of $err_\sigma = \|\boldsymbol{\sigma}_{tested}^g - \boldsymbol{\sigma}_{ref}^g\|_{eq}/\|\boldsymbol{\Sigma}\|_{eq}$ observed did not get higher than 10^{-6} which is very negligible. Using an Nvidia RTX3090 GPU, it takes between 15 s. for $D = 0$ (187,000 elements) and 26 s. for $D = \infty$ (373,000 elements) to perform one *S-1G* simulation, and 65 s. (634,000 elements) for one *S-Agg* simulation.

The material chosen for these simulations and the rest of the article is the 316L austenitic stainless steel (face centered cubic structure) which the material parameters were taken from Guilhem's work [38]. This material shows a significant crystal elastic anisotropy, which is great to put to test the CA model. Its crystal structure is face-centered cubic, and only the 12 octahedral slip systems are considered for the calculation of the NRSS. The material parameters are listed in Appendix B.

3.3. Comparison between FE and CA models predictions

3.3.1. Prediction of $\Delta\boldsymbol{\sigma}_{n_i}^{g^{eff}}(d)$ and $\boldsymbol{\sigma}_0^g(d)$

Before evaluating the CA model's accuracy in predicting the grain mean stress tensor $\boldsymbol{\sigma}^g$, the accuracy of its components used to calculate it, $\Delta\boldsymbol{\sigma}_{n_i}^{g^{eff}}(d)$ and $\boldsymbol{\sigma}_0^g(d)$, needs to be assessed. In the following, to standardize the notations, $\boldsymbol{\sigma}_0^g$ will be denoted as $\Delta\boldsymbol{\sigma}_{n_0}^{g^{eff}}$, meaning that n_0 refers to the observed central grain g. Also, $\Delta\boldsymbol{\sigma}_{n_i\ FE}^{g^{eff}}$ and $\Delta\boldsymbol{\sigma}_{n_i\ CA}^{g^{eff}}$ will denote the predictions obtained with the FE model and the linear regression from Eqs. (6) and (8), respectively.

For each relative position $\overrightarrow{gn_i}$ and each grain g depth, 100 values of $\Delta\boldsymbol{\sigma}_{n_i}^{g^{eff}}(d)$ are extracted from *S-1G* simulations, with one value corresponding to each random orientation of the central grain. Within these 100 values, a random subset of 40 values is used for the multilinear regression to fit the tensor $\boldsymbol{U}^{\overline{gn_i},d}$ and $\boldsymbol{U}^{\vec{0},d}$, while the other 60 are used to validate the prediction accuracy.

Instead of relying on the R-squared statistic, which is inadequate for evaluating the prediction error of a stress tensor, different statistical tool was employed to calculate the differences in predictions. The variable used to assess the prediction differences between the two models is defined as follows:

$$\Delta\boldsymbol{\sigma}_{n_i}^{diff} = \frac{\|\Delta\boldsymbol{\sigma}_{n_i\ FE}^{g^{eff}} - \Delta\boldsymbol{\sigma}_{n_i\ CA}^{g^{eff}}\|_{eq}}{\langle\|\Delta\boldsymbol{\sigma}_{n_i\ FE}^{g^{eff}}\|_{eq}\rangle_{ori_{n_i}}} \tag{14}$$

where $\langle\bullet\rangle_{\overline{gn_i}}$ is the average of the 60 values obtained for each orientation attributed to the grain n_i.

In Table 1 are presented the statistics of $\Delta\boldsymbol{\sigma}_{n_i}^{diff}$ per neighboring grains layers and grain n_i's depth. In the table, $\langle\Delta\boldsymbol{\sigma}_{n_i}^{diff}\rangle_{\overline{SEL}}$ designs the average over all the values of $\Delta\boldsymbol{\sigma}_{n_i}^{diff}$ obtained for each orientation, each selected relative position $\overrightarrow{gn_i}$ and each selected depth. Similarly, $std(\bullet)_{\overline{SEL}}$ and $max(\bullet)_{\overline{SEL}}$ are their standard deviation and maximum, respectively. Two points can be drawn out of these data.

First, apart from the case where the central grain is located on the free surface (in that case the central grain is a half Kelvin's cell), the predictions errors are remotely constant no matter the grain n_i's depth and distance from the grain g. An average prediction error of $\boldsymbol{\sigma}_0^g{}_{CA}$ lower than 0.3% is observed, with a standard deviation lower than 0.1% and a maximum error lower than 0.5%. This high precision of $\boldsymbol{\sigma}_0^g{}_{CA}$ predictions is relevant as it represents in average more than half of a grain total stress (Eq. (1b)). An average prediction error of $\Delta\boldsymbol{\sigma}_{n_i}^g{}_{CA}$ of ~1% is observed, with a standard deviation lower than 0.1% and a maximum error lower than 4%. These data are also very satisfying, showing the excellent accuracy of the model.

The second point concerns grains located on the free surface (depth equals 0) as a loss in accuracy is observed. The average prediction error of $\boldsymbol{\sigma}_0^g{}_{CA}$ is 1%, which is 3 times larger than the error observed in depth. Similarly, the standard deviation and maximum errors are up to 0.3% and 1.9%, respectively. These values are still significantly small, which is great, knowing the importance of predicting $\boldsymbol{\sigma}_0^g{}_{CA}$ accurately. On the same note, $\Delta\boldsymbol{\sigma}_{n_i}^g{}_{CA}$ also shows similar trends: the average prediction error goes up to 4.6% when ~1.1% was observed for the other depth. Similar ratios are also observed for all the other relative positions and other statistical values. Nonetheless, these values remain low and the accuracy of the model excellent.

Concerning the identification of the tensors $\boldsymbol{U}^{\vec{0},d}$ and $\boldsymbol{U}^{\overline{gn},d}$, the number of FE simulations required to fit their components could have been reduced, but due to the speed of the FE software used, it was not necessary to look for an optimization of the process, but one could imagine a smaller specific set of crystallographic orientations leading to the same precision and reducing the amount of FE simulations.

3.3.2. Prediction of $\boldsymbol{\sigma}^g(d)$ and $\tau^{*g}(d)$

In this section, the CA model predictions of a grain's stress tensor and highest NRSS for the *S-Agg* aggregates are confronted to those obtained from the FE model. To highlight the importance of each improvement brought to the CA model (new approximation formulation, depth dependency, 4 layers of neighbors accounted), five different variants of the CA model were tested, where each variant has some or all of the upgrades mentioned in Section 2.2. The 65×804 predictions obtained with each FE and CA models are compared in Table 2:

Table 1

Statistics of $\boldsymbol{\Delta\sigma}_{n_i}^{diff}$ in % depending on n_i's depth and distance from the grain g. $\boldsymbol{\Delta\sigma}_{n_i}^{diff}$ is used to assess $\boldsymbol{\sigma}_0^g(d)$ and $\boldsymbol{\Delta\sigma}^{g^{eff}}$ predictions differences between the CA and FE models.

[%]	Relative positions	Grain n_i's depth						
	$\overline{gn_i}$ considered	0	1	2	3	4	5 to 9	∞
Average	$\boldsymbol{\sigma}_0^g\ (n_i = n_0)$	1.0	0.3	0.2	0.2	0.2	0.2	0.2
	$0 < \|\overline{gn_i}\| \leq 2$	4.6	1.4	1.1	1.0	1.0	1.0	1.0
	$2 < \|\overline{gn_i}\| \leq 4$	4.2	1.1	0.9	0.8	0.8	0.8	0.8
$\langle \boldsymbol{\Delta\sigma}_{n_i}^{diff} \rangle_{\overline{SEL}}$	$4 < \|\overline{gn_i}\| \leq 6$	3.9	1.0	0.9	0.8	0.7	0.7	0.7
	$6 < \|\overline{gn_i}\| \leq 8$	3.8	1.0	0.9	0.8	0.8	0.7	0.7
Standard deviation	$\boldsymbol{\sigma}_0^g\ (n_i = n_0)$	0.3	0.1	0.1	0.0	0.0	0.0	0.0
	$0 < \|\overline{gn_i}\| \leq 2$	2.4	0.5	0.4	0.3	0.3	0.3	0.3
	$2 < \|\overline{gn_i}\| \leq 4$	2.3	0.6	0.4	0.3	0.3	0.3	0.3
$std\left(\boldsymbol{\Delta\sigma}_{n_i}^{diff}\right)_{\overline{SEL}}$	$4 < \|\overline{gn_i}\| \leq 6$	2.1	0.5	0.4	0.3	0.3	0.3	0.3
	$6 < \|\overline{gn_i}\| \leq 8$	2.1	0.5	0.4	0.3	0.3	0.3	0.3
Maximum	$\boldsymbol{\sigma}_0^g\ (n_i = n_0)$	1.9	0.5	0.4	0.3	0.3	0.3	0.3
	$0 < \|\overline{gn_i}\| \leq 2$	14.6	3.6	2.5	2.2	2.3	2.3	2.3
	$2 < \|\overline{gn_i}\| \leq 4$	13.9	3.8	2.5	2.1	2.1	2.1	2.1
$max\left(\boldsymbol{\Delta\sigma}_{n_i}^{diff}\right)_{\overline{SEL}}$	$4 < \|\overline{gn_i}\| \leq 6$	12.7	3.9	2.6	2.2	2.2	2.2	2.3
	$6 < \|\overline{gn_i}\| \leq 8$	12.8	3.5	2.5	2.4	2.3	2.4	2.4

Table 2

Comparison between the FE model and several variants of the CA model. *S-Agg* simulations results from both models are used to assess the CA model accuracy to predict a grain's stress tensor and NRSS, depending on the model's version and the grain's depth. Different variants of the CA model are presented to highlight the importance of each improvement brought to the CA model.

Average	CA model's	Grain's depth d										All depths
	version	0	1	2	3	4	5	6	7	8	≤ 9	included
Stress error	NEW4	4.1	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.3	3.2	3.4
	OLD4	6.7	5.6	5.4	5.2	5.2	5.2	5.2	5.2	5.1	5.1	5.4
$\frac{\|\boldsymbol{\sigma}_{FE}^g - \boldsymbol{\sigma}_{CA}^g\|_{eq}}{\|\boldsymbol{\sigma}_{FE}^g\|_{eq}}$	D∞	13.3	8.1	5.3	4.2	3.8	3.6	3.4	3.4	3.3	3.2	5.0
	NEW3	4.5	3.7	3.7	3.8	3.8	3.8	3.8	3.9	3.8	3.8	3.9
[%]	D01	8.5	10.6	⊘	⊘	⊘	⊘	⊘	⊘	⊘	⊘	9.6
	SC	16.4	14.9	15.6	15.4	15.4	15.3	15.3	15.3	15.1	15.3	15.4
NRSS error	NEW4	2.0	1.6	1.6	1.5	1.6	1.6	1.6	1.6	1.5	1.5	1.6
	OLD4	2.9	2.5	2.4	2.3	2.4	2.3	2.3	2.3	2.3	2.3	2.4
$\frac{|\tau_{FE}^{*g} - \tau_{CA}^{*g}|}{\tau_{FE}^{*g}}$	D∞	5.2	3.2	2.4	2.0	1.8	1.7	1.7	1.6	1.6	1.5	2.2
	NEW3	2.3	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.8	1.9
[%]	D01	4.4	4.9	⊘	⊘	⊘	⊘	⊘	⊘	⊘	⊘	4.7
	SC	8.8	7.7	7.6	6.8	6.6	6.8	6.8	6.7	6.9	7.0	7.2

- The NEW4 version, based in Eq. (3) , includes the 3 mentioned improvements and shows the best accuracy among all the CA model's variants, with an average stress prediction error of 3.4% and an average NRSS prediction error of 1.6%. These results are very satisfying, however, further details regarding this matter will be provided later.

- OLD4 accounts for 4 layers of neighboring grains and is depth dependent, but Eq. (2b) is used to approximate the neighborhood effect instead of Eq. (3). The prediction errors increased by a factor 1.5, showing that the new approximation of the neighborhood effect (Eq. (3)) is much more accurate than the initial one proposed in [32] (Eq. (2b)). This is even more relevant for the case $d = 0$ where the average prediction error is increased by a factor 1.6.

- D∞ uses Eq. (3) approximation and accounts for 4 layers of neighboring grains, but all grains are considered in deep-depth ($d = \infty$). Comparing D∞ to NEW4 shows that the free surface affects only the grains less deep than $d = 5$. A much more significant accuracy drops is observed for D∞ in comparison to NEW4 for the grains located at the surface ($d = 0$ or 1).

- NEW3 uses Eq. (3) approximation and is depth dependent, but only accounts for 3 layers of neighboring grains. Comparing NEW3 to NEW4 shows that reducing the radius of accounted neighboring grains has decreased the accuracy in a non-negligible amount, going from an average stress prediction error of 3.4% to 3.9% and an average NRSS prediction error of 1.6% to 1.9%. This

indicates that increasing the range of neighboring grains considered in the model could be a potential solution for improving its accuracy, up to a certain point.

- D01 includes all upgrades but considers only the grains visible from the surface (grains located at $d = 0$ and $d = 1$) for the calculation of the neighborhood effect. The difference in precision between the model versions NEW4 and D01 underscores the risk of relying solely on 2D analysis based on EBSD surface images. Considering only the visible grains leads to an average error of 9.6% for the stress and 4.7% for the NRSS. These errors increase to 11.1% and 10.5%, respectively, when considering only the grains with high stress concentration from *S-Agg-2* aggregates. This demonstrates the critical importance of accounting for the full grain environment when performing stress field analysis, rather than relying only on surface observations.

