Image Anal Stereol 2022;41:25-39 doi: 10.5566/ias.2642
Original Research Paper

STATISTICAL ASSESSMENT OF STRESS REDISTRIBUTION IN LOADED POLYCRYSTALS

IVA KARAFIÁTOVÁ✉,[1], ZBYNĚK PAWLAS[1] AND LUDĚK HELLER[2]

[1]Department of Probability and Mathematical Statistics, Faculty of Mathematics and Physics, Charles University, Sokolovská 83, 18675 Praha 8, Czech Republic, [2]Institute of Physics of the Czech Academy of Sciences, Na Slovance 2, 18221 Praha 8, Czech Republic
e-mail: karafiatova@karlin.mff.cuni.cz, pawlas@karlin.mff.cuni.cz, heller@fzu.cz
(Received November 16, 2021; revised January 11, 2022; accepted January 11, 2022)

ABSTRACT

This work deals with the analysis of stress redistribution in a polycrystal due to external loading, anisotropy of elastic properties, and microstructure characteristics. A statistical method that enables assessing relationships between stress fields and microstructure features of interest is suggested. The notion of generalised semivariogram is introduced and used to determine the extent of spatial dependence in multivariate random fields. Afterwards, it is allowed to perform the tests of independence based on the distance correlation coefficient. The detected non-spatial dependencies are further examined, focusing on the identification of the actual type of heteroscedasticity. The method is aimed at analysing large computational datasets resulting from numerical simulations of stress redistribution in polycrystals under external loads. It is demonstrated on datasets computed on a realistic microstructure of a NiTi wire subjected to tension while considering uniform and preferential lattice orientation distributions and various degrees of elastic anisotropy. The method shows for the considered microstructure and loading that the degree of elastic anisotropy does not affect the dependencies contrarily to the lattice orientation distribution.

Keywords: distance correlation, elastic anisotropy, generalised semivariogram, heteroscedasticity, lattice disorientation, test of independence.

INTRODUCTION

The deformation behaviour of polycrystalline metals and alloys is driven by the evolution of the internal stress field in response to external loads. The internal stress field is proportional to the stretch and the distortion of interatomic bonds is described at the continuum level by a second-order strain tensor $\varepsilon_{ij}\,(i,j=1,2,3)$ that is related to the second-order stress tensor σ_{kl} $(k,l=1,2,3)$ through the elastic constants forming a fourth-order elasticity tensor $C_{ijkl}\,(i,j,k,l=1,2,3)$. The stress-strain relationship is formulated by the following linear relationship defined in the space of the crystal lattice as

$$\sigma_{ij}=\sum_{k=1}^{3}\sum_{l=1}^{3}C_{ijkl}\varepsilon_{kl},\quad i,j=1,2,3. \tag{1}$$

The components of the strain/stress tensors $\varepsilon_{ij}/\sigma_{ij}$ denote strain/stress acting on the plane normal to the direction i along the direction j of the chosen coordinate system. This can be, for instance, the Cartesian coordinate system for which the integer indices $1,2,3$ relate to axes x,y,z, respectively, or the cylindrical coordinate system for which the integer indices $1,2,3$ relate to axes r,φ,z.

The stress-strain fields within a polycrystal are ruled by the deformation interactions among individual crystals, called grains, constituting the polycrystals. To simulate and understand these interactions, the polycrystals are modelled by tessellation-based methods that can be supported by experimental data to generate realistic microstructure models. For example, the 3DXRD experimental method (Sedmák *et al.*, 2016) providing information on microstructure features in terms of grain volumes, grains' centres of mass, and grain-wise averaged strains and stresses (Fig. 1 (b)) have been used to reconstruct a realistic microstructure model of a wire made from a NiTi shape memory alloy (Otsuka and Wayman, 1998; Heller *et al.*, 2020) (Fig. 1 (a)) being used in biomedical applications. Such microstructure models represented by a set of convex polyhedra (Fig. 1 (c)) are used to simulate stress-strain fields using numerical methods such as the Finite Element Method (FEM) (Barbe *et al.*, 2001). FEM is applied to solve the problem of internal force equilibrium formulated by a set of partial differential equations including boundary conditions representing the loading conditions and deformation constraints (Zienkiewicz *et al.*, 2013). The FEM assumes a polynomial evolution of the displacement field on the space of individual elements created by the meshing of individual grains (Fig. 1 (d)). The displacements at the elements' vertices are

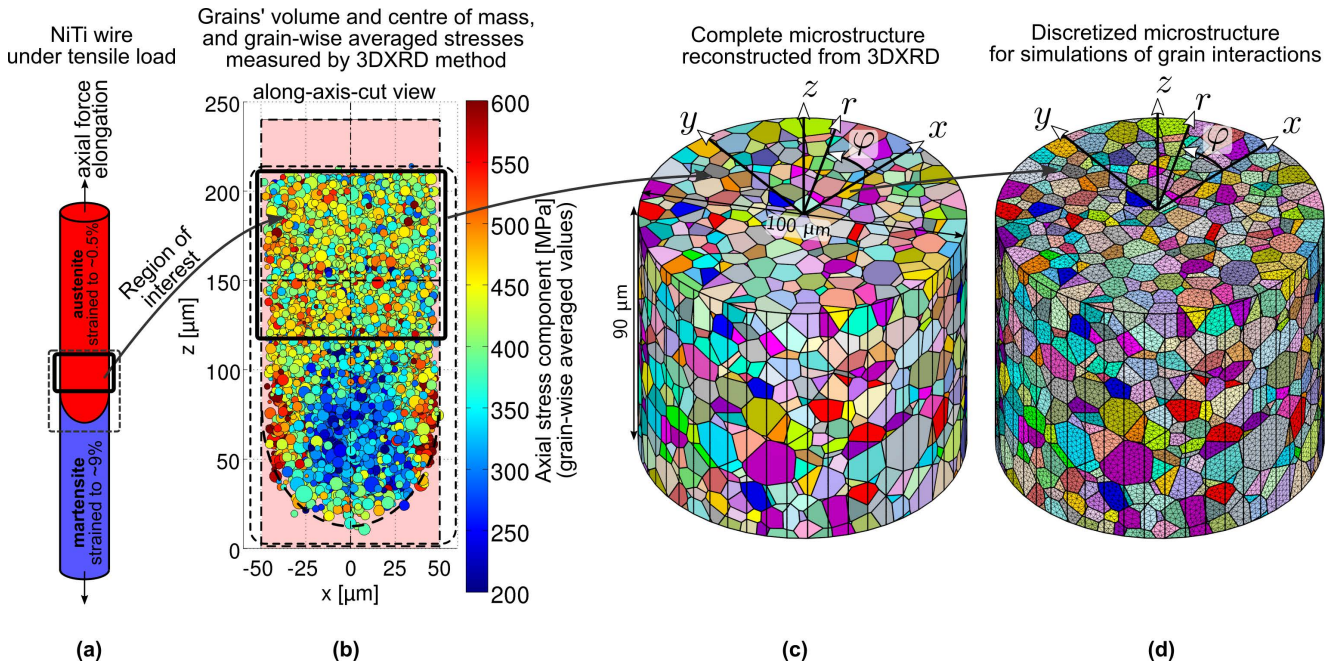


Fig. 1: Example of microstructure reconstruction and modelling used to simulate stress-strain fields in polycrystals subjected to external loads published in Heller *et al.* (2020). The microstructure of a NiTi wire (a) was reconstructed from experimental data provided by the 3DXRD method (b) using tessellation modelling (Petrich *et al.*, 2019) resulting in a set of convex polyhedra (c) that are further discretised by finite elements (d) for purposes of numerical simulations of stress-strain fields.

the solution of the simulation, while the stress-strain fields are calculated from the displacement fields and Eq. (1). Besides the tessellation-based methods, less computationally demanding self-consistent modelling methods are being used to estimate the effects of crystal orientations on the distribution of stresses and strains in loaded polycrystals (Willot *et al.*, 2020).

The stress-strain redistribution inside the externally loaded polycrystal stems from deformation processes in individual grains that depend on the orientation of the loading with respect to the crystal lattice, which is in each of the grains oriented differently with respect to the fixed sample's coordinate system. This is illustrated in Fig. 2 (a), where the orientation of the cubic lattice is depicted in each of the grains of a longitudinal cross-section of the microstructure model of the NiTi wire shown in Fig. 1 (c). The lattice orientation is further visualised by grain colours reflecting the orientation of the wire axis in the grains' cubic lattices according to the colour coding explained in Fig. 2 (b). The blue shades prevail, meaning that the crystals are preferentially oriented so that the wire axis lies near the body diagonal of the crystal lattice.

The mismatch between elastic properties of neighbouring grains is determined by the anisotropy of elastic properties and the distribution of disorientations (smallest misorientation) between grains. For materials with the cubic crystal lattice, such as the one of the present study, the elasticity tensor C_{ijkl} (Eq. (1)) is determined by three independent constants defined in the coordinate system attached to the basal direction of the cubic lattice

$$C_{11} = C_{iiii}, \quad C_{12} = C_{iijj} = C_{jjii}, \quad C_{44} = C_{ijij} = C_{jiji},$$

$i, j = 1,2,3,\ i \neq j$. The stress-strain relationship (Eq. (1)) can be expressed in matrix form using Voigt notation as follows

$$\begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sigma_{23} \\ \sigma_{13} \\ \sigma_{12} \end{pmatrix} = \begin{pmatrix} C_{11} & C_{12} & C_{12} & 0 & 0 & 0 \\ & C_{11} & C_{12} & 0 & 0 & 0 \\ & & C_{11} & 0 & 0 & 0 \\ & sym. & & C_{44} & 0 & 0 \\ & & & & C_{44} & 0 \\ & & & & & C_{44} \end{pmatrix} \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{22} \\ \varepsilon_{33} \\ 2\varepsilon_{23} \\ 2\varepsilon_{13} \\ 2\varepsilon_{12} \end{pmatrix},$$

where components of stress and strain tensors form two vectors and the elasticity tensor is substituted by 6×6 symmetric stiffness matrix determined by the three constants in the case of the cubic crystal. The extent of anisotropy of the cubic crystal is defined by the coefficient of elastic anisotropy A that is derived from the three independent constants as the ratio of the shear stiffness in the basal plane, C_{44}, to the shear stiffness in the cube diagonal plane, $C' = (C_{11} - C_{12})/2$, i.e., $A = C_{44}/C'$ (Zener , 1948).

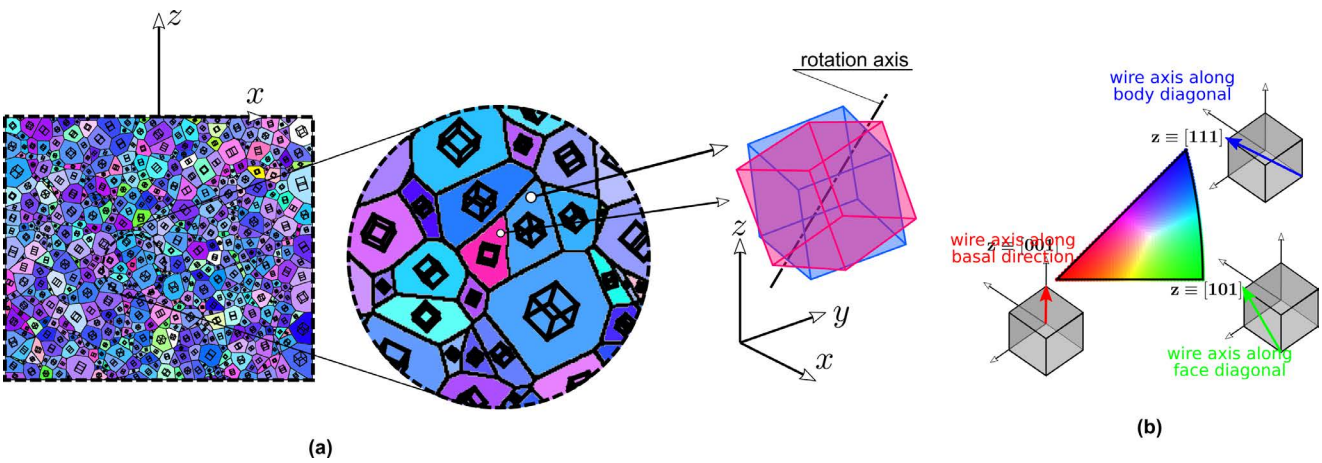


Fig. 2: Visualisation of lattice orientations in individual grains (a) by planar projections of cubic lattices and by colours selected according to the colour coding (b).

The disorientation between grains is another factor contributing to the mismatch in grain stiffnesses. The distribution of disorientations is determined by the distribution of crystal lattice orientations (Bunge, 1969). Besides a uniform distribution of crystal lattice orientations inherited from the stochastic process of metal solidification, materials may possess preferential lattice orientations due to the thermomechanical processing applied after the solidification. For example, the NiTi wire is produced by hot rolling and cold drawing involving a large, lattice orientation dependent, plastic deformation process. Such a process results in cubic lattice rotations such that their body diagonals align preferentially along the wire axis direction as shown in Fig. 3 (b1). This so-called texture leads to a different disorientation angle distribution (Fig. 3 (c1)) in comparison to uniformly distributed lattice orientations (Fig. 3 (b2)) that result in the disorientation distribution shown in Fig. 3 (c2). The textured material possesses a lower frequency of high disorientation angles that leads to a lower scatter of axial stresses, as evident by comparing the axial stress fields of the textured material (Fig. 3 (d1)) and the material with the uniform distribution of lattice orientations (Fig. 3 (d2)). The distribution of grain orientations determines macroscopic elastic properties of polycrystals (Hirsekorn, 1990).

Additionally to microstructure parameters, including material parameters (e.g., elastic anisotropy) and crystal lattice orientation, the redistribution of stresses and strains in polycrystals is driven by microstructure morphology (grains' size, shape, and grain boundaries). The knowledge of the relationship between microstructure parameters and stress-strain redistribution enables us to understand and predict the material behaviour or optimise the materials' performance.

The redistribution of the stress-strain fields is usually characterised by averages and standard deviations measured within a representative volume element (RVE) originally introduced in Hill (1963). In simple words, RVE is defined such that these characteristics do not change further when increasing the observation volume beyond that of RVE. However, the variables in two adjacent RVEs could be correlated. Therefore, the notion of an uncorrelated volume element (UVE) was introduced in Sanei and Fertig (2015). The average (nominal) values of stress and strain characterise the effective macroscopic deformation properties, while their standard deviations characterise the local extremes of stress and strain. While the global effective properties are considered in structural engineering, the standard deviations are considered for the prediction of critical stress states for the initiation of other deformation properties. Hence, both statistical characteristics of the stress-strain fields are relevant for the materials' performance.

In this paper, we suggest such an identification method that searches for both spatial and non-spatial dependencies. First, the identification of UVE is presented. It empowers us to investigate non-spatial dependencies on sets of nodes that are at distances larger than the range of UVE, thus eliminating the local effects. Second, we test if considered microstructure parameters affect the stress distribution within grains and/or at grain boundaries. The grain and grain-boundary stress distributions are assessed by average and standard deviation of stresses within individual grains and grain boundaries, respectively. The distribution of average stresses within individual grains thus expresses the distribution of inter-granular stresses, while the distribution of standard deviations of stresses within individual grains expresses the distribution of intra-granular