- SC refers to the commonly used Self Consistent scheme, which consists of disregarding the neighborhood effect in the CA model ($\boldsymbol{\sigma}_{SC}^g(d) = \boldsymbol{\sigma}_0^g(d)$). Comparing the SC and FE models is in fact equivalent to quantifying the neighborhood effect $\boldsymbol{\Delta\sigma}_N^g(d)$. The neighborhood effect represents on average 15.4% of a grain total stress, with a 90th percentile value of 23.7% and a maximum of 69.9% observed for the *S-Agg-3* aggregates. These values show how important it is to accurately capture the neighborhood effect when the environment is specifically set to generate stress concentrations, as it is in *S-Agg-2* and *S-Agg-3* aggregates.
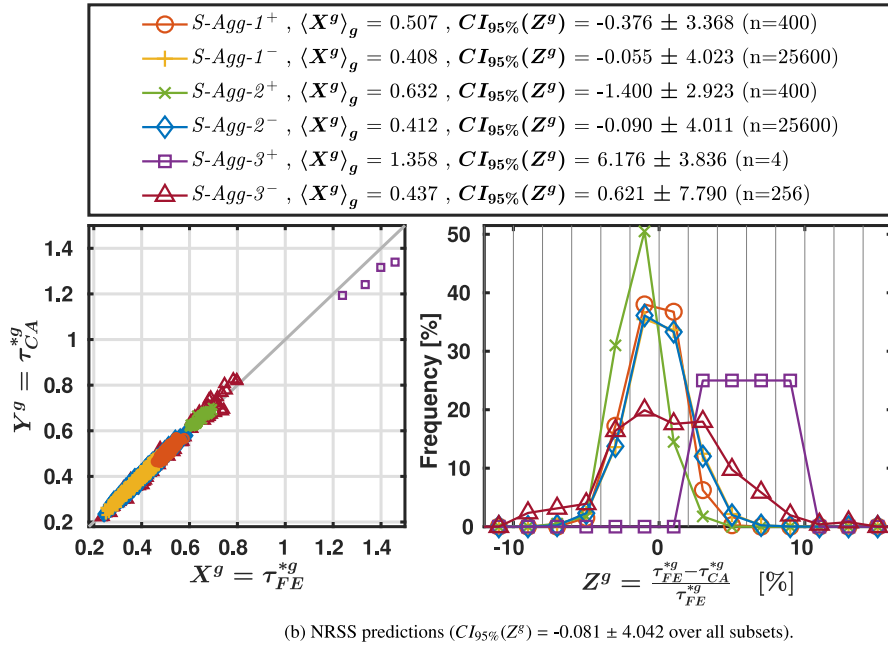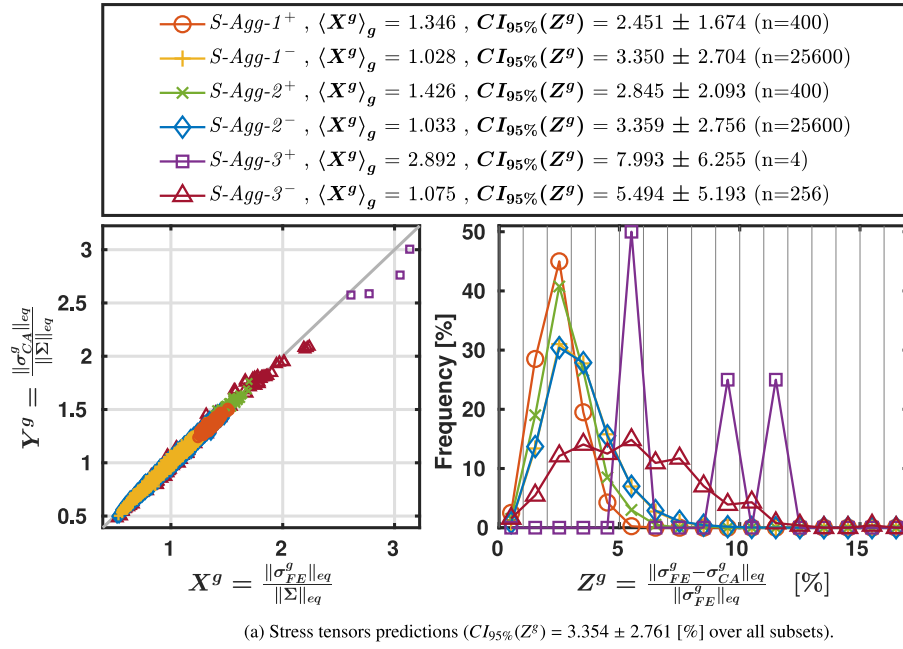
(a) Stress tensors predictions ($CI_{95\%}(Z^g) = 3.354 \pm 2.761$ [%] over all subsets).

(b) NRSS predictions ($CI_{95\%}(Z^g) = -0.081 \pm 4.042$ over all subsets).

Fig. 4. Scatter plot of FE and CA-NEW4 model predictions, along with the distribution of the model differences across the 65×804 values divided into various subsets. The variable n refers to the number of grains included in the subset. The exponent sign $\bullet^+$ refers to the predictions of the grains showing the highest X value among the 65 grains considered within an aggregate. $\bullet^-$ refers to all the remaining grains. Thus, *S-Agg-3$^+$* counts only 4 points.

- Regardless of the model under consideration, surface grains (d = 0) consistently exhibit a prediction accuracy noticeably lower than that of the other grains.

Each modification brought to the model has improved the accuracy of the CA model when compared to the FEM. To provide a thorough assessment of the model NEW4's accuracy, Fig. 4 displays the predictions of σ^g and τ^{*g} from both models, along with the distributions of their prediction differences. Data are divided into various subsets to apprehend the models' differences depending on the grain-neighborhood configuration. *S-Agg$^+$* refers to the grains showing the highest value among the 65 grains considered within an aggregate for each *S-Agg* simulation. On the figure is also displayed the 95% confidence interval calculated using the empirical rule as follows:

$$CI_{95\%}(Z^g) = \langle Z^g \rangle_g \pm 2 * std(Z^g)_g \tag{15}$$

This formulation assumes that Z^g follows a normal distribution, which is a reasonable assumption for our data. Several points can be noticed from these data:

- The two models' predictions exhibit remarkable consistency, with all data points closely aligning along the $X = Y$ axis. Over all the data, the following 95% confidence intervals of the two models differences are obtained:

  - $CI_{95\%}(\frac{\|\sigma^g_{FE} - \sigma^g_{CA}\|_{eq}}{\|\sigma^g_{FE}\|_{eq}}) = 3.354 \pm 2.761$ [%]

  - $CI_{95\%}(\frac{\tau^{*g}_{FE} - \tau^{*g}_{CA}}{\tau^{*g}_{FE}}) = -0.081 \pm 4.042$ [%]

– $CI_{95\%}(\tau_{FE}^{*g} - \tau_{CA}^{*g}) = 0.000 \pm 0.016$

- Except for *S-Agg-3*, there is no notable decrease in accuracy observed between the *S-Agg+* and *S-Agg−* subsets, despite the significant differences in stress levels among their grains. In absolute terms, a slight decrease of accuracy is observed for τ_{CA}^{*g} in Fig. 4(b) from *S-Agg−* to *S-Agg+* subsets. However, proportionally to the predicted grain value, the trend is reversed, similar to σ_{CA}^{g} in Fig. 4(a).

- *S-Agg-3+* exhibits extremely high stress levels. Nonetheless, the accuracy of predictions for these extreme configurations remains relatively similar to that of other predictions. Their values of τ^{*g} are twice as high as those of *S-Agg-1+* and even *S-Agg-2+*, which are the maximum observed after millions of random grain-neighborhood generations. This demonstrates that when the entire grain neighborhood is perfectly oriented to generate stress concentration, extremely high stress levels can be attained. In the case of textured material, these stress levels might be more achievable than in the case of fully random aggregates.

- The CA model predictions for *S-Agg-2+* are less accurate than those for *S-Agg-1+*. τ^{*g} are in average overestimated by the CA model for *S-Agg-2+* in comparison to the FE model, whereas an average closer to zero is observed for *S-Agg-1+*. This could be explained by the difference of stress level between the two subsets ($\langle \tau_{FE}^{*g} \rangle_g = 0.507$ for *S-Agg-1+* versus 0.632 for *S-Agg-2+*), suggesting that the CA model tends to overestimate grains exhibiting high stress levels. However, this hypothesis is not corroborated by the data from the *S-Agg-3+* subset, where the CA model underestimates the stress levels. Nonetheless, this subset is not large enough to draw conclusive results. Another reason for this discrepancy could be the methodology used to generate *S-Agg-2* aggregates. They are generated using the CA model, by keeping the grain-neighborhood configurations that exhibit the highest NRSS among 10^8 random configurations. Consequently, by selecting the highest predicted value, the likelihood of $\tau_{CA}^{*g} > \tau_{FE}^{*g}$ is increased, leading to skewed statistics. The best way to confirm these hypotheses would be to also simulate 10^8 aggregates randomly generated using the FE model. However, undertaking such a large scale of simulations would require a significant amount of time and resources. Therefore, the necessity of models such as the CA model becomes apparent, as they provide a more efficient means of generating large datasets for analysis and comparison.

- Grains located at the surface (d = 0) exhibit lower prediction accuracy compared to other grains, as previously observed in Table 2. Similar to Fig. 4(b), different subsets of data are presented in Fig. 5: predictions from *S-Agg-1+* and *S-Agg-2+* subsets are further divided based on the grains' depth. In the case of the *S-Agg-2+* subset, when the stress concentration is located at d = 1, 2 and 9, the confidence intervals are constant. However, when d = 0, the interval width is multiplied by $0.024/0.016 \approx 1.5$. This trend is less pronounced in the case of the *S-Agg-1+* subset. The explanation mentioned above could also be applied to this discrepancy.

The primary objective of the CA model is to accurately predict the NRSS of grains, enabling the identification of grain-neighborhood configurations that lead to high stress concentration within very large-scale aggregates. This information is crucial for pinpointing the grains most likely to initiate crack and determining their associated stress levels. Therefore, the confidence intervals illustrated in Fig. 5 are the most relevant to the problematic. Assuming that the intervals from *S-Agg-2+* subsets are skewed due to the explanation mentioned above, the intervals from *S-Agg-1+* would be the most accurate and adapted to the problematic for assessing the model's precision in comparison to the FE model:

- For surface grains (d = 0):

  – $CI_{95\%}(\tau_{FE}^{*g} - \tau_{CA}^{*g}) = -0.002 \pm 0.018$

  – $CI_{95\%}(\frac{\tau_{FE}^{*g} - \tau_{CA}^{*g}}{\tau_{FE}^{*g}}) = -0.472 \pm 3.652$ [%]

- For in-depth grains (d > 0):

  – $CI_{95\%}(\tau_{FE}^{*g} - \tau_{CA}^{*g}) = -0.001 \pm 0.016$

  – $CI_{95\%}(\frac{\tau_{FE}^{*g} - \tau_{CA}^{*g}}{\tau_{FE}^{*g}}) = -0.110 \pm 3.129$ [%]

An estimation error of approximately ±4% (±0.02 in absolute terms) validates the precision of the CA model, demonstrating its effectiveness in identifying areas of high stress concentration within polycrystalline aggregates.

However, determining whether stress concentrations are more likely to be located at the surface or within the bulk material presents a more challenging task. The range of prediction error is comparable to the variations in NRSS observed across different depths. For instance, in the subset *S-Agg-2+*, the average $\langle \tau_{FE}^{*g} \rangle_g$ is 0.666 for d = 0 and 0.620 for d > 0, yielding a difference of 0.046, which is only twice the magnitude of the prediction error interval. Fig. 6 displays the box plots of τ^{*g} predicted by FE and CA models for various subsets of data and grain depths (each box plot is based on 100 values). Although there are discrepancies between the box plots of each model (CA box plots being shifted towards higher values compared to FE box plots), the trends observed across different grain depths are consistent for both models. On average, grains located at d = 0 exhibit the highest stress, whereas those at d = 1 demonstrate the lowest stress levels. Grains at d = 2 and d = 9 fall somewhere in between, but closer to the stress levels observed at d = 1. This shows that despite the CA model exhibiting prediction errors of similar magnitude as the average variations in NRSS attributable to the grain's depth, it remains capable of pinpointing where grains experiencing the highest stress are likely to be located: either at the surface or within the bulk material.

4. Stress field analysis by mean of the FEM model

Before undertaking large-scale statistical analyses of aggregate stress fields using the CA model, a preliminary investigation is performed with the FEM model. This analysis first examines the linear-elastic stress field using *S-Agg* simulaitons, followed by an assessment incorporating elasto-plastic behavior for selected *S-Agg* aggregate configurations.