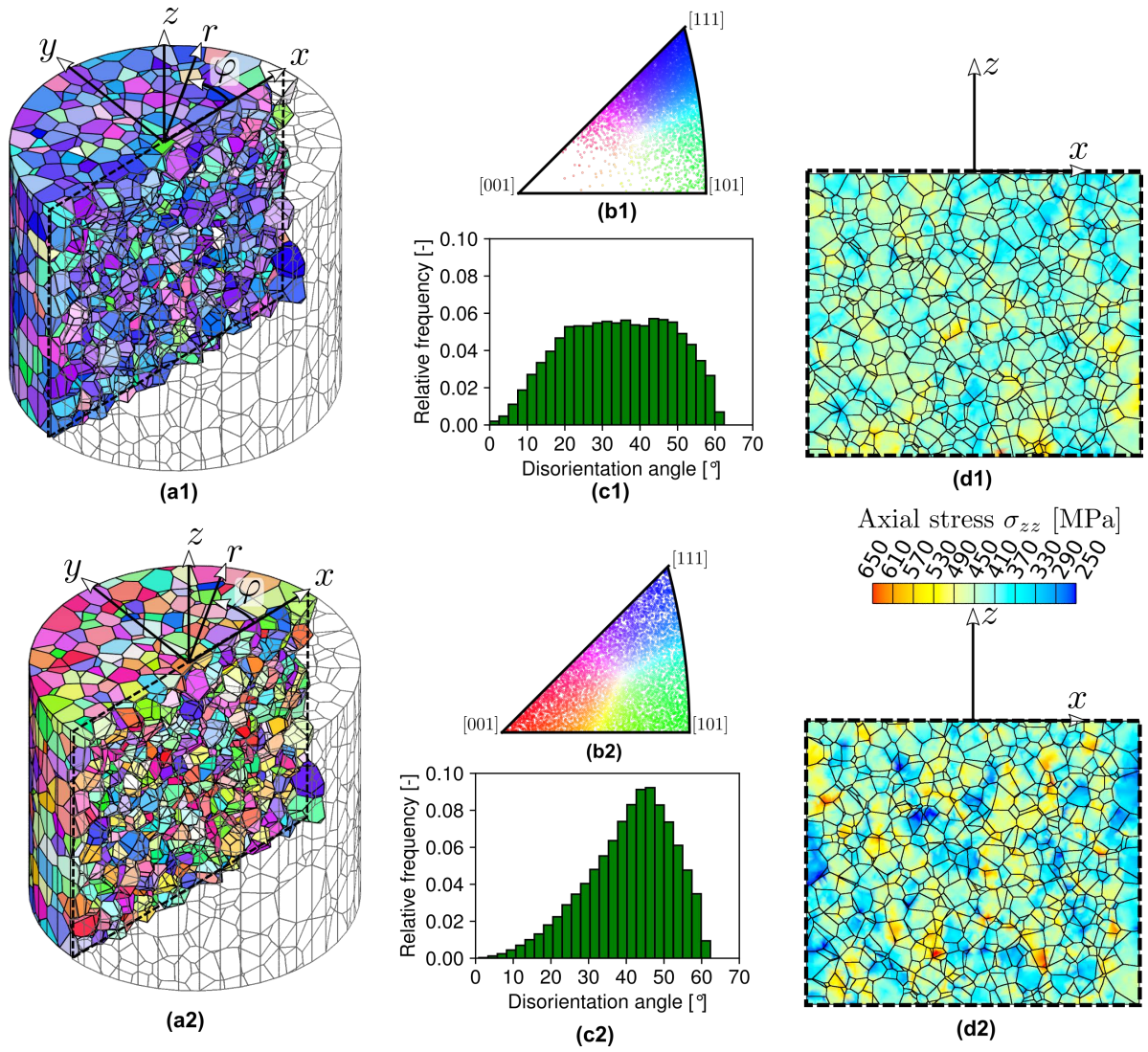


Fig. 3: Effect of the preferential lattice orientation on the stress field. The behaviour of σ_{zz} (d1, d2) differs for the textured material (a1, b1, c1) and the material having a uniform distribution of crystal orientation (a2, b2, c2).

stresses. The grain-boundary stress assesses local stresses using the same statistical parameters that are evaluated at the grain boundaries. Formal statistical tests are used to evaluate the dependencies of grain and grain-boundary stress distributions on microstructure parameters. Third, in case of rejecting the independence hypothesis, we further suggest investigating whether the data are heteroscedastic, i.e., whether grain to grain or boundary to boundary fluctuations of the stress distribution increase or decrease with the increase of a microstructure parameter.

MATERIAL

In general, the problem at hand relates to the broader issue of material-property relationship seeking for the identification of relationships between microstructure properties and macroscopic deformation behaviour in order to optimise currently used materials or design new materials (Fullwood *et al.*, 2010; Adams *et al.*, 2013; Brough *et al.*, 2017). Specifically, we are interested in statistical methods that asses the effect of microstructure on distributions of inter- and intra-granular stresses and local stresses near grain boundaries.

The method is tested on simulation data published in Heller *et al.* (2020). The dataset is related to a single microstructure tessellation model approximating a real NiTi wire microstructure loaded by tension along the wire axis. The approximation was performed by using Laguerre tessellation fitted to experimental data collected in a 3DXRD experiment, providing information on grain volumes, centres of mass positions, and lattices' orientation as

schematically shown in Fig. 1. The method was applied to assess the effect of preferential lattice orientation, the magnitude of elastic anisotropy, and grain boundary orientation. Therefore, five datasets were used. Two datasets account for a very low elastic anisotropy and different lattice orientation distributions compared in Fig. 3: 1) experimentally identified preferential orientation of [111] lattice directions along the wire axis, denoted as *texture*; 2) uniform lattice orientation distribution, denoted as *uniform*. Three other datasets account for the experimentally identified preferential lattice orientation distribution and three different elastic anisotropies: *low*, *mid* and *high*, obtained by different choices of C_{ijkl} in Eq. (1). Table 1 provides an overview of the datasets used in this work, including their abbreviations used hereinafter.

Table 1: Overview of used datasets including their abbreviations, where $A = 2C_{44}/(C_{11} - C_{12})$.

	Lattice ori.	Elastic constants			A
		C_{11}	C_{12}	C_{44}	
uniform	uniform	169.00	141.00	33.00	2.36
texture	preferred 111	169.00	141.00	33.00	2.36
low	preferred 111	162.45	145.45	34.00	4.00
mid	preferred 111	159.00	147.63	35.00	6.16
high	preferred 111	152.90	143.50	35.40	7.53

From a mathematical point of view, we deal with a tessellation (Fig. 1 (c)) with multiple mark sets with values in multidimensional metric spaces. Similar marked tessellation was also studied in Pawlas *et al.* (2020) to test the independent marking of a tessellation marked with crystallographic orientations. In this paper, one set of marks, which is considered to be the set of response variables, is formed by the tensor of mean values ($Stress_g$ *[avg]*), the distribution of which characterizes the inter-granular stress fluctuations, and the tensor of standard deviations of the stress components ($Stress_g$ *[std]*), which characterizes the intra-granular stress fluctuations. Both these tensors are calculated from all elements constituting the grain. For the local stresses near grain boundaries, we again consider the two stress tensors. However, in this case, we compute the mean ($Stress_b$ *[avg]*) and the standard deviation ($Stress_b$ *[std]*) only from the elements lying on the grain boundaries. As the explanatory variables, we consider the following set of microstructure features:

– *Lattice ori* – lattice orientation described by a unit quaternion,

– *Load ori* – deviation angle of the [111] crystallographic direction of a grain from the loading axis,

– *Dis [...]* – disorientation angles with neighbouring grains (mean, minimum, maximum, and standard deviation),

– *Volume* – volume of a grain normalised by the average volume within the sample.

Regarding the grain boundary microstructure, we use these variables:

– *Area* – size of the boundary,

– *Boundary ori* – three-component normal vector describing the tilt of the grain boundary,

– *Dis angle* – disorientation angle between the pair of neighbouring grains,

– *Vol ratio* – ratio of the volume of the larger grain to the volume of the smaller one sharing the grain boundary.

The geometrical characteristics are computed based on the parametric representation of the tessellation.

METHODS

To determine the character of the desired relationships, we use various approaches and distinguish spatial and non-spatial dependence. First, the spatial dependence of the stress field needs to be filtered out. This can be achieved by the usage of the random field and the generalised semivariogram applied to a suitably large subsample of elements, which empowers the determination of UVE. Once the UVE is determined, further analysis follows, which proceeds only with the grain and boundary variables. For a spatially independent subsample of grains, the tests of independence between the stress values and microstructure variables are applied. For pairs for which the independence hypothesis is rejected, heteroscedasticity is examined.

SPATIAL DEPENDENCE

When dealing with image data of polycrystalline materials, spatial dependencies are often present. Thus, for further analysis, it is essential to determine the *range* of spatial dependence, which is the distance beyond which the data are independent. The range determines the size of the UVE. More precisely, the UVE is estimated by a ball with a diameter equal to the estimated range. This plays an important role because it enables us to select spatially uncorrelated subsamples.

Spatially distributed data on $D \subseteq \mathbb{R}^3$ are modelled by a spatial random field, which is a collection of random variables $\{Z(x),\ x \in D\}$ with, most frequently, real values. Let $\{Z(x),\ x \in D\}$ be a stationary random field. To express the spatial dependence of the random field, it is common to use the so-called *semivariogram*, defined as

$$\gamma(h) = \frac{1}{2}\mathbb{E}\big[Z(x+h) - Z(x)\big]^2, \quad h \in D - D, \qquad (2)$$

where $D - D = \{h : h = y - x, x \in D, y \in D\}$. If there exists a finite limit $\lim_{\|h\|\to\infty} \gamma(h)$ called the *sill*, we define the *range* C as the minimal value for which the semivariogram hits its sill, i.e.,

$$C = \min\left\{ s \geq 0 : \gamma(x) = \lim_{\|h\|\to\infty} \gamma(h),\ \text{for all } \|x\| \geq s \right\}.$$

The range indicates the distance within which values at two points in D are spatially correlated. In geostatistics, it is common to define the *practical range* (or *effective range*) as the distance at which the semivariogram achieves 95% of its sill, see (Schabenberger and Gotway, 2005, Section 1.4.2).

However, this approach is used only for spatially distributed data with real values. A generalisation to multidimensional data is not so often considered in the literature. Such a generalisation is needed for our purposes due to the desire to assess the spatial correlation of the stress tensor values. More specifically, let us consider a stationary spatial random field with values in a general metric space $(\mathbb{X}, d)$ with dimension higher than one. We suggest defining the *generalised semivariogram* as

$$\gamma(h) = \frac{1}{2}\mathbb{E}\big[d\,(Z(x+h), Z(x))\big]^2, \quad h \in D - D. \qquad (3)$$

It is easily seen that the semivariogram defined by Eq. (2) is a special case of Eq. (3) on the Euclidean space $\mathbb{X} = \mathbb{R}$ with $d(x,y) = |x - y|$. Another method of measuring spatial correlations on multidimensional spaces is used, e.g., in Abdallah *et al.* (2015).

Let $\{Z(x),\ x \in D\}$ be a motion-invariant (i.e., stationary and isotropic) random field with values in $\mathbb{X}$. It means that the distribution of distances $d\,(Z(x), Z(y))$ depends only on the distance $\|x - y\|$ between the corresponding ground points. Following Matheron's estimator (Schabenberger and Gotway, 2005, Section 4.4.1), we define a non-parametric estimator of $\gamma(r)$ for $r \geq 0$ as

$$\hat{\gamma}(r) = \frac{1}{2|N(r)|} \sum_{x,y \in N(r)} \Big[d\,(Z(x), Z(y)) \Big]^2,$$

where $N(r) = \{(x,y) \in D \times D : \|x - y\| = r\}$. However, since for fixed $r \geq 0$, the set $N(r)$ contains very little or no pair, we divide the interval $[0, R)$, $R > 0$, into k equidistant bins of width $\delta = R/k$. Then γ is estimated for representing points $r_i = (i - 1/2)\delta$ by

$$\hat{\gamma}(r_i) = \frac{1}{2|N_i|} \sum_{x,y \in N_i} \Big[d\,(Z(x), Z(y)) \Big]^2, \quad i = 1, \ldots, k, \quad (4)$$

where $N_i = \{(x,y) \in D \times D : \|x - y\| \in [r_i - \delta/2, r_i + \delta/2)\}$.

We intend to determine the range of spatial dependence so that the data are correlated if the ground points are within this distance and the correlation is negligible outside it. From $\hat{\gamma}$ given by Eq. (4), we can naturally estimate the sill and the practical range. Then we can define the UVE as a ball with a diameter equal to the estimated practical range.

To apply this procedure to the stress field, the values of element-wise stress tensor can be viewed as a realisation of a spatial random field. We focus only on the cubic structure of a polycrystalline material. The corresponding stress tensor (σ_{kl}) $(k,l = 1,2,3)$ can then be represented as a real symmetric 3×3 matrix. Thus, we consider a spatial random field with ground points in D and values in the metric space $(\mathbb{R}^{3\times3}, d_1)$ with the ℓ_1-metric d_1 defined as

$$d_1(x,y) = \sum_{i,j=1}^{3} |x_{ij} - y_{ij}|, \quad x, y \in \mathbb{R}^{3\times3}. \qquad (5)$$

This metric depends on the choice of the basis. However, the present study deals with a simple uniaxial loading along the axis of a cylindrical sample. Hence the natural choice of the cylindrical coordinate system (z-axis along the loading direction being the wire axis) points its z-axis along the average orientation of the principal stress. Therefore, the metric given by Eq. (5) can be considered as very close to a metric operating with principal stresses. The choice of the metric for more general types of loading and stress states requires further investigations. One can consider the distances discussed in Angulo (2014) or Yamaji and Sato (2006).

NON-SPATIAL DEPENDENCE

Assume that the random field $\{Z(x),\ x \in D\}$ is observed at spatial locations $x_i \in D$. Once the range C is determined, a sample $Z(x_1), \ldots, Z(x_n)$ of practically independent variables is obtained by subsampling the original locations so that the distance between all pairs is greater than C, i.e., $\|x_i - x_j\| > C$, $i, j = 1, \ldots, n$.

When we consider a tessellation composed of grains Ξ_i, a subsample of practically independent grains is obtained as follows. First, for each grain Ξ_i denote its centre of mass by X_i, estimate its

diameter, and denote it as r_i. Then randomly order the grains. The first grain is included in the subsample. Each additional grain, say Ξ_j, is included only if it satisfies the condition $\|X_i - X_j\| - (r_i + r_j) > C$ for all preceding grains Ξ_i, otherwise it is rejected. An example of this sequential procedure is shown in Fig. 4, where accepted grains are highlighted. The resulting subsample contains only grains with a mutual distance greater than C.

We have obtained subsamples that could be regarded as spatially independent data, which are used to analyse the non-spatial dependencies. We use a specific test that makes it possible to test the independence between variables with values in higher and not necessarily equal dimensions. In case of rejecting the independence hypothesis, we further focus on the heteroscedasticity of the stress variables.

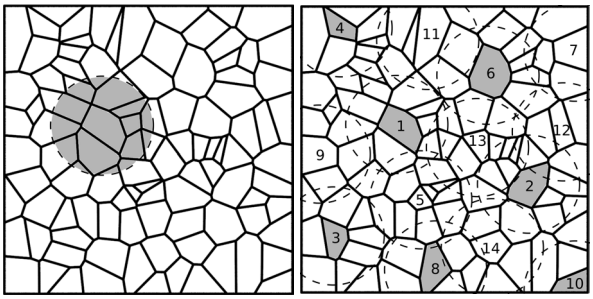


Fig. 4: Example of a tessellation with uncorrelated volume (left) and subsample of spatially uncorrelated grains (right).

The procedure is to be demonstrated on three examples of dependent and, for simplicity, real-valued variables depicted in Fig. 5. The dependence is apparent in all three examples. We see that the dependence is non-linear and even non-monotone. Moreover, in the latter two examples, the data are also heteroscedastic. The variability in Fig. 5 (b) monotonically decreases while there is a non-monotonic relationship in Fig. 5 (c).

The procedure proceeds as follows. First, the independence test is performed. Then, in the case of rejecting the independence hypothesis, heteroscedasticity is studied. The values of both Pearson and Spearman correlation coefficients are close to zero in all three examples presented in Fig. 5. Therefore, we prefer to apply the independence test based on the distance correlation, which will be explained later. The next step is the heteroscedasticity investigation method, which is based on clustering the data according to the explanatory variable. This is followed by computation of the correlation coefficient between the cluster means of the explanatory variable and the standard deviations (stds) of the response variables within a single cluster. Finally, formal statistical tests of uncorrelatedness and independence are performed.