4.1. Linear-elastic FEM analysis

The extreme values observed for *S-Agg-3+* in Fig. 4(b) theoretically represent the maximum values a grain can possibly reach, but statistically, they are highly improbable. Conversely, *S-Agg-2+* provides the maximum values that should be expected within a dataset of 10^8 grains. These values are detailed in Table 3 as a function of the grain's depth. In-depth grains, having more neighboring grains, would exhibit the highest possible value if all these neighbors were perfectly aligned to generate stress concentration, as observed in the table for *S-Agg-3+*. Following the same logic, surface grains should theoretically exhibit the lowest value. However, this is not observed in the results. Surface grains actually display the second-highest value. One potential explanation for this discrepancy could be the stronger individual influence of neighboring grains on surface grains compared to in-depth grains. This could be attributed to the fact that surface grains constitute only half of a Kelvin's cell. On the other hand, since in-depth grains have 536 neighboring grains compared to 292 for surface grains, it is more probable for the neighbors of surface grains to all be aligned and generate stress concentration compared to in-depth grains. Additionally, due to
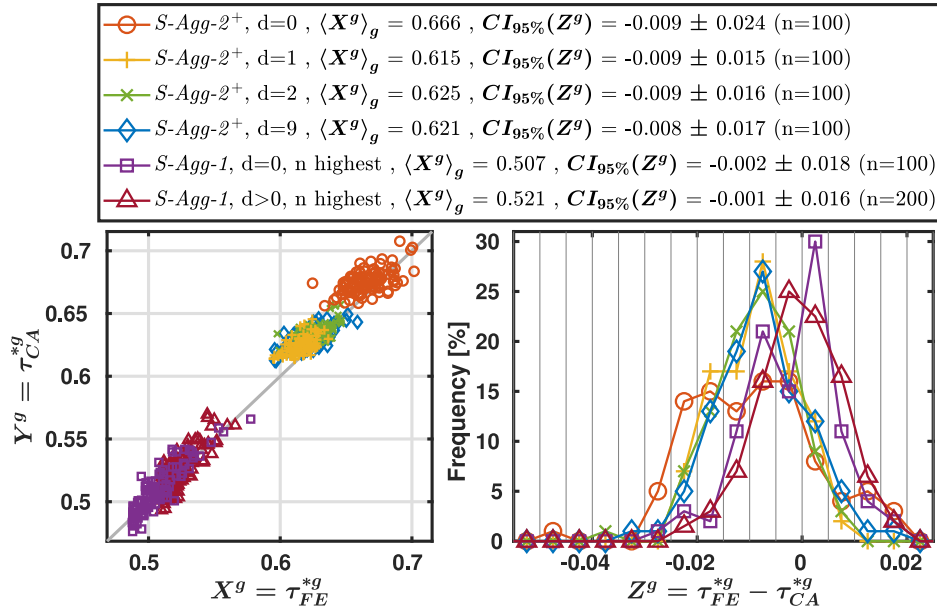
Fig. 5. Scatter plot of FE and CA-NEW4 model predictions of *S-Agg* highest NRSS values, along with the distribution of the model differences across the 800 values. The exponent sign $\bullet^+$ refers to the predictions of the grains showing the highest τ_{FE}^{*g} among the 65 grains considered within an aggregate. Over all subsets, $CI_{95\%}(Z^g) = -0.006 \pm 0.019$.
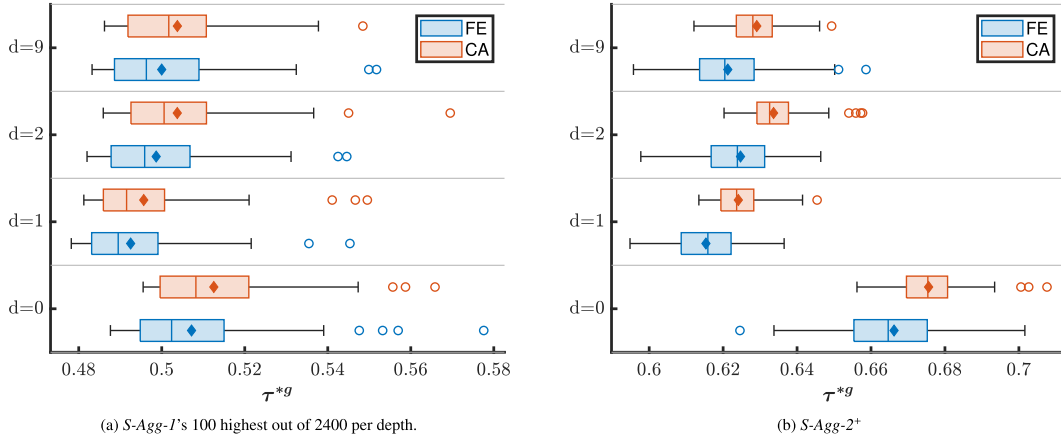

(a) *S-Agg-1*'s 100 highest out of 2400 per depth.

(b) *S-Agg-2*$^+$

Fig. 6. Box plot of τ^{*g} predicted by FE and CA models for different grain's depth and subsets of data (100 values were used for each box plot): (a) *S-Agg-1*'s 100 highest values out of the 6 grains located at $d = i \times$ among the 400 *S-Agg-1* random aggregates; (b) *S-Agg-2*$^+$ subset.

Table 3
Variation of τ_{FE}^{*g} depending on the grain's depth in the case of high stress concentration configurations.

	$d =$	0	1	2	9
$\langle \tau_{FE}^{*g} \rangle_g$	*S-Agg-2*$^+$	6.66	6.15	6.25	6.21
	S-Agg-3$^+$	1.40	1.24	1.33	1.46

the stronger influence of surface grains' neighbors, they are expected to have higher values than in-depth grains, as observed in the table for *S-Agg-2*$^+$.

On another note, the variations of the grains' second moment depending on the grains' depth is an interesting value to consider. Up to this point, only the first moment (average value within a grain) has been observed. The CA model is limited to predicting the first moment, while the second moment (the standard deviation of all the grain's elements) can be derived from FE simulations. This value is relevant for characterizing the heterogeneity within a grain. The second moment of a grain's stress level is calculated as follows:

$$std_{\boldsymbol{\sigma}}^g = \sqrt{\sum_{el \; from \; g} \frac{v_{el}}{v_g} \cdot \left(\| \boldsymbol{\sigma}^{el} - \boldsymbol{\sigma}^g \|_{eq} \right)^2} \tag{16}$$

where v_{el} and v_g are the element and grain volumes, respectively. Fig. 7 illustrates the variations in the second moment of grains as a function of several factors: the grain's average stress level, the grain's depth, and the prediction differences between the FE and CA models. Three key observations can be drawn from the figure:

- There appears to be no evident correlation in Fig. 7(a) between the second moment and the prediction error of the CA model. This suggests that even in cases where high heterogeneities exist within the grain, the CA model can predict a grain's mean stress level with consistent accuracy.
- As the stress concentration within the aggregate increases, so does the second moment within the grains of the aggregate. The boxes in Fig. 10 display higher $std_{\boldsymbol{\sigma}}$ for *S-Agg-3* compared to *S-Agg-2*, and they are higher for *S-Agg-2* compared to *S-Agg-1*. This can be explained by the fact that neighboring grains can induce stresses in different directions, leading to stress heterogeneity within the grain. Grains with high stress levels (such as those in the *S-Agg-2*$^+$ subset) have neighbors with strong influences, but these neighbors may not all pull in the exact same direction, but still resulting in a stress concentration within the grain. This explanation is also
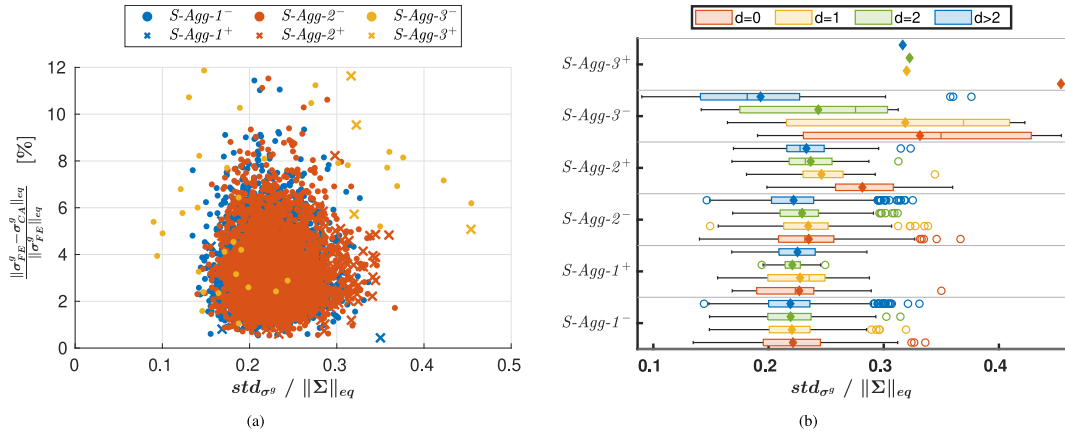
Fig. 7. Variations of the grains' stress level second moment as a function of (a) the FE and CA models predictions differences, (b) different subset of data based on grains' depth and stress level.

supported by the observations of *S-Agg-3$^+$* grains, which do not exhibit significantly high standard deviations compared to their very high stress levels. This is because their entire neighborhood is perfectly oriented to increase stress within the grain in the same direction, thus introducing no significant heterogeneities. However, their neighboring grains (*S-Agg-3$^-$*) show higher standard deviations than the other aggregates.

- Surface grains are more likely to display a higher second moment than their neighbors. For each subset of data in Fig. 7(a), the boxes display higher std_σ^g the closer the grains are to the surface. The explanation from the previous point can also be applied to this observation. Surface grains are influenced by their neighbors on only one side, leading to heterogeneities within the grain. Additionally, grains located at d = 0 are only half of a Kelvin's cell, which could be another reason for this observation.

At equal stress levels, grains with a higher second moment should be more likely to plastify first. Within the *S-Agg-2$^+$* subset, surface grains exhibit higher second moments, making them even more likely to plastify first in addition to their higher stress levels.

4.2. Elasto-plastic FEM analysis

4.2.1. Finite element model additional feature

The above analysis was conducted under the assumption that the grain most likely to initiate a crack first is the one with the highest NRSS during the elastic load. To validate this assumption, elasto-viscoplastic FEM simulations were performed using the aggregates *S-Agg* presented in Section 3.2. The grains' mechanical behavior is no longer considered purely elastic and has been extended to an elasto-viscoplastic behavior. A cyclic load with a stress rate $R = -1$, a constant strain rate of $\dot{E}_{33} = 10^{-3}$ s^{-1}, and a strain amplitude of 0.06% is applied to the aggregate's matrix. This amplitude was chosen just below the material elastic limit, allowing the matrix surrounding the grains to still be assumed as linear-elastic. As in Section 3.2, only the behavior of the 65 core-grains with their four whole neighboring grain layers was reported. These simulations are not intended to study the cyclic behavior of the grain, but rather to confirm the correlation between the grains' NRSS and their susceptibility to initiate crack formation in the presence of a free surface.

The Meric-Cailletaud's model [40] was employed to characterize the crystal elasto-viscoplastic behavior, defined as:

$$\dot{\gamma}_s^g = \left(\frac{\tau_s^g - \chi_s^g}{|\tau_s^g - \chi_s^g|} \right) \dot{v}_s^g \tag{17a}$$

with $$\dot{v}_s^g = \left(\frac{\left(|\tau_s^g - \chi_s^g| - r_s^g \right)^+}{K} \right)^n \tag{17b}$$

where $\dot{\gamma}_s^g$ is the grain *g* slip rate of the slip system *s*; $(\cdot)^+$ denotes the operator taking the positive part of its argument; v_s^g is the system *s* cumulative viscoplastic slip; *K* and *n* are material parameters that characterize the viscous effect; χ_s^g and r_s^g are, respectively, the kinematic and isotropic hardening. The chosen parameters are derived from the works of Y. Guilhem [38] on 316L stainless steel in HCF. Only the 12 octahedral slip systems were considered. All simulations presented in this article were conducted using the equations and material parameters listed in Appendix B. For further details on the model, readers are encouraged to refer to the following articles [17,38,40].

FE simulations were conducted using ABAQUS software, with the Meric-Cailletaud model implemented through the UMAT function. Due to the computational weight of such simulation, the mesh shown in Fig. 3(b) was retained, although it was not sufficient to achieve viscoplastic-variable convergence. However, it was deemed adequate for capturing the overall trends. For a more comprehensive analysis of the grain's cyclic behavior, a finer mesh and a larger aggregate would be required.

The total cumulative viscoplastic slip v_Σ^g is employed to evaluate grain fatigue damage, a variable commonly used in crack nucleation criteria. It is defined as the summation of the cumulative viscoplastic slip $v_s^g(t)$ at the instant *t* across all 12 octahedral slip systems:

$$v_\Sigma^g(t) = \sum_{s=1}^{12} v_s^g(t) \tag{18}$$

In this study, 20 aggregates from the *S-Agg-2* set (five from each stress concentration depth $d = \{0, 1, 2, 9\}$) and an additional five aggregates from the *S-Agg-1* set were randomly selected, resulting in a total of 25 simulations. Each simulation provides 65 values of v_Σ^g. Only 1/4 of a cycle was simulated, except for five simulations (one for each subset) where 4 cycles were simulated (due to server issues, only 1.5 cycles were simulated for two of them).