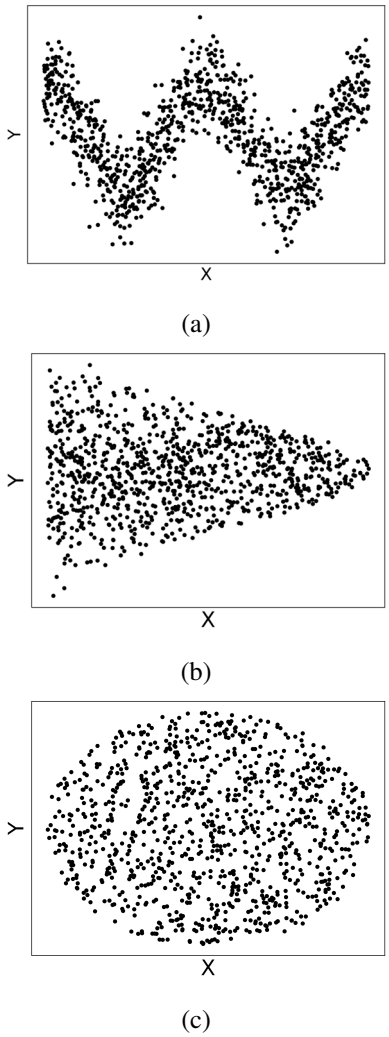


Fig. 5: Simulated examples of dependent variables.

Test of independence

To test the independence between random vectors, we use the so-called distance correlation coefficient introduced in Székely *et al.* (2007); Székely and Rizzo (2009) for a pair (X, Y) of random vectors having finite first moments. The advantage of the distance correlation coefficient is that, unlike both Pearson and Spearman coefficients, the distance correlation between two random variables is zero if and only if these variables are independent. This allows performing statistical tests of independence, not only uncorrelatedness. The test can be performed using the permutation bootstrap method. Moreover, using the distance correlation, we can deal with variables with values in spaces with higher and not necessarily equal dimensions. However, unlike Pearson and Spearman coefficients the test based on the distance correlation coefficient does not reveal the type of the relationship.

The definition of the distance correlation coefficient was extended in (Lyons, 2013) to random elements with values in metric spaces of strong negative type. This property is needed for the equivalence saying that the distance correlation between two random elements is zero if and only if these elements are independent.

We say that a metric space $(\mathbb{X}, d)$ is of negative type if

$$\sum_{i=1}^{n}\sum_{j=1}^{n} c_i c_j d(x_i, x_j) \leq 0$$

for all $n \in \mathbb{N}$, $x_1, \ldots, x_n \in \mathbb{X}$, and $c_1, \ldots, c_n \in \mathbb{R}$ such that $\sum_{i=1}^{n} c_i = 0$. A metric space $(\mathbb{X}, d)$ is of strong negative type if it is of negative type and

$$\int_{\mathbb{X}\times\mathbb{X}} d(x_1, x_2)\, \mathrm{d}(\nu_1 - \nu_2)^2(x_1, x_2) = 0 \quad \Longleftrightarrow \quad \nu_1 = \nu_2$$

for all ν_1, ν_2 probability measures on $\mathbb{X}$ with finite first moment (Lyons, 2013).

For the sake of simplicity, we describe only a sketch of the independence test, for more details see Székely *et al.* (2007) or Lyons (2013). Let $(X_1, Y_1), \ldots, (X_n, Y_n)$ be a random sample of random elements (X, Y) with values in metric spaces of strong negative type $(\mathbb{X}, d_X)$, $(\mathbb{Y}, d_Y)$, respectively. For $j, k = 1, \ldots, n$ compute the distance between the j-th and k-th component of the vector $(X_1, \ldots, X_n)$ using the d_X metric and apply the same to the vector $(Y_1, \ldots, Y_n)$ with the metric d_Y. This gives rise to the matrices $(a^X_{j,k})$ and $(a^Y_{j,k})$, respectively. From these matrices we construct matrices $(A^X_{j,k})$ and $(A^Y_{j,k})$ as follows:

$$A_{j,k} = a_{j,k} - \frac{1}{n}\sum_{i=1}^{n} a_{j,i} - \frac{1}{n}\sum_{i=1}^{n} a_{i,k} + \frac{1}{n^2}\sum_{i=1}^{n}\sum_{l=1}^{n} a_{i,l}.$$

The empirical distance correlation coefficient is then calculated as

$$\mathrm{dCorr}_n(X, Y) = \frac{\mathrm{dCov}_n(X, Y)}{\sqrt{\mathrm{dVar}_n(X)\,\mathrm{dVar}_n(Y)}},$$

if $\mathrm{dVar}_n(X)\,\mathrm{dVar}_n(Y) > 0$, otherwise 0, where

$$\mathrm{dCov}^2_n(X, Y) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} A^X_{j,k} A^Y_{j,k},$$

$$\mathrm{dVar}^2_n(X) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} (A^X_{j,k})^2,$$

$$\mathrm{dVar}^2_n(Y) = \frac{1}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} (A^Y_{j,k})^2.$$

The corresponding p-value of the permutation bootstrap test is calculated as

$$p = \frac{1}{N+1}\left[1 + \sum_{i=1}^{N} \mathbf{1}\left\{\mathrm{dCorr}^{(i)}_n(X, Y) \geq \mathrm{dCorr}_n(X, Y)\right\}\right],$$

where $\mathrm{dCorr}^{(i)}_n(X, Y)$ denotes the empirical coefficient calculated from the random sample $X_1, \ldots, X_n$ and a randomly permuted sample $Y^{(i)}_1, \ldots, Y^{(i)}_n$ obtained from the random sample $Y_1, \ldots, Y_n$.

As shown in Table 2, the null hypothesis of the test concerning the examples from Fig. 5 is rejected in all three cases at the level $\alpha = 0.05$.

Table 2: Results of independence test for examples from Fig. 5.

	Fig. 5 (a)	Fig. 5 (b)	Fig. 5 (c)
dCorr_n	0.220	0.210	0.090
p-value	$\ll \mathbf{0.005}$	$\ll \mathbf{0.005}$	**0.006**

To apply the distance correlation approach to our data, we need appropriate metrics d_X and d_Y. Since $(\mathbb{R}^{3\times3}, d_1)$, with d_1 defined by Eq. (5), is a space of negative type, according to Lyons (2013, Corollary 3.18), $d_1^{1/2}$ can be used to measure the distances between stress tensors. Let $\mathbb{S}^3$ denote the set of unit quaternions and let d_{dis} denote the disorientation angle. It can be shown that $(\mathbb{S}^3, d_{\mathrm{dis}})$ is a metric space of negative type and thus, $d_{\mathrm{dis}}^{1/2}$ can be used to measure the distances between lattice orientations. Similarly, the metric

$$d^{1/2}_{\alpha}(u, v) = \sqrt{\arccos\left(\frac{|\sum_{i=1}^{3} u_i v_i|}{\|u\|\|v\|}\right)},$$

can be used to measure the distances between two boundary orientations represented by lines through the origin, spanned by non-zero vectors u and v, which are the normal vectors of two boundaries.

Heteroscedasticity investigation

Let X (e.g., disorientation angle) and Y (e.g., grain-wise averaged stress tensor) be random elements for which the test introduced above rejects the null hypothesis about the mutual independence. The natural desire is to quantify the manner of dependence. Given the variable X, we focus only on determining the heteroscedasticity of Y, i.e., the case when the conditional variance of Y is non-constant. There are already developed procedures for testing the heteroscedasticity (see, e.g., Breusch and Pagan (1979); Koenker and Bassett (1982)). Nevertheless,

our question of interest is how the standard deviation/variance of the response variable depends on the explanatory variable. For higher-dimensional data, this task is not straightforward. Therefore, we study the dependence for all components separately. In some cases, such reduction to individual components does not make sense, e.g., for lattice orientation.

The proposed approach lies in clustering the data based on the values of the explanatory variable followed by computation of the means and stds within clusters. The final step comprises the calculation of the correlation coefficients from which conclusions about the dependence can be made. Knowledge of the correlation coefficients between these vectors provides information about the shape or at least the manner of dependence in the original data. This is useful when dealing with some large datasets, in which case the depicted graphs can be quite cluttered.