Using these longer simulations (65 × 5 values), it can be demonstrated that the values of v_Σ^g at 1/4 cycle are proportional to the values of v_Σ^g at 1.5 and 4 cycles. Their regressions exhibit an excellent R^2 coefficients:

$$v_\Sigma^g(1.5 \ cycles) \approx 4.89 \cdot v_\Sigma^g(1/4 \ cycle), \quad (R^2 = 0.999), \quad (5 \ x \ 65 \ values) \tag{19a}$$

$$v_\Sigma^g(4 \ cycles) \approx 14.6 \cdot v_\Sigma^g(1/4 \ cycle), \quad (R^2 = 0.998), \quad (3 \ x \ 65 \ values) \tag{19b}$$

Even though the elasto-viscoplastic behavior includes softening, it should only appear after several cycles, which explains the excellent R^2 coefficients. This correlation allows us to study the grain cyclic damage by only focusing on the value at the end of the first tensile load (1/4 cycle) which considerably reduce the computational time
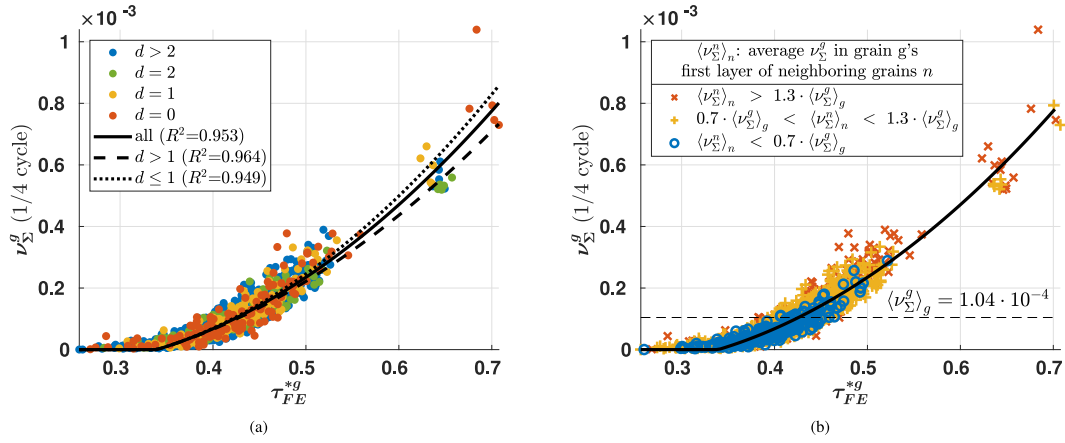
Fig. 8. Correlation between the grains' values of v_Σ^g at 1/4 cycle predicted by the FE model and their NRSS τ_{CA}^{*g} predicted by the CA model along with a second order polynomial regression. (a) color coded as a function of the grain's depth; (b) color coded as a function of the grain's neighbor plasticity level.

Table 4
Eq. (20) fitting parameter b obtained for different subsets of data ($a = 9.11 \cdot 10^{-4}$ and $\tau_{crit}^* = 0.34$).

Selected grains	all	$d > 1$	$d \leq 1$	$d = 0$	$d = 1$	$d = 2$	$d = 9$
$b = [\times 10^{-3}]$	3.43	2.93	3.85	3.88	3.79	2.68	3.07

needed to perform the elasto-viscoplastic simulations. For simplicity, $v_\Sigma^g(1/4\ cycle)$ will be denoted as v_Σ^g in the following.

4.2.2. Analysis of the elasto-plastic FEM simulations

Fig. 8 show the grains' plastic damage at 0.06% deformation (1/4 cycle) predicted by the FE model (v_Σ^g) as a function of the grains' NRSS at 0.01% deformation (before plasticity) predicted by the CA model (τ_{CA}^{*g}). The figure confirms the correlation between v_Σ^g and τ_{CA}^{*g}, and a second order polynomial function was used to express v_Σ^g as a function of τ_{CA}^{*g}:

$$v_\Sigma^g \approx f_{v_\Sigma^g}(\tau_{CA}^{*g}) = \begin{cases} 0 & \text{if} \quad \tau_{CA}^{*g} < \tau_{crit}^* \\ a \times (\tau_{CA}^{*g} - \tau^*) + b \times (\tau_{CA}^{*g} - \tau_{crit})^2 & \text{else} \end{cases}$$

(20)

τ_{crit}^* corresponds to the minimum NRSS value necessary for the slip system to reach its critical resolved shear stress $r_0 = 40$ MPa. Thus, τ_{crit}^* is equal to $40/(E_y \cdot 0.06\%) = 0.34$. The coefficients a and b have been fitted using all the available data, resulting in values of 9.11×10^{-4} and 3.43×10^{-3}, respectively, with a coefficient of determination $R^2 = 0.953$, confirming the validity of the correlation.

It appears that in Fig. 8(a), for the same value of τ_{CA}^{*g}, grains located at the surface ($d \leq 1$, represented by orange and yellow dots) exhibit higher v_Σ^g values compared to those below ($d > 1$, represented by blue and green dots). This observation is further supported by fitting Eq. (20) parameter b to different subsets of data corresponding to grains' depth. While keeping the parameters a and τ_{crit}^* constant, Table 4 presents the fitted parameter b obtained for different grains' depths. Surface grains demonstrate a higher value of b, indicating that, for a given NRSS, surface grains undergo greater plastic damage than grains located below. However, although a larger dataset would be necessary to accurately estimate the influence of grain depth, it can be argued that the differences observed between the fitting curves are negligible in comparison to the overall trend. Hence, the primary correlation for fatigue damage within the grains predominantly lies with their NRSS, while the influence of grain depth appears to be negligible.

Another observation from Fig. 8 is that grains deviating the most from the fitting curve are those with a neighborhood that has undergone more plastic deformation than the average. In Fig. 8(b), grains are color-coded into three categories based on the plasticity level of their close neighborhood: grains with a highly plastified close neighborhood are represented by red crosses in the background, those with a mildly plastified close neighborhood are denoted by yellow plus signs, and grains with a non-plastified close neighborhood are depicted as blue circles in the foreground. It can be observed that grains deviating from the fitting curve are grains with a plastified neighborhood. This observation could be explained by the fact that when a grain undergoes plastic deformation, its apparent stiffness tensor varies, thereby affecting its influence on its neighbors (Eqs. (6) and (7)). Its influence can either be amplified or diminished depending on its relative position with the neighboring grain. Furthermore, the observation that grains with high plasticity levels are also surrounded by a plastified neighborhood (as seen in Fig. 8(b), where all grains displaying $\tau^{*g} > 0.5$ are either yellow or red) suggests an escalation of plasticity levels. These grain plasticity levels may increase stress levels in their neighborhood, further augmenting its plastic strain and creating a sort of snowball effect. Consequently, there may be an agglomeration of grains with higher values of v_Σ^g, which are ideal sites for crack initiation.

The computation of a grain's NRSS by the CA model only considers the elastic behavior of its neighborhood. However, due to the influence of grains' plastic strain on their environment, τ_{CA}^{*g} is not a perfect indicator of grain damage rate. Nonetheless, grains most affected by neighboring plasticity would be those next to a grain undergoing a higher plastic rate and thus with a very high NRSS. The probability of two adjacent grains having very high NRSS values is almost zero. Therefore, since the stress level applied to the aggregate is low in HCF-VHCF, the neighborhood effect variation due to plastic strains is negligible. Thus, τ_{CA}^{*g} remains a reliable indicator for identifying critical configurations that lead to high damage rates and, consequently, crack initiation.

To conclude on the correlation between v_Σ^g and τ^{*g}, according to the regression function $f_{v_\Sigma^g}(\tau_{CA}^{*g})$, a grain with a NRSS of 0.62 (the expected maximum NRSS in the case of GEO1 in Fig. 11) should exhibit a value of $v_\Sigma^g = 5.26e-04$. This represents a damage rate 5 times faster than the average grain ($\langle v_\Sigma^g \rangle_g = 1.04e-04$) and 1.7 times faster than the maximum rate observed in the random distributions S-Agg-1 ($v_\Sigma^g = 3.18e-04$). The damage rate increasing exponentially with τ^{*g} justifies even more the need to study aggregates such as S-Agg-2 rather than random aggregates, highlighting the importance of identifying such configurations for deeper studies. However, it is essential to acknowledge the limitations of this correlation. Using the three sigma rules, a 95% confidence interval of $f_{v_\Sigma^g}(\tau_{CA}^{*g})$ can be estimated using the 20 grains displaying a NRSS higher than 0.6:

- $CI_{95\%}(v_\Sigma^g - f_{v_\Sigma^g}(\tau_{CA}^{*g})) = 0.033 \pm 1.923 \ [\times 10^{-4}]$
- $CI_{95\%}(\dfrac{v_\Sigma^g - f_{v_\Sigma^g}(\tau_{CA}^{*g})}{v_\Sigma^g}) = -2.435 \pm 25.02 \ [\%]$

A data set larger than 20 values would be required to obtain a more accurate estimate of the confidence interval. However, the range of the interval highlights the lack of precision in the correlation for grains displaying high NRSS.

5. Probabilistic approach of the highest NRSS localization

Considering the observations from Section 4.2, which indicate that grains with the highest NRSS are the ones most likely to plastify first and thus initiate a crack, the probability of obtaining a high value of τ^{*g} based on a grain's depth was analyzed. This study helps assess the conditions under which a crack is more likely to initiate, whether near the surface or deeper within the material.

5.1. Independent random variables' definitions

In order to proceed, several variables, terms and equations need to be introduced:

- X_g^d represents the dimensionless shear stress τ^{*g} of a random grain g surrounded by a random neighborhood, located at a depth d. In the following, we will assume N_d random variables $X_1^d, \ldots, X_{N_d}^d$ that are identically and independently distributed with a common probability distribution function (PDF) $P(X_g^d = \tau^{*g}) = f_X^d(\tau^{*g})$. The left-tailed cumulative distribution function (CDF) is noted $P(X_g^d \leq \tau^{*g}) = F_X^d(\tau^{*g})$.

- $Y_{N_d}^d = max(X_1^d, \ldots, X_{N_d}^d)$ denotes the highest value of τ^{*g} observed among N_d random grains located at a depth d. It is PDF and CDF noted $f_Y^d(\tau^{*g}, N_d)$ and $F_Y^d(\tau^{*g}, N_d)$, respectively, can be defined as a function of $f_X^d(\tau^{*g})$:

$$f_Y^d(\tau^{*g}, N_d) = N_d \times f_X^d(\tau^{*g}) \times F_X^d(\tau^{*g})^{(N_d-1)} \quad (21a)$$

$$F_Y^d(\tau^{*g}, N_d) = F_X^d(\tau^{*g})^{N_d} \quad (21b)$$

The probability to have at least one grain among N_d grains at a depth d with a value of τ^{*g} higher than y is $P(Y_{N_d}^d > y) = 1 - F_Y^d(y)$.

- $Z^D = max(Y_{N_{d_1}}^{d_1}, \ldots, Y_{N_{d_I}}^{d_I})$ denotes the highest value of τ^{*g} observed among $\sum_{i=1}^{i=I} N_{d_i}$ random grains located at different depths d_i from the set of depth $D = \{d_1, \ldots, d_I\}$. It is PDF and CDF noted $f_Z^D(\tau^{*g}, \{N_{d_1}, \ldots, N_{d_I}\})$ and $F_Z^D(\tau^{*g}, \{N_{d_1}, \ldots, N_{d_I}\})$, respectively, can be defined as a function of $f_X^d(\tau^{*g})$:

$$f_Z^D(\tau^{*g}, \{N_{d_1}, \ldots, N_{d_I}\}) = \sum_{i=1}^{i=I} N_{d_i} \times f_X^{d_i}(\tau^{*g}) \times F_X^{d_i}(\tau^{*g})^{-1}$$

$$\times \prod_{j=1}^{j=I} F_X^{d_j}(\tau^{*g})^{(N_{d_j})} \quad (22a)$$

$$F_Z^D(\tau^{*g}, \{N_{d_1}, \ldots, N_{d_I}\}) = \prod_{j=1}^{j=I} F_X^{d_j}(\tau^{*g})^{(N_{d_j})} \quad (22b)$$

In the case of a depth-set of a single layer ($D = \{d_i\}$) we have then $f_Z^D(\tau^{*g}, \{N_{d_i}\}) = f_Y^{d=d_i}(\tau^{*g}, N_{d_i})$ and $F_Z^D(\tau^{*g}, \{N_{d_i}\}) = F_Y^{d=d_i}(\tau^{*g}, N_{d_i})$. The probability to have at least one grain among all grains within a range of depth with a value of τ^{*g} higher than z is $P(Z^D > z) = 1 - F_Z^D(z)$.

- $U^{D_1, D_2} = Z^{D_2} - Z^{D_1}$ denotes the difference between of the highest value of τ^{*g} observed among random grains located in the two different sets of depth D_1 and D_2. It is PDF and CDF noted $f_U^{D_1, D_2}(\tau^{*g}, D_1, D_2)$ and $F_U^{D_1, D_2}(\tau^{*g}, D_1, D_2)$ respectively can be defined as a function of $f_Z^d(\tau^{*g})$:

$$f_U^{D_1, D_2}(u) = \int_{-\infty}^{\infty} f_Z^{D_1}(x, \{N_{d_{1,1}}, \ldots, N_{d_{1,I}}\})$$

$$\cdot f_Z^{D_2}(x - u, \{N_{d_{2,1}}, \ldots, N_{d_{2,I}}\}) \cdot dx \quad (23a)$$

$$F_U^{D_1, D_2}(u) = \int_{-\infty}^{u} f_U^{D_1, D_2}(x) \cdot dx \quad (23b)$$

The probability that the difference between the maximum τ^{*g} within the depth-sets D_1 and D_2 is lower or equal to a given value u is $P(U^{D_1, D_2} \leq u) = P(Z^{D_2} - Z^{D_1} \leq u) = F_U^{D_1, D_2}(u)$. Therefore, $P(U^{D_1, D_2} \leq 0) = P(Z^{D_2} \leq Z^{D_1}) = F_U^{D_1, D_2}(0)$ is the probability that the grain with the highest value of τ^{*g} among all the grains from the two depth-sets D_1 and D_2 is located in the depth-set D_1.

5.2. Analysis and fitting of τ^{*g} distributions

To calculate $P(U^{D_1, D_2} \leq 0)$, it is essential to evaluate the two functions $f_X^d(\tau^{*g})$ and $F_X^d(\tau^{*g})$, particularly for the extreme values of τ^{*g}. To achieve this, the CA model proposed in Section 2.2 and tested in Section 3.3 was employed to estimate the distribution functions of τ^{*g} for different grain depths. Its efficiency and accuracy enable the estimation of τ^{*g} for a large number of grain-neighborhood configurations within a short calculation time. For each depth d under study, 10^{10} random grain-neighborhood configurations were generated, and the NRSS of each configuration was estimated using the CA model. Among these 10^{10} configurations, every 10^8 block, the configuration leading to the highest NRSS value in the block was saved and utilized for the specific configurations of *S-Agg-2* from FEM simulations, as defined in Section 3.2. The methodology for generating these distributions proceeded as follows:

- A pool of 100,000 random crystallographic orientations was generated using the quaternion representation [41,42]. For each orientation, depth, and relative position $\overline{gn}$, the values $\sigma_0^g(d) / \Delta\sigma_n^g(d)$ were preliminary calculated according to Eqs. (6), (8).
- The central grain and its four layers of neighboring grains were randomly assigned orientations from the initial pool. The central grain's NRSS was calculated according to Eqs. (3) and (9) based on the given orientation configuration and utilizing the preliminarily calculated values of $\sigma_0^g(d) / \Delta\sigma_n^g(d)$. This step was repeated 10^{10} times for each studied depth.

Because of the straightforward nature of the CA model, the aforementioned algorithm is easily parallelizable. As a result, it was implemented in CUDA-FORTRAN, leveraging the parallel computing capabilities of graphics processing units (GPUs). With an Nvidia RTX3090 GPU, estimating the τ^{*g} values for 10^7 grain-neighborhood configurations takes approximately 40 s. Thus, the total time required to generate the distributions for 10^{10} evaluations across all selected depths d from 0 to 9 and ∞ is roughly 122 h.

The cumulative distribution of the 10^{10} predicted values of τ^{*g} is depicted in Fig. 9 for each grain depth. Overall, these distributions closely resemble each other, except for $d = 0$, which exhibits a slightly wider distribution. The distribution's mean is approximately 0.410 at $d = 0$, and it fluctuates around this value within a range of ± 0.02 as the depth increases. It quickly converges back to 0.410 at around $d = 7$. Despite minor differences, these average values closely match those obtained from the 400 random *S-Agg-1* aggregates discussed in Section 3.3, as shown in Table 5. The observed discrepancies at the surface may be attributed to either the limited accuracy of the CA model for $d = 0$ or the insufficient number of *S-Agg-1* aggregate simulated to precisely capture the average NRSS, possibly due to a wider distribution of τ^{*g} for $d = 0$.

Given that material parts are composed of millions of grains, it is more pertinent to focus on the extreme values. A closer examination of the extreme values of τ^{*g} in Fig. 9(b), using a scale $log_{10}(1 - cdf)$,
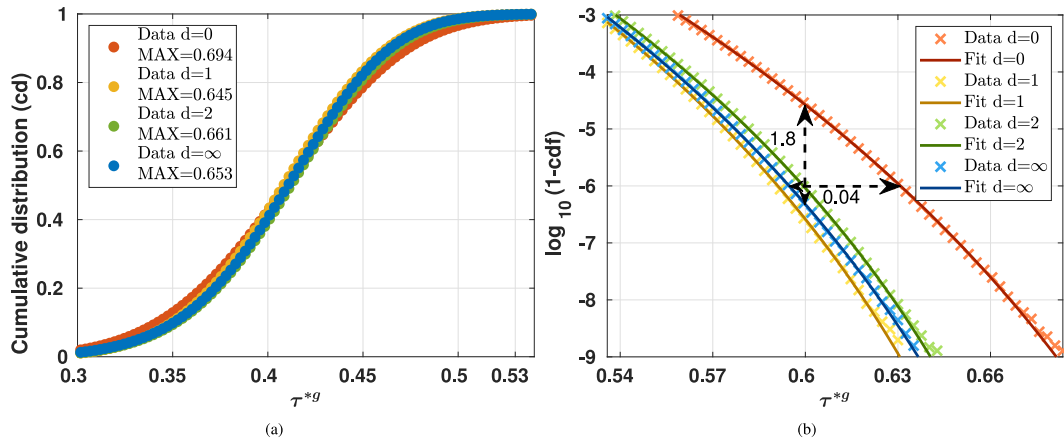
Fig. 9. Distribution of τ^{*g} predicted by the CA model: (a) Cumulative distribution of 10^{10} evaluations of random grain-neighborhood crystallographic orientation configurations for various grain depths d; (b) Fitting of the distribution by means of the generalized Pareto function. The scale $log_{10}(1-cdf)$ is utilized to zoom in on the extreme values of τ^{*g} with very low probability.

Table 5
Average value of τ^{*g} per depth predicted by the FE or CA models for different dataset sizes: 2400 value per depth from *S-Agg-1* random aggregates versus a dataset of 10^{10} random configurations.

$\langle \tau^{*g} \rangle_g$	$d =$	0	1	2	3	4	5	6	7 to 9	∞
2400 values	FE	0.407	0.406	0.411	0.410	0.410	0.412	0.411	0.410	$\varnothing$
(*S-Agg-1*)	CA	0.409	0.406	0.412	0.411	0.410	0.412	0.411	0.410	$\varnothing$
10^{10} values	CA	0.410	0.408	0.412	0.412	0.412	0.411	0.411	0.410	0.410

Table 6
Generalized Pareto parameters for each studied grain's depth.

$d =$	0	1	2	3	4	5	6	7	8	9	∞
κ_0^d	0.559	0.533	0.539	0.538	0.537	0.537	0.536	0.536	0.536	0.535	0.535
ξ^d	−0.050	−0.056	−0.056	−0.057	−0.054	−0.050	−0.050	−0.050	−0.051	−0.051	−0.052
θ^d	0.012	0.010	0.011	0.011	0.011	0.010	0.010	0.010	0.010	0.010	0.010

reveals that the distribution for $d = 0$ differs from the others. It indicates that a grain at $d = 0$ has approximately $10^{1.8}$ times more likelihood to reach a value higher than 0.6 compared to a grain at $d = \infty$. In another perspective, the probability of a grain at $d = 0$ exceeding 0.63 is 10^{-6}, while a grain at $d = \infty$ has the same probability of exceeding 0.59 (0.04 lower). The distribution for $d = 1$ exhibits the lowest probability of extreme values. Subsequently, the distribution for $d = 2$ shows a slight increase compared to $d = 1$, becoming the second-highest probability after $d = 0$. From $d = 2$ onwards, the probability gradually decreases towards $d = \infty$.

To predict the probability of the distribution extreme values, the peak over threshold method [43] with the generalized Pareto distribution (GPD) [44] were used. The PDF $f_X^d(\tau^{*g})$ and CDF $F_X^d(\tau^{*g})$ are then expressed as follows:

$$f_X^d(\tau^{*g}) = P(X^d > \kappa_0^d) \left(\frac{1}{\theta^d} \left(1 + \xi^d \frac{\tau^{*g} - \kappa_0^d}{\theta^d} \right) \right)^{-(1/\xi^d + 1)} \quad \forall \tau^* \geq \kappa_0^d$$

(24a)

$$F_X^d(\tau^{*g}) = (1 - P(X^d > \kappa_0^d)) + P(X^d > \kappa_0^d)$$
$$\times \left(1 + \xi^d \frac{\tau^{*g} - \kappa_0^d}{\theta^d} \right)^{-1/\xi^d} \quad \forall \tau^* \geq \kappa_0^d$$

(24b)

The selected thresholds κ_0^d are chosen so that $P(X^d \leq \kappa_0^d) = (1 - 10^{-3})$, which is equivalent to $P(X^d > \kappa_0^d) = 10^{-3}$. This chosen value of 10^{-3} is sufficiently low considering that a typical mechanical part is composed of far more than 10^3 grains, yet not too low to maintain an excellent fit with the distribution data. The parameters ξ^d and θ^d do not have a physical interpretation; they only control the curvature and the rate of decrease of the curve. The fitted curves for $d = 0, 1, 2, \infty$

are plotted in Fig. 9(b) along with their corresponding distributions, and the fitted parameters for each depth are listed in Table 6. The lowest R^2 value obtained from the fitting curve is 0.9998, indicating the goodness of fit of the GPD parameters. The obtained parameters align with the observations mentioned in the paragraph above. For the values of τ^{*g} below κ_0^d, a normal distribution was used. The normal distribution's parameters are adjusted to ensure the continuity of f_X^d and F_X^d at $\tau^{*g} = \kappa_0^d$.

5.3. Probabilistic localization of τ^{*g} highest value in mechanical parts

In this section, the probability of the highest τ^{*g} value and its localization in a real-size mechanical part is studied. To do so, the dimensions of a standard cylindrical fatigue test specimen, following ASTM E-466 specifications, were chosen as the reference geometry (GEO1). Its cylindrical working section has a height h_c of 25 mm and a diameter $\varnothing_c$ of 6.5 mm. To explore the influence of the ratio between the number of surface and in-depth grains on the localization probability of the grain with the highest NRSS, two additional specimen geometries were derived. The second specimen (GEO2) halves the height and doubles the diameter compared to SP1, resulting in twice the number of in-depth grains while maintaining the same number of surface grains. Conversely, the third geometry (GEO3) quadruples the height and halves the diameter, doubling the number of surface grains while retaining the same number of in-depth grains. The number of grains per depth N_d was determined based on the specimen dimensions. An illustration alongside equations estimating the number of grains per depth, depending on the cylinder's dimensions, is provided in Fig. 10.
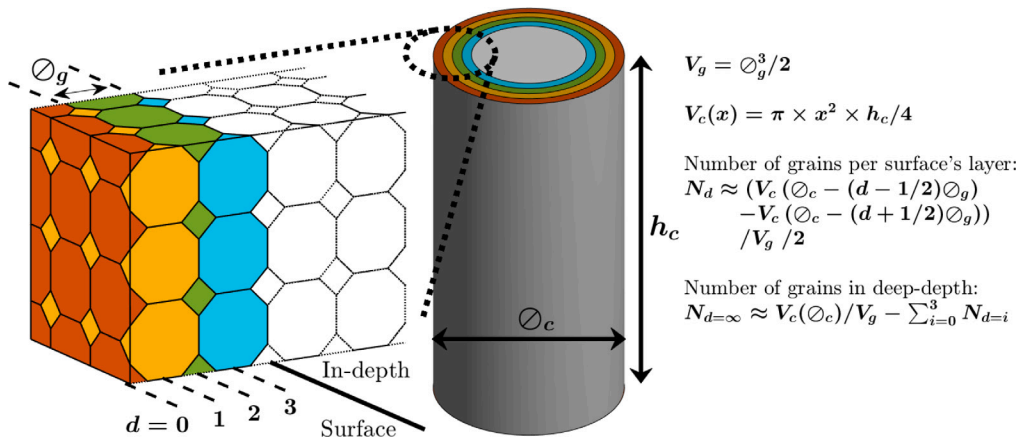
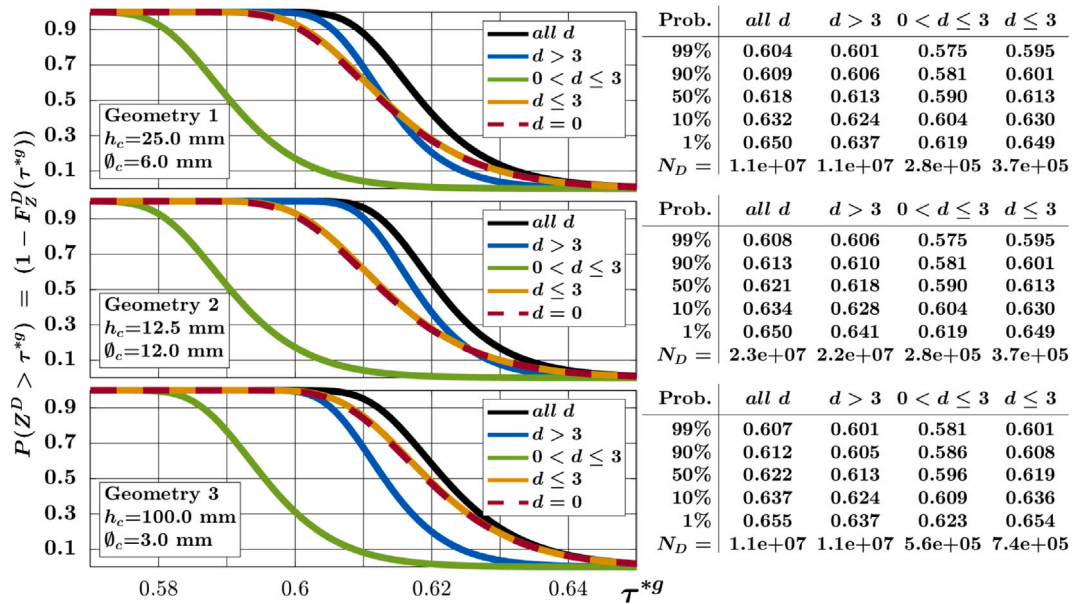**Fig. 10.** Illustration of the grains counting per grain-layers in a cylinder.