The proposed method is the following. Arrange the random sample $(X_1,Y_1),\dots,(X_n,Y_n)$ into a matrix Z with n rows and 2 columns. Sort the rows of Z by the values of the vector $(X_1,\dots,X_n)$ and denote the resulting matrix as $Z'=(X',Y')$. Divide Z' into k groups/clusters with number of components $n_1,\dots,n_k$ with $\sum_{i=1}^{k} n_i = n$. Next, compute vector $\bar{X}'_{\cdot}$ of group means of X' and vector $Y^{\star}_{\cdot}$ of group stds of Y', i.e.,

$$\bar{X}'_{\cdot} = \left(\frac{1}{n_1}\sum_{i=1}^{n_1} X'_{1,i}, \dots, \frac{1}{n_k}\sum_{i=1}^{n_k} X'_{k,i} \right),$$

$$Y^{\star}_{\cdot} = \left(\sqrt{\frac{1}{n_1}\sum_{i=1}^{n_1} \left(Y'_{1,i} - \bar{Y}'_{1\cdot}\right)^2}, \dots, \sqrt{\frac{1}{n_k}\sum_{i=1}^{n_k} \left(Y'_{k,i} - \bar{Y}'_{k\cdot}\right)^2} \right),$$

where $\bar{Y}'_{\cdot} = \left(\bar{Y}'_{1\cdot},\dots,\bar{Y}'_{k\cdot}\right)$ denotes the vector of group means of Y'. The sketch of the procedure is given in Eq. (6).

$$\begin{pmatrix} X_1 & Y_1 \\ \vdots & \vdots \\ X_n & Y_n \end{pmatrix} \xrightarrow{\text{sort by } X} \begin{pmatrix} X'_1 & Y'_1 \\ \vdots & \vdots \\ X'_n & Y'_n \end{pmatrix} \xrightarrow{\text{clustering}} \begin{pmatrix} X'_{1,1} & Y'_{1,1} \\ \vdots & \vdots \\ X'_{1,n_1} & Y'_{1,n_1} \\ \vdots & \vdots \\ X'_{k,1} & Y'_{k,1} \\ \vdots & \vdots \\ X'_{k,n_k} & Y'_{k,n_k} \end{pmatrix} \xrightarrow[\text{stds of } Y]{\text{means of } X} \begin{pmatrix} \bar{X}'_{1\cdot} & Y^{\star}_{1\cdot} \\ \vdots & \vdots \\ \bar{X}'_{k\cdot} & Y^{\star}_{k\cdot} \end{pmatrix} \tag{6}$$

In the last step, we compute the correlation coefficients between vectors $\bar{X}'_{\cdot}$ and $Y^{\star}_{\cdot}$ and perform the corresponding tests of coefficients being zero. We use both Pearson and Spearman as well as the distance correlation coefficient. The Pearson correlation coefficient measures the strength of a linear relationship while Spearman's rank correlation coefficient the strength of a monotonic relationship. Therefore, we conclude that the relationship is non-monotonic if the null hypothesis is rejected for the distance correlation only and is not rejected for both Pearson and Spearman coefficients. Rejecting the null hypothesis about uncorrelatedness using the Pearson test would suggest that the relationship between the variables is linear. However, due to the large power of this test, a linear dependence may also be suggested in the case of monotonic but non-linear relationships. Thus, for the sake of simplicity, we decide to reduce the number of options and exclude the linear relationship. We also record the sign of the Pearson correlation coefficient to determine whether the relationship is positive or negative. The classification of the dependence is as follows.

– *None* – the test based on the distance correlation does not reject.

– *Non-monotone* – the test based on the distance correlation rejects, but both Pearson and Spearman do not reject.

– *Monotone* $(-)$ – the test based on the distance correlation rejects and also Pearson or Spearman. Additionally, the sign of the Pearson coefficient is -1.

– *Monotone* $(+)$ – the test based on the distance correlation rejects and also Pearson or Spearman. Additionally, the sign of the Pearson coefficient is 1.

An important question is how to cluster the components of Z'. One possibility is to divide Z' into approximately equally large groups. The disadvantage of this choice is that dissimilar values of X' may occur in one group. This may lead to larger stds of X' within a single group. To avoid this, one can divide Z' into groups based on the stds of X'. In practice, we fix an upper bound of the std and add another X'_i to a cluster until the bound is exceeded. However, this approach can give rise to the problem of small cluster sizes n_i. Moreover, the property of generality could be lost.

To compromise between these two approaches, we use the k-means clustering method (see e.g., MacQueen (1967)). This iterative method finds a locally optimal solution by minimising the distance between each component and its nearest cluster mean. It seems that a good choice of the initial solution may

be the division of Z' into approximately equally large clusters.

The remaining question is the choice of k, i.e., the number of clusters. For large values, the problem with low values of n_i may occur again. On the contrary, for small values of k, unrealistic stronger correlations may appear. For example, there are only two observations for the choice $k = 2$. Thus, the correlation would most likely be strong. To conclude, with a fixed number of observations N, the need is to construct as many clusters as possible as well as to construct clusters with sufficiently many components. For this reason, as a rule of thumb we propose the choice $k = \lfloor\sqrt{N}\rfloor$, where $\lfloor\cdot\rfloor$ denotes the floor function. The scheme of the proposed procedure is summarised in Algorithm 1.

Algorithm 1

1. Compute $(\bar{X}'_\cdot, Y^\star_\cdot)$ as in Eq. (6).
2. Compute p-values of the tests based on distance, Pearson and Spearman correlation coefficients.
3. Apply classification.

However, since we use only the spatially independent subsample, the results may depend on the choice of the subsample. Therefore, it is convenient to use multiple subsamples (obtained by choosing different seeds) and combine the results. For this purpose, we use the extreme rank envelope test (Myllymäki *et al.*, 2017). The scheme of the procedure is summarised in Algorithm 2. The idea is to compute the correlation coefficients together with the p-value of the corresponding test of significance between means and stds for each subsample. Then apply the same to the permuted pairs. In more detail, the vector of means stays fixed while we are randomly permuting the vector of stds. This gives rise to a table of size (#subsamples $\times$ (#permutations+1)) of p-values. Finally, the extreme rank envelope test can be applied to this table to determine the extremeness of the original unpermuted pair. The procedure is summarised in Algorithm 2.

Algorithm 2

1. S times repeat:
   - I. Subsample $\big((X_1,Y_1),\ldots,(X_{n_S},Y_{n_S})\big)^T$.
   - II. Compute $(\bar{X}'_\cdot, Y^\star_\cdot)$ as in Eq. (6).
   - III. Compute p-values of the tests based on distance, Pearson and Spearman correlation coefficients.
   - IV. P times randomly permute $Y^\star_\cdot$ and compute p-values of the tests.

   $\rightarrow$ Three vectors of p-values of length $(P+1)$.
2. Apply the extreme rank envelope test to all three matrices (type $S \times (P+1)$) $\rightarrow$ p-values.
3. Apply classification.

Table 3 provides a summary of the results of the simpler method described in Algorithm 1 applied to the examples from Fig. 5 with the response Y and the explanatory variable X. There are p-values corresponding to the uncorrelatedness and independence tests with sign of the Pearson coefficient. The p-values corresponding to the tests for which the null hypothesis is rejected at the level $\alpha = 0.05$ are highlighted. The results correspond to our expectations. The method did not reveal any heteroscedasticity in Fig. 5 (a). The variance in Fig. 5 (b) is monotonically decreasing. The method discovered a non-monotonic relationship between X and the variance of Y in Fig. 5 (c).

Table 3: The p-values of statistical tests for the data presented in Fig. 5.

	Fig. 5 (a)	Fig. 5 (b)	Fig. 5 (c)
Pearson	0.479 (−)	$\ll$ **0.005** (−)	0.907 (−)
Spearman	0.330 (−)	$\ll$ **0.005** (−)	0.990 (−)
Distance	0.519	$\ll$ **0.005**	**0.013**
Dep. type	None	Mon. (−)	Non-mon.

RESULTS

SPATIAL DEPENDENCE

Our dataset is comprised of 7 685 grains with more than 1 500 000 elements. According to the procedure, to characterise the heteroscedasticity in the data, the first step is the determination of the spatial dependencies. For this purpose, we use the generalised semivariogram defined by Eq. (3). We estimate this function from a random field of stress tensors at element-wise locations, which give more refined

information about the range of dependence than only grain-wise averaged stress tensors. However, since estimating the semivariogram over all elements would be computationally demanding, we choose a random sample of elements in the following way. From each grain, we randomly select $1/5$ of the contained elements. This leads to a sample composed of approximately 300 000 elements. These sampled elements are used in the estimator given by Eq. (4).

Fig. 6 shows the estimates of the generalised semivariograms of stress tensors for different choices of elastic constants. The bin width in Eq. (4) is $\delta = r_{i+1} - r_i = 0.5$. We see that the estimated ranges for different choices of the elastic constants are comparable, around 6 µm. Note that the mean grain diameter of the underlying tessellation is approximately 4–5 µm. Thus, we conclude that the stress values mostly affect the stress values of the neighbouring grains only.