$$V_g = \varnothing_g^3/2$$

$$V_c(x) = \pi \times x^2 \times h_c/4$$

Number of grains per surface's layer:
$$N_d \approx (V_c\,(\varnothing_c - (d-1/2)\varnothing_g) \\ - V_c\,(\varnothing_c - (d+1/2)\varnothing_g)) \\ /V_g\,/2$$

Number of grains in deep-depth:
$$N_{d=\infty} \approx V_c(\varnothing_c)/V_g - \sum_{i=0}^{3} N_{d=i}$$

Prob.	all d	d > 3	0 < d ≤ 3	d ≤ 3
99%	0.604	0.601	0.575	0.595
90%	0.609	0.606	0.581	0.601
50%	0.618	0.613	0.590	0.613
10%	0.632	0.624	0.604	0.630
1%	0.650	0.637	0.619	0.649
$N_D =$	1.1e+07	1.1e+07	2.8e+05	3.7e+05

Prob.	all d	d > 3	0 < d ≤ 3	d ≤ 3
99%	0.608	0.606	0.575	0.595
90%	0.613	0.610	0.581	0.601
50%	0.621	0.618	0.590	0.613
10%	0.634	0.628	0.604	0.630
1%	0.650	0.641	0.619	0.649
$N_D =$	2.3e+07	2.2e+07	2.8e+05	3.7e+05

Prob.	all d	d > 3	0 < d ≤ 3	d ≤ 3
99%	0.607	0.601	0.581	0.601
90%	0.612	0.605	0.586	0.608
50%	0.622	0.613	0.596	0.619
10%	0.637	0.624	0.609	0.636
1%	0.655	0.637	0.623	0.654
$N_D =$	1.1e+07	1.1e+07	5.6e+05	7.4e+05

Fig. 11. Probability of at least one grain having an NRSS exceeding τ^{*g} within a given range of grain depth for each studied geometry. N_D represents the number of grains within the considered grain depth. GEO1 denotes the geometry with realistic dimensions. GEO2 doubles the number of grains in depth-set D_∞, while $D_{0<d\leq3}$ and $D_{d\leq3}$ remain constant. GEO3 doubles the number of surface grains while maintaining a constant number in depth.

For each geometry, the most probable values of Z^D (Eq. (22)) and U^{D_1,D_2} (Eq. (23)) were examined for different depth-sets D:

- $D_{all} = \{all\ layers\ d\}$: all the sample's grains.
- $D_{d>3} = \{all\ layers\ d > 3\}$: all the in-depth grains. An arbitrary limit of 4 surface layers was chosen to set the limit between the surface and the in-depth grains as illustrated in Fig. 10.
- $D_{d\leq3} = \{d = 0, d = 1, d = 2, d = 3\}$: all the surface grains.
- $D_{0<d\leq3} = \{d = 1, d = 2, d = 3\}$: all the surface grains except those located at d = 0 which are only a half Kelvin's cell.

Grains were modeled as Kelvin cells with a diameter $\varnothing_g$ of 50 μm, typical for stainless steel 316L grain size [38,45]. As a result, the chosen threshold between surface and in-depth grains corresponds to 100 μm.

The probabilities $P(Z^D > \tau^{*g})$ (Eq. (22)) and $P((Z^{D_\infty} - Z^{D_{d\leq3}}) \leq u)$ (Eq. (23)) are illustrated in Figs. 11 and 12, respectively, for each geometry and various depth-sets. Several points can be discussed from these figures:

- An increase in either the number of grains within the specimen or its surface area corresponds to higher anticipated maximum

NRSS values and, consequently, a reduced fatigue life, as observed experimentally [46]. Within the entire specimen GEO1, the highest NRSS expected should fall within the range of 0.604 to 0.650 with a 99% chance, with a 50% chance to be either above or below 0.618. These values increase for GEO2 and GEO3 due to different factors. In GEO2, the total number of grains doubles, thereby enhancing the likelihood of encountering higher stress levels. Conversely, in GEO3, although the total number of grains remains constant, the surface grain count doubles. Since surface grains are more prone to undergo higher NRSS compared to those within the material depth, this consequently increases the expected highest NRSS within the whole part.

- The highest NRSS within the surface grains is predominantly influenced by the grains located at $d = 0$. In Fig. 11, the curves for the depth-sets $D_{d\leq3}$ and $D_{d=0}$ overlap each other, whereas the curve for $D_{0<d\leq3}$ exhibits lower values. This is due to the fact that the number of grains in each of these layers is identical and that grains located at $d = 0$ have a much higher chance to hit a higher NRSS than other surface-layers. Thus, the grain with the highest NRSS is most likely to be located at d = 0 rather than 1, 2 or 3. Therefore, the following approximation can be made:
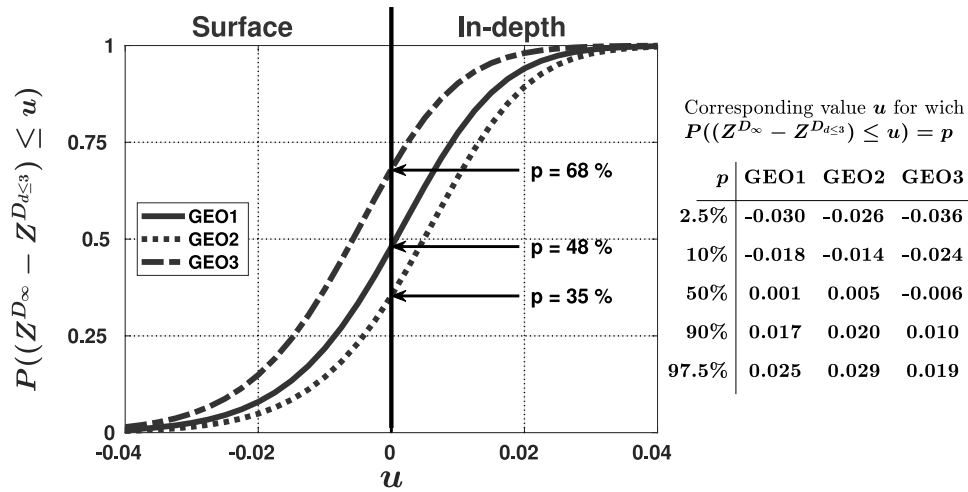
Fig. 12. Probability that the difference between the highest NRSS among grains from depth-set D_∞ and grains from depth-set $D_{d\leq3}$ is equal to u for each studied geometry. GEO1 denotes the geometry with realistic dimensions. GEO2 doubles the number of grains in depth-set D_∞, while $D_{d\leq3}$ remains constant. GEO3 doubles the number of surface grains while maintaining a constant number in depth.
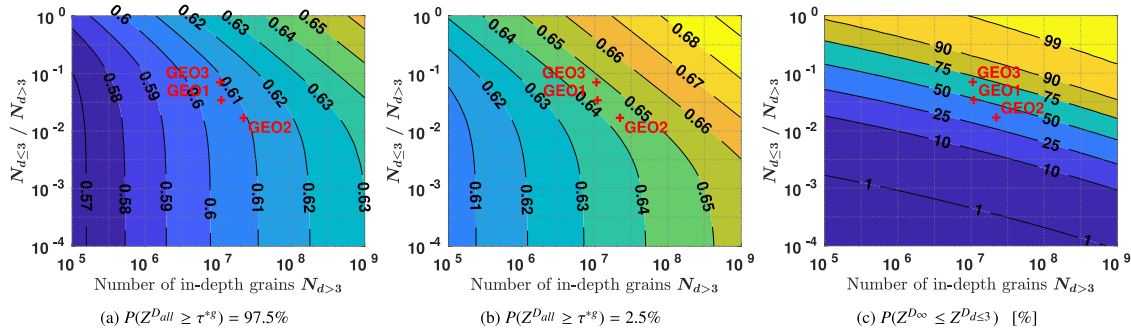

Fig. 13. Variables contour plot based on specimen dimensions: (a–b) lower and upper bounds of the 95% confidence interval for the expected highest τ^{*g} within the specimen; (c) Probability of the grain with the highest τ^{*g} being located at the surface.

$$F_Z^{D_{d\leq3}}(\tau^{*g}, \{N_{d=0}, N_{d=1}, N_{d=2}, N_{d=3}\}) \approx F_Y^{d=0}(\tau^{*g}, \{N_{d=0}\}) \quad (25)$$

- The ratio of surface area to volume of the specimen influences the location where cracks initiate. In Fig. 11, the curves corresponding to the surface grains (orange) and the curves corresponding to the in-depth grains (blue) do not intersect at the same point, depending on the geometry: GEO2 shows higher chances of a higher τ^{*g} value in-depth, while GEO3 displays higher chances of a higher τ^{*g} value at the surface. As observed in Fig. 12 for GEO1, there is a 48% chance that $Z^{D_{\inf}} \leq Z^{D_{d\leq3}}$, indicating a 48% probability that the highest NRSS is located at the surface. When the ratio of surface to in-depth grains is halved (GEO2), this probability decreases to 35%, and when the ratio is doubled (GEO3), the probability increases to 68%.

- In the case of GEO1, the disparity between the highest NRSS of surface and in-depth grains is expected to be within [−0.030; 0.025] with 95% confidence, representing ±5% of the highest expected NRSS ($P(Z^{all\ d} > 0.618 = 50\%)$). This range is of the same order as the CA model prediction margin of error, estimated at ±4%. Although Section 3.3.2 demonstrated that both the FE and CA models exhibited similar trends in the discrepancies between surface and in-depth grains, this small predicted range might not be significant enough to definitively determine whether cracks will initiate on the surface or in-depth.

On a more general note, Fig. 13 displays the 95% confidence interval of the expected highest NRSS and the probability of whether this highest NRSS is located at the surface or in-depth for any specimen size and surface area to specimen volume ratio. The figure illustrates

the broad spectrum of potential applications of the CA model and the information it can provide. By utilizing these contour maps, users can pinpoint the expected range of the highest NRSS and its probable location, which can then be employed to generate more suitable aggregates for fatigue life analysis using more detailed simulation methods, akin to what was accomplished for *S-Agg-2* aggregates.

5.4. Textured orientations pool

Previously, grain orientations were randomly generated using the quaternion method, which prevented any specific texture. Due to manufacturing processes, crystallographic texture gradients can be present within a mechanical part. This prompts consideration of the impact of introducing a crystallographic texture at the surface on Fig. 9.

Initially, orientations were chosen randomly from a pool of 100,000 random orientations. Here, three new orientation pools were generated to introduce texture, from which orientations are randomly distributed to grains. The first pool uses an EBSD map from an extruded aluminum sheet. Although aluminum differs from the 316L steel alloy, this provides a realistic texture for a theoretical CA model application. The second pool emphasizes high Schmid factor (SF) orientations: 10^6 randomly generated orientations were sorted into categories based on SF values, from which a specified number of orientations were randomly selected per category. This selection generated a new orientation pool with a high concentration of orientations having SF values above 0.48. The third pool only includes orientations leading to high-stress concentrations in *S-Agg-3$^+$* configurations. In this section, the "random pool" refers to the original orientation set, the "extrusion pool" to the EBSD-derived set, the "high SF pool" to the second set, and the "S-
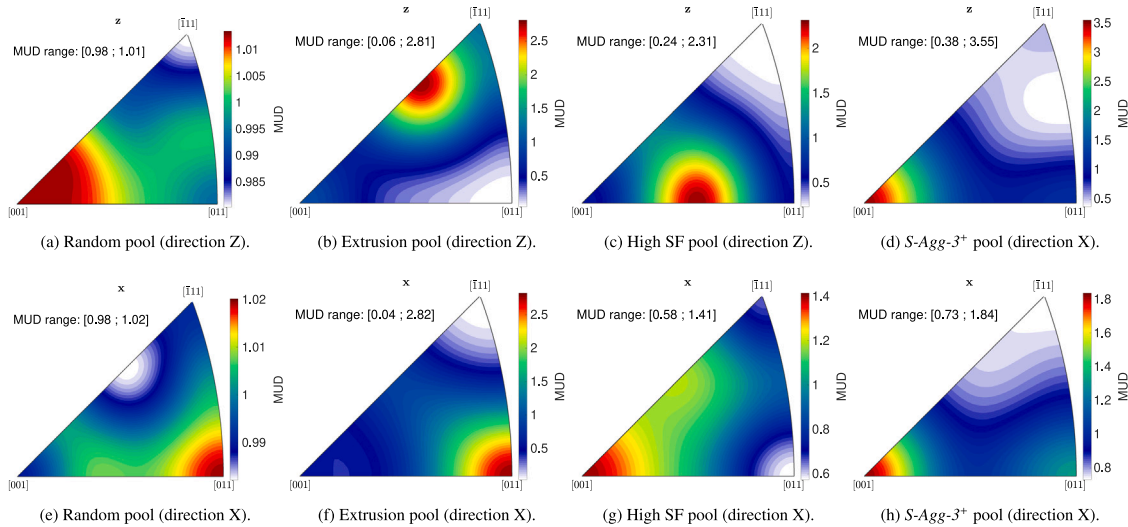
Fig. 14. ODF of all the orientation pools generated. The direction X refers to the free surface normal and the direction Z to the loading direction.

Agg-3⁺ pool" to the third. Fig. 14 presents the orientation distribution function (ODF) for each pool.

The process described in Section 5.2 has been repeated for $d = 0$, providing new NRSS distribution functions. Fig. 15 compares these newly obtained distribution functions, based on the textured orientation pools, with those shown in Fig. 9(a). No significant differences are observed between the "high SF" and "random" pools, while the "S-Agg-3⁺" pool exhibits slightly lower NRSS values. The "extrusion" pool, however, shows notably lower NRSS values. Regarding the "extrusion" pool, the inverse pole figures (IPF) for the unit triangle, along with the Young's modulus and SF contour plots shown in Fig. 16, reveal that its orientations are concentrated around below-average Young's modulus values and low SF values. This distribution explains the lower NRSS observed. More generally, the lower NRSS in textured orientation pools can be linked to the orientation distributions within S-Agg-2⁺ configurations. Fig. 17 shows the inverse pole figures (IPF) of the orientations assigned to the central grain at $d = 0$, where stress is concentrated ([0 0 0]), and for three neighboring grains: the one positioned above ([0 0 2]), the one opposite the surface ([2 0 0]), and a side neighbor ([0 2 0]). Each position displays a unique preferred orientation that leads to high NRSS. The central grain prefers the hkl crystallographic direction [−3 8 11] aligned with the loading axis (Z), the neighbor opposite the surface aims towards [0 0 1], and the neighbor above towards [−1 1 1]. The side neighbor shows no preference, likely because this position corresponds to a surface half Kelvin cell, thus exerting less influence. Each position exhibits distinct orientation preferences, highlighting the necessity for a heterogeneous orientation pool to achieve high NRSS alignment across different relative positions.

As mentioned in [32], a single grain orientation can simultaneously increase stress in one neighbor while reducing it in another, emphasizing the need for orientation variety to achieve high NRSS. The strong textures limit the range of orientations configurations, reducing the likelihood of high stress concentrations.

It is important to note that this section serves as a hypothetical application of the CA model, whose parameters depend on the material macroscopic stiffness and are not calibrated for gradient properties. Therefore, while this analysis highlights that orientation diversity is key to reach high NRSS, further investigation is needed to fully understand the effect of texture gradients.

6. Conclusions

In an attempt to better understand the mechanisms governing crack localization within samples subjected to stress level typically below
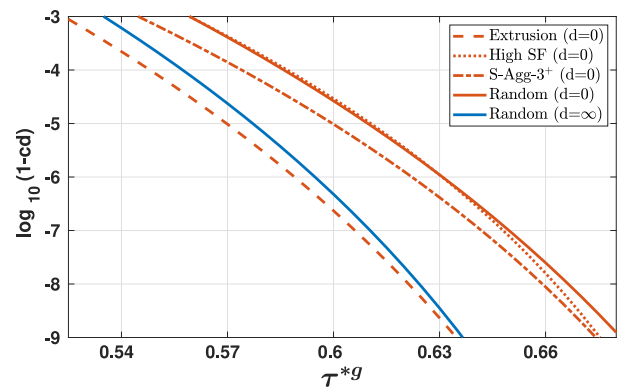

Fig. 15. Fitted distribution by means of the generalized Pareto function of τ^{*g} predicted by the CA model using different orientation pools. The scale $log_{10}(1 - cdf)$ is utilized to zoom in on the extreme values of τ^{*g} with very low probability.

the material elastic limit, an analytical model has been developed. Discrepancies in stress levels between grains located at different depths within a sample were investigated using a data-driven model known as the CA model. Initially developed by Bretin et al. [32], the model has been enhanced to incorporate surface effects in addition to the neighborhood effect. The aim of the CA model is to efficiently predict the grain mean stress tensor and NRSS based on the specific crystallographic orientations and depths of the grain and its neighbors. Due to its computational efficiency, the model can estimate in a few minutes the stress level of millions of grains within polycrystalline aggregates under any loading, while accounting for the specific crystallographic heterogeneities surrounding the grain. Specifically, this study focuses on the grain mean NRSS in single-phase polycrystals with an FCC crystal structure and high elastic anisotropy under linear-elastic loading conditions. The aggregates were represented using Kelvin's structure to minimize sources of heterogeneity, focusing primarily on discrepancies due to grains' crystallographic orientation and depth relatively to a free surface.

The NRSS under elastic loading was chosen as the primary focus, based on the assumption that the grain with the highest NRSS would plastify the fastest and thus be the most susceptible to initiate a crack first. This assumption, previously highlighted by Bretin et al. [17] for an infinite medium, remains relevant in the presence of an idealized free surface without any defects. The correlation NRSS — damage
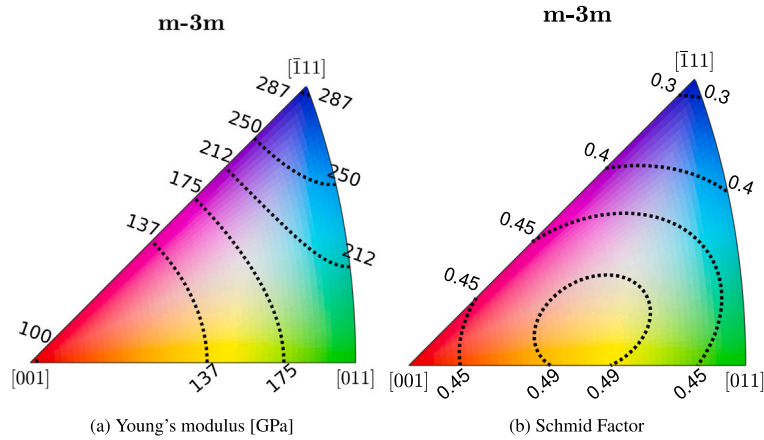
(a) Young's modulus [GPa]

(b) Schmid Factor

Fig. 16. Contour plots of the orientation Young's modulus and Schmid's factor along the direction of the inverse pole figure in the standard stereo-triangle.
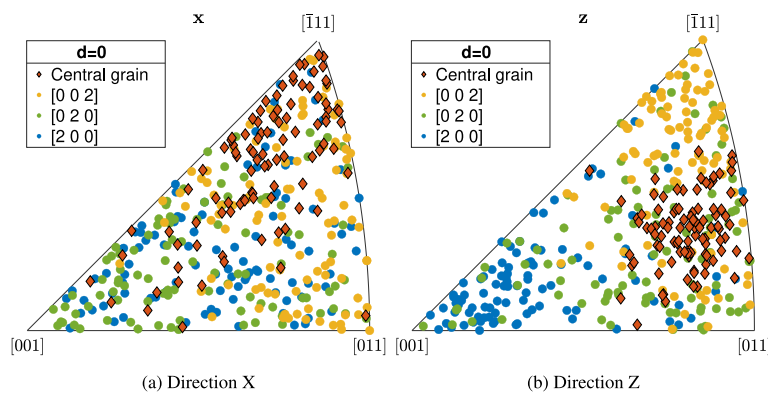

(a) Direction X

(b) Direction Z

Fig. 17. Inverse pole figure in the standard stereo-triangle of orientations from *S-Agg-2⁺* grain-neighborhood configurations. Only the grain located at $d = 0$ in which stress is concentrated and three of its neighbors are represented. One hundred orientations are displayed for each position.

rate shown by FEM simulations comforts the assumption even though some exception were found. For a given NRSS, surface grains display higher damage rates than in-depth grains, however a larger dataset is necessary to fully evaluate the extent of these differences. Nonetheless, the magnitude of this disparity seems negligible when compared to the variations of damage rate observed for different NRSS values. A quadratic function was fitted to the data to express the grain damage rate as a function of its NRSS, yielding a satisfactory regression coefficient.

To validate the accuracy of the updated CA model, comparisons were made between its predictions and those of the FE method. The model was tested on different types of aggregates: some with fully randomized orientations and others with high stress concentrations located at various depths. The model exhibited excellent accuracy in predicting both the grain mean stress tensors and NRSS. Even under conditions of high stress concentration, where stress levels were twice as high as the applied stress, the model maintained its accuracy. Some minor loss in accuracy was observed for surface grains, but it remained within an acceptable margin of error. Overall, for grains experiencing high stress levels, the NRSS was predicted within a margin of error of plus or minus 4% at a 95% confidence interval.

Despite the high accuracy of the CA model, its precision is comparable to the discrepancies observed between surface and in-depth grains. Leveraging the model's efficiency, the NRSS of millions of different grain-neighborhood configurations with random crystallographic orientations were predicted. Using this data, a probabilistic analysis of the highest NRSS within a mechanical part was conducted depending on its

geometry. The observed NRSS discrepancies between the highest NRSS from surface and in-depth grains within a typical fatigue test sample are approximately 5% of their expected value at most. This difference is not significant enough to establish a clear distinction between surface and in-depth grains.

Using textured orientation pools, it was shown that the highest expected NRSS value decreases as the pool becomes more textured. A wide variety of orientations is necessary for critical configurations to occur. Since surface grains are generally more textured than in-depth grains, these observations would suggest that in-depth grains are more likely to reach a higher NRSS value.

While the differences observed between in-depth and surface grains may not be significant enough to definitively determine whether a crack would initiate in-depth or on the surface, the CA model has proven to be an excellent tool for identifying probable configurations responsible for high stress concentration. Identifying these configurations is crucial, as they could lead to significant discrepancies in damage rates. For instance, the highest expected NRSS among millions of grains is approximately 25% higher than among a thousand. This disparity results in a damage rate 2.2 times faster in the former than in the latter, according to the regression function between NRSS and damage rate. The CA model effectively identifies configurations that lead to high stress levels at any grain depth, which can then be further analyzed using full-field models for deeper insights. Other criteria could also be added to the model such as slip transmission and blockage due to grain boundary misorientation between grains to distinguish grains with similar value of NRSS.
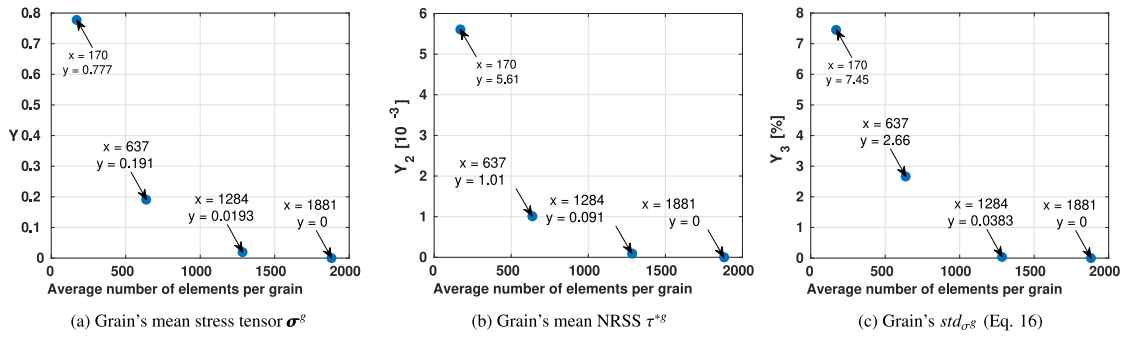
Fig. A.18. Evolution of the maximum differences of the studied variables FE predictions relatively to the reference finest mesh, depending on the mesh finesse. Convergence of the grains (a) stress tensor (b) NRSS, and (c) stress second moment.

Table B.7
Material parameters of the 316L steel [38,47].

E_y^{eff} (GPa)	v^{eff}		$\mathbb{C}_{1111}^{eff}$ (GPa)	$\mathbb{C}_{1122}^{eff}$ (GPa)		
196	0.280		251	97.9		
E_y^{cry} (GPa)	v^{cry}	G^{cry} (GPa)	$\mathbb{C}_{1111}^{cry}$ (GPa)	$\mathbb{C}_{1122}^{cry}$ (GPa)	$\mathbb{C}_{1212}^{cry}$ (GPa)	
100	0.3882	122	197	125	122	
K (MPa s^{-n})	n	r_0 (MPa)	Q (MPa)	b	c_χ (MPa)	d
12	11	40	10	3	40	1500
$h1$	$h2$	$h3$	$h4$	$h5$	$h6$	
1	1	0.6	12.3	1.6	1.8	

CRediT authorship contribution statement

R. Bretin: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **P. Bocher:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition.

Declaration of competing interest

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

Acknowledgments

Appendix A. Mesh convergence studies

To assess the level of mesh refinement required for the FE simulations presented in this article, simulations were conducted using different mesh sizes. One *S-Agg-1* and four *S-Agg-2* aggregates from Section 3.2 were selected for the convergence test. Four mesh sizes were tested: very coarse (170 elements per grain), coarse (637 elements per grain), medium (1284 elements per grain), and fine (1881 elements per grain). In Fig. A.18, the predictions obtained from each mesh are compared to those of a reference mesh, which in this case corresponds to the fine mesh. The convergence of three variables was studied.

• The grain mean stress tensor σ^g:

$$Y_1 = max\left(\frac{\|\sigma_{msh}^g - \sigma_{ref}^g\|_{eq}}{\|\Delta\sigma_{ref}^g\|_{eq}}\right)_g \quad (A.1)$$

• The grain mean NRSS τ_s^{*g}:

$$Y_2 = max\left(|\tau_{s\ msh}^{*g} - \tau_{s\ ref}^{*g}|\right)_{s,g} \quad (A.2)$$

• The stress standard deviation within a grain $std\sigma^g$ (Eq. (16)):

$$Y_3 = max\left(\frac{|std_{\sigma^g\ msh} - std_{\sigma\ ref}^g|}{std_{\sigma\ ref}^g}\right)_g \quad (A.3)$$

Based on these figures, it can be observed that the mesh used in this article (medium mesh with an average of 1284 elements per grain) demonstrates convergence of the studied variables towards the reference values within a reasonable margin of error.

Appendix B. Material constitutive law and parameters

The crystal stiffness tensor is assumed cubic (3 independent variables) and its values were taken from [38]. The effective elastic properties of the material were assumed to be isotropic, with a Young's modulus E_y^{eff} of 196 GPa and a Poisson's ratio v^{eff} of 0.280. These values were obtained through the homogenization of FE simulations of a polycrystalline aggregate RVE submitted to an elastic loading by following the method described in [20]. The stiffness tensors components are listed in Table B.7.

The kinematic and isotropic hardening from the slip rate defined in Eq. (17) are respectively defined as follow:

$$\chi_s^g = c_\chi \alpha_s^g \quad (B.1a)$$

with $\quad \dot{\alpha}_s^g = \dot{\gamma}_s^g - d\alpha_s^g|\dot{\gamma}_s^g| \quad (B.1b)$

$$r_s^g = r_0 + Qb\sum_{u=1}^{12} h_{su}\rho_u^g \quad (B.1c)$$

with $\quad \dot{\rho}_s^g = (1 - b\rho_s^g)|\dot{\gamma}_s^g| \quad (B.1d)$

where ρ_s^g and α_s^g are internal state variables; b and Q are material parameters characterizing the isotropic hardening; d and c_χ are material parameters characterizing the kinematic hardening. The hardening matrix $\boldsymbol{h}$ reflects the self-hardening and the latent hardening between slip systems. All these parameters are listed in Tables B.7–B.8.

Data availability

Data will be made available on request.

Table B.8
Slip system list and their hardening matrix.

Plane	Direction	System #	A2	A3	A6	B2	B4	B5	C1	C3	C5	D1	D4	D6
$(\bar{1}11)$	$[0\bar{1}1]$	A2	h_1	h_2	h_2	h_4	h_5	h_5	h_3	h_5	h_6	h_3	h_6	h_5
$(\bar{1}11)$	$[101]$	A3	h_2	h_1	h_2	h_5	h_3	h_6	h_5	h_4	h_5	h_6	h_3	h_5
$(\bar{1}11)$	$[110]$	A6	h_2	h_2	h_1	h_5	h_6	h_3	h_6	h_5	h_3	h_5	h_5	h_4
(111)	$[0\bar{1}1]$	B2	h_4	h_5	h_5	h_1	h_2	h_2	h_3	h_6	h_5	h_3	h_5	h_6
(111)	$[\bar{1}01]$	B4	h_5	h_3	h_6	h_2	h_1	h_2	h_6	h_3	h_5	h_5	h_4	h_5
(111)	$[1\bar{1}0]$	B5	h_5	h_6	h_3	h_2	h_2	h_1	h_5	h_5	h_4	h_6	h_5	h_3
$(11\bar{1})$	$[011]$	C1	h_3	h_5	h_6	h_3	h_6	h_5	h_1	h_2	h_2	h_4	h_5	h_5
$(11\bar{1})$	$[101]$	C3	h_5	h_4	h_5	h_6	h_3	h_5	h_2	h_1	h_2	h_5	h_3	h_6
$(11\bar{1})$	$[1\bar{1}0]$	C5	h_6	h_5	h_3	h_5	h_5	h_4	h_2	h_2	h_1	h_5	h_6	h_3
$(1\bar{1}1)$	$[011]$	D1	h_3	h_6	h_5	h_3	h_5	h_6	h_4	h_5	h_5	h_1	h_2	h_2
$(1\bar{1}1)$	$[\bar{1}01]$	D4	h_6	h_3	h_5	h_5	h_4	h_5	h_5	h_3	h_6	h_2	h_1	h_2
$(1\bar{1}1)$	$[110]$	D6	h_5	h_5	h_4	h_6	h_5	h_3	h_5	h_6	h_3	h_2	h_2	h_1

References

[1] McDowell D, Gall K, Horstemeyer M, Fan J. Microstructure-based fatigue modeling of cast A356-T6 alloy. Eng Fract Mech 2003;70(1):49–80.

[2] Ishihara S, McEvily A. On the early initiation of fatigue cracks in the high cycle regime. In: Proceedings of the 12th international conference on fracture, Ottawa, ON, Canada. 2009, p. 12–7.

[3] McDowell D, Dunne F. Microstructure-sensitive computational modeling of fatigue crack formation. Int J Fatigue 2010;32(9):1521–42.

[4] Castelluccio GM, McDowell DL. Mesoscale modeling of microstructurally small fatigue cracks in metallic polycrystals. Mater Sci Eng A 2014;598:34–55.

[5] Marines-Garcia I, Paris PC, Tada H, Bathias C. Fatigue crack growth from small to long cracks in very-high-cycle fatigue with surface and internal "fish-eye" failures for ferrite-perlitic low carbon steel SAE 8620. Mater Sci Eng A 2007;468:120–8.

[6] Hong Y, Lei Z, Sun C, Zhao A. Propensities of crack interior initiation and early growth for very-high-cycle fatigue of high strength steels. Int J Fatigue 2014;58:144–51.

[7] Lei Z, Xie J, Sun C, Hong Y. Effects of loading condition on very-high-cycle fatigue behaviour and dominant variable analysis. Sci China Phys Mech Astron 2014;57(1):74–82.

[8] McDowell DL. Basic issues in the mechanics of high cycle metal fatigue. Int J Fract 1996;80(2):103–45.

[9] Mukherjee S, Kundu A, De PS, Mahato JK, Chakraborti P, Shome M, Bhattacharjee D. Insitu investigation of tensile deformation behaviour of cold-rolled interstitial-free high-strength steel in scanning electron microscope. Mater Sci Eng A 2020;776:139029.

[10] Stinville JC, Echlin MP, Texier D, Bridier F, Bocher P, Pollock TM. Sub-grain scale digital image correlation by electron microscopy for polycrystalline materials during elastic and plastic deformation. Exp Mech 2016;56(2):197–216.

[11] Berger A, Witz J-F, El Bartali A, Sadat T, Limodin N, Dubar M, Najjar D. Experimental investigation of early strain heterogeneities and localizations in polycrystalline α-Fe during monotonic loading. Int J Plast 2022;153:103253.

[12] Brenner R, Lebensohn RA, Castelnau O. Elastic anisotropy and yield surface estimates of polycrystals. Int J Solids Struct 2009;46(16):3018–26.

[13] Bennett V, McDowell D. Polycrystal orientation distribution effects on microslip in high cycle fatigue. Int J Fatigue 2003;25(1):27–39.

[14] Sweeney C, Vorster W, Leen S, Sakurada E, McHugh P, Dunne F. The role of elastic anisotropy, length scale and crystallographic slip in fatigue crack nucleation. J Mech Phys Solids 2013;61(5):1224–40.

[15] Phung N, Favier V, Ranc N. Evaluating Schmid criterion for predicting preferential locations of persistent slip markings obtained after very high cycle fatigue for polycrystalline pure copper. Int J Fatigue 2015;77:115–27.

[16] Gehrig F, Wicht D, Krause M, Böhlke T. FFT-based investigation of the shear stress distribution in face-centered cubic polycrystals. Int J Plast 2022;157:103369.

[17] Bretin R, Levesque M, Pilvin P, Bocher P. Neighborhood effect within polycrystalline materials: Elastoplastic micromechanical analysis by cellular automaton and finite element. Int J Fatigue 2020;137:105634.

[18] Tanaka K, Mura T. A dislocation model for fatigue crack initiation. J Appl Mech 1981;48(1):97–103.

[19] Sauzay M. Cubic elasticity and stress distribution at the free surface of polycrystals. Acta Mater 2007;55(4):1193–202.

[20] Bretin R, Levesque M, Bocher P. Neighborhood effect on the strain distribution in linearly elastic polycrystals: Part 1 – finite element study of the interaction between grains. Int J Solids Struct 2019;176–177:36–48.

[21] Barbe F, Decker L, Jeulin D, Cailletaud G. Intergranular and intragranular behavior of polycrystalline aggregates. Part 1: F.E. model. Int J Plast 2001;17(4):513–36.

[22] Guilhem Y, Basseville S, Curtit F, Stephan J-M, Cailletaud G. Investigation of the effect of grain clusters on fatigue crack initiation in polycrystals. Int J Fatigue 2010;32(11):1748–63.

[23] Roters F, Eisenlohr P, Bieler T, Raabe D. Crystal plasticity finite element methods: in materials science and engineering. Wiley; 2011.

[24] Lavergne F, Brenner R, Sab K. Effects of grain size distribution and stress heterogeneity on yield stress of polycrystals: A numerical approach. Comput Mater Sci 2013;77:387–98.

[25] Nguyen T-K, Sab K, Bonnet G. Green's operator for a periodic medium with traction-free boundary conditions and computation of the effective properties of thin plates. Int J Solids Struct 2008;45(25):6518–34.

[26] Gélébart L. A simple extension of FFT-based methods to strain gradient loadings - application to the homogenization of beams and plates with linear and non-linear behaviors. J Theor Comput Appl Mech 2022.

[27] Hill R. A self-consistent mechanics of composite materials. J Mech Phys Solids 1965;13(4):213–22.

[28] Lebensohn RA, Liu Y, Ponte Castañeda P. On the accuracy of the self-consistent approximation for polycrystals: comparison with full-field numerical simulations. Acta Mater 2004;52(18):5347–61.

[29] Mareau C, Berbenni S. An affine formulation for the self-consistent modeling of elasto-viscoplastic heterogeneous materials based on the translated field method. Int J Plast 2015;64:134–50.

[30] Liu Z, Bessa MA, Liu WK. Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials. Comput Methods Appl Mech Engrg 2016;306:319–41.

[31] Krause M, Böhlke T. Maximum-entropy based estimates of stress and strain in thermoelastic random heterogeneous materials. J Elasticity 2020;141(2):321–48.

[32] Bretin R, Levesque M, Bocher P. Neighborhood effect on the strain distribution in linearly elastic polycrystals: Part 2 – cellular automaton. Int J Solids Struct 2019;176–177:49–67.

[33] Petr L. Fatigue crack nucleation and microstructure. ASM Handb 1996;96–109.

[34] Liu C, Zhao M-C, Zhao Y-C, Zhang L, Yin D-F, Tian Y, Shan Y-Y, Yang K, Atrens A. Ultra-high cycle fatigue behavior of a novel 1.9 GPa grade super-high-strength maraging stainless steel. Mater Sci Eng A 2019;755:50–6.

[35] Tian J, Wang W, Li H, Yang K, Jiang Z. Understanding main factors controlling high cycle fatigue crack initiation and propagation of high strength maraging stainless steels with Ti addition. Mater Sci Eng A 2021;805:140589.

[36] Sauzay M, Jourdan T. Polycrystalline microstructure, cubic elasticity, and nucleation of high-cycle fatigue cracks. Int J Fract 2006;141(3–4):431–46.

[37] Barbe F, Forest S, Cailletaud G. Intergranular and intragranular behavior of polycrystalline aggregates. part 2: Results. Int J Plast 2001;17(4):537–63.

[38] Guilhem Y, Basseville S, Curtit F, Stephan J, Cailletaud G. Numerical investigations of the free surface effect in three-dimensional polycrystalline aggregates. Comput Mater Sci 2013;70:150–62.

[39] Xu X, Lunt D, Thomas R, Babu RP, Harte A, Atkinson M, da Fonseca J, Preuss M. Identification of active slip mode in a hexagonal material by correlative scanning electron microscopy. Acta Mater 2019;175:376–93.

[40] Méric L, Poubanne P, Cailletaud G. Single crystal modeling for structural calculations: Part 1—Model presentation. J Eng Mater Technol 1991;113(1):162–70.

[41] Altmann S. Rotations, quaternions, and double groups. Dover Publications; 2005.

[42] Shoemake K. Uniform random rotations. In: Graphics gems III (IBM version). Elsevier; 1992, p. 124–32.

[43] Coles S. An introduction to statistical modeling of extreme values. Springer; 2014.

[44] MATLAB. Version 9.9.0 (R2020b). Natick, Massachusetts: The MathWorks Inc.; 2020.

[45] Mineur M, Villechaise P, Mendez J. Influence of the crystalline texture on the fatigue behavior of a 316L austenitic stainless steel. Mater Sci Eng A 2000;286(2):257–68.

[46] Tridello A, Niutta CB, Berto F, Paolino D. Size-effect in very high cycle fatigue: A review. Int J Fatigue 2021;153:106462.

[47] Huntington H. The elastic constants of crystals. 7, 1958, p. 213–351,