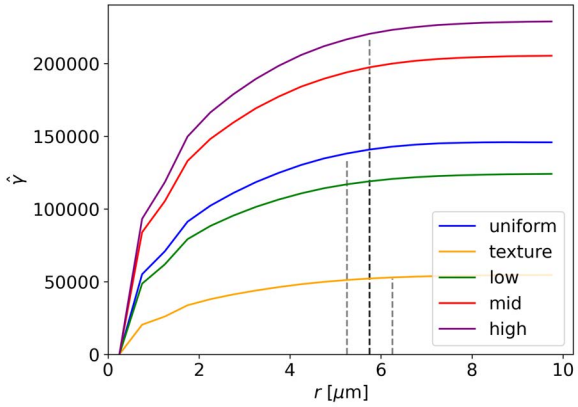


Fig. 6: Estimates of the generalised semivariogram for all five datasets.

NON-SPATIAL DEPENDENCE

Test of independence

In the previous section, it was shown that the stress tensor values are affected by grains at a certain distance. Hence, to apply the independence test mentioned above, which is applicable only to i.i.d. samples, we choose a subsample of practically spatially uncorrelated grains. Such a subsample consists of 430 grains for dataset *uniform*, 335 grains for dataset *texture*, and 382 grains for the remaining datasets. This should guarantee the independence and identical distribution of components within the subsample.

Table 4 provides a summary of the results of the independence test between the grain microstructure variables and $Stress_g$ *[avg]*, $Stress_g$ *[std]*. The cells corresponding to the dependent pairs, i.e., the independence hypothesis was rejected, are highlighted. It is conclusive that the stress field values depend on the lattice orientation for all datasets, which corresponds to the natural dependence due to the directional anisotropy of elastic properties. The microstructure with uniformly distributed crystal orientations showed lower dependence on microstructure parameters compared to the textured one as the test of independence was not rejected only for lattice orientation, volume and maximum disorientation angle, where the former two affect both $Stress_g$ *[avg]* and $Stress_g$ *[std]*, while the last one affects only intra-granular stresses ($Stress_g$ *[std]*). The textured microstructure showed additional dependencies of the stress field values on load orientation, minimum and mean disorientation angle. The variance of disorientation of neighbouring grains affected only average grain stresses. Finally, the elastic anisotropy does not seem to systematically affect the dependencies.

Regarding the grain boundaries, to avoid the internal dependencies within samples, the range of spatial dependence C should be again taken into consideration. This time we choose to take a subsample of the grain boundaries such that none of the pairs has centres closer than C. This subsample is comprised of 1104 grain boundaries for the dataset *uniform*, 791 grain boundaries for the *texture*, and 947 grain boundaries for the remaining datasets.

Table 5 provides a summary of the results of the independence test between the boundary microstructure variables and both $Stress_b$ *[avg]* and $Stress_b$ *[std]* as the results for these tensors are identical. The cells corresponding to the dependent pairs are again highlighted. We see that for all datasets, we reject the independence hypothesis between stress on the grain boundary and its area irrespectively of texture and degree of the elastic anisotropy. In addition, we reject the independence hypothesis between stress on the grain boundary and disorientation for textured microstructure and all anisotropy degrees. The stress seems to be independent of both the boundary orientation and the volume ratio of the neighbouring grains.

To fully understand the relationships between the stress distribution and the microstructure variables, these results may be compared with the analysis through the Pearson correlation coefficient (applied to individual components of the multidimensional variables). However, the values of the Pearson coefficient are almost zero for all examined pairs except for the pairs formed by *Area* and the components of $Stress_b$ *[std]* corresponding

Table 4: Dependencies of stress tensors values on microstructure grain variables. The cells corresponding to the dependent pairs, i.e., the independence hypothesis was rejected, are highlighted.

	Dis [avg]		*Dis [min]*		*Dis [max]*		*Dis [std]*		*Lattice ori*		*Load ori*		*Volume*	
$Stress_g$	*[avg]*	*[std]*	*[avg]*	*[std]*	*[avg]*	*[std]*	*[avg]*	*[std]*	*[avg]*	*[std]*	*[avg]*	*[std]*	*[avg]*	*[std]*
uniform						✓			✓	✓			✓	✓
texture	✓	✓	✓	✓		✓	✓		✓	✓	✓	✓	✓	✓
low	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓
mid	✓	✓	✓	✓		✓	✓		✓	✓	✓	✓	✓	✓
high	✓	✓	✓	✓		✓	✓		✓	✓	✓	✓	✓	✓

to datasets with non-uniformly distributed crystal orientations. In this particular case, the values are significantly high. Thus, we can conclude that for these datasets, the values of $Stress_b$ *[std]* are higher for larger boundaries.

Table 5: Dependencies within grain boundaries (both $Stress_b$ *[avg]* and $Stress_b$ *[std]*). The cells corresponding to the dependent pairs, i.e., the independence hypothesis was rejected, are highlighted.

	Area	*Boundary ori*	*Dis angle*	*Vol ratio*
uniform	✓			
texture	✓		✓	
low	✓		✓	
mid	✓		✓	
high	✓		✓	

Heteroscedasticity

Next, we apply the method from Algorithm 2 with choices $S = 99$ and $P = 199$. However, the variables of interest have values in higher dimensions. Thus, we need to look at their components individually. The results are shown in Fig. 7. Since all datasets with preferred lattice orientation behave very similarly irrespectively of the elastic anisotropy we distinguish, for simplicity, only the datasets *uniform* and *texture*.

It can be noted that, in the case of the microstructure with uniformly distributed lattice orientations, the stress within the whole grain is related to:

– *Volume* – the smaller volume, the larger distributions of inter- and intra-granular stresses, i.e., the larger variance of mean stress (*Stress [avg]*) and standard deviation of stress (*Stress [std]*) within grains,

– *Dis [max]* – the grains with higher maximum disorientation have a smaller variance of intra-granular stresses (*Stress [std]*),

– *Load ori* – no dependence, which is in agreement with usually assumed macroscopic elastic anisotropy of polycrystals with uniformly distributed lattice orientations.

The local stress at the grain boundary depends on the area of the grain boundary (*Area*) – the grain boundaries with smaller area have larger mean stress and stress inhomogeneity at grain boundaries.

The same dependencies and their trends apply to the textured microstructure with the preferred lattice orientations. However, the following additional dependencies of the stress within the whole grain are detected:

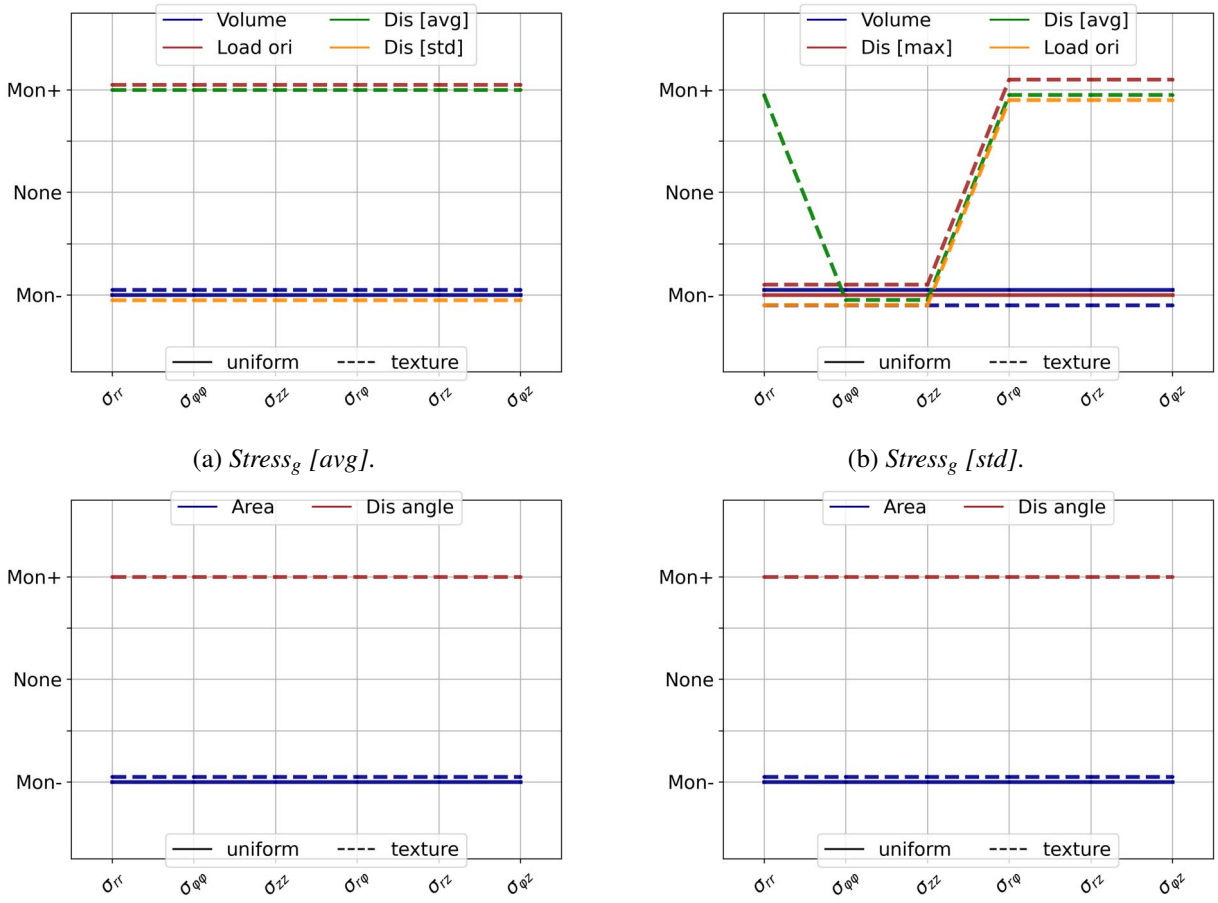
– *Load ori* – the grains having their [111] lattice direction closer to the loading direction show higher variance of mean stresses for all stress components while a smaller variance of intra-granular normal stresses and larger variance of intra-granular shear stresses,

– *Dis [avg]* – the grains with higher mean disorientation have a larger variance of mean stress,

– *Dis [std]* – the grains with higher fluctuation of disorientations have a smaller variance of mean stress,

– *Dis [min]* and *Dis [max]* – not a unique trend was identified.

The local stress at the grain boundary of textured microstructure depends additionally on the disorientation of neighbouring grains (*Dis angle*) – the grain boundaries having larger disorientation show



(a) *Stress$_g$ [avg]*. (b) *Stress$_g$ [std]*.

(c) *Stress$_b$ [avg]*. (d) *Stress$_b$ [std]*.

Fig. 7: Heteroscedasticity type of the stress components within grains (a, b) and on grain boundaries (c, d).

a larger variance of mean stresses and larger variance of stress fluctuations at grain boundaries.

DISCUSSION

The results showed that for the considered microstructure and loading, the degree of elastic anisotropy does not essentially affect the dependencies contrarily to the lattice orientation distribution that was found to affect the dependencies. We were able to detect some dependencies between stress and microstructure. Moreover, we characterised the actual type of trend of variance of stress conditionally on microstructure parameters.

What might still be an object of discussion is the choice of the families as well as the recommended lower bound of the number of observations contained in each cluster. The higher performance of the method would be unequivocally ensured by a higher number of observations. However, collecting such a considerably large dataset may be an issue. Moreover, processing such a dataset may be time-demanding. With a large dataset, we can afford to have enough clusters to accurately test the dependence as well as to be endowed with enough observations inside each cluster to compute the mean and standard deviation with higher precision. In the case of spatially independent data, the method simplifies to no subsampling and no extreme rank envelope test. Thus, for such a dataset, fewer observations are satisfactory.

A disadvantage of the proposed method is the necessity of reducing the multivariate variables down to single components. When dealing with complex variables such as the grain orientation, the reduction becomes complicated. The reduction of the stress tensor is quite natural from the origin of the variable.

The method, however, provides only insight into the dependence of the stress components on single variables. This helps to understand the individual relationships, but in case of an intention to model the whole structure, a deeper grasp is needed. For example, the whole vector of disorientations should be

taken into consideration instead of only the vector of characteristics. An important note is also that, when investigating the dependencies, multicollinearity may be present. Thus, it is important to be careful in explaining the results. To check the multicollinearity, it is possible to use the independence test with distance correlation coefficient or another already established method (see, e.g., Kumar (1975); O'Brien (2007)). In the presented analysis, there was nothing to indicate that the microstructure parameters are dependent.

ACKNOWLEDGMENTS

We thank the reviewers for their helpful comments. We kindly acknowledge financial support from the Grant schemes at Charles University project no. CZ.02.2.69/0.0/0.0/19_073/0016935 (I. Karafiátová) and the Czech Science Foundation by project no. 19-04412S (Z. Pawlas) and project no. 20-14114S (L. Heller). The work is supported by Operational Programme Research, Development and Education financed by European Structural and Investment Funds and the Czech Ministry of Education, Youth and Sports (Project No. SOLID21 CZ.02.1.01/0.0/0.0/16_019/0000760).

REFERENCES

Abdallah B, Willot F, Jeulin D (2015). Stokes flow through a Boolean model of spheres: Representative volume element. Transp Porous Media 109(3):711–26.

Adams BL, Kalidindi SR, Fullwood DT (2013). Microstructure-Sensitive Design for Performance Optimization. Elsevier Inc.

Angulo J (2014). Structure tensor image filtering using Riemannian L^1 and L^∞ center-of-mass. Image Anal Stereol 33(2):95–105.

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

Breusch TS, Pagan AR (1979). A simple test for heteroscedasticity and random coefficient variation. Econometrica 47(5):1287–94.

Brough DB, Wheeler D, Kalidindi SR (2017). Materials knowledge systems in Python - A data science framework for accelerated development of hierarchical materials. Integr. Mater Manuf Innov 6(1):36–53.

Bunge HJ (1969). Texture Analysis in Materials Science. Butterworth-Heinemann.

Fullwood DT, Niezgoda SR, Adams BL, Kalidindi SR (2010). Microstructure sensitive design for performance optimization. Prog Mater Sci 55(6):477–562.

Heller L, Karafiátová I, Petrich L, Pawlas Z, Shayanfard P, Beneš V, Schmidt V, Šittner P (2020). Numerical microstructure model of NiTi wire reconstructed from 3D-XRD data. Model. Simul. Mater. Sc. 28(5):055007.

Hill R (1963). Elastic properties of reinforced solids: Some theoretical principles. J Mech Phys Solids 11(5):357–72.

Hirsekorn S (1990). Elastic properties of polycrystals: a review. Texture Microstruct. 12:1–14.

Koenker R, Bassett GJ (1982). Robust tests for heteroscedasticity based on regression quantiles. Econometrica 46:43–61.

Kumar TK (1975). Multicollinearity in regression analysis. Rev Econ Stat 57(3):365–6.

Lyons R (2013). Distance covariance in metric spaces. Ann Probab 41(5):3284–305.

MacQueen JB (1967). Some methods for classification and analysis of multivariate observations. Berkeley Symposium on Mathematical Statistics and Probability 5(1):281–97.

Myllymäki M, Mrkvička T, Grabarnik P, Seijo H, Hahn, U (2017). Global envelope tests for spatial processes. J R Stat Soc B 79:318–404.

O'Brien RM (2007). A caution regarding rules of thumb for variance inflation factors. Quality and Quantity 41:673–90.

Otsuka K, Wayman CM (1998). Shape Memory Materials. Cambridge University Press.

Pawlas Z, Karafiátová I, Heller L (2020). Random tessellations marked with crystallographic orientations. Spatial Statistics 39:100469.

Petrich L, Staněk J, Wang M, Westhoff D, Heller L, Šittner P, Krill CE, Beneš V, Schmidt V (2019). Reconstruction of grains in polycrystalline materials from incomplete data using Laguerre tessellations. Microsc Microanal 25(3):743–52.

Sanei, SHR and Fertig, RS (2015). Uncorrelated volume element for stochastic modeling of microstructures based on local fiber volume fraction variation. Compos Sci Technol 117:191–8.

Sedmák P, Pilch J, Heller L, Kopeček J, Wright J, Sedlák P, Frost M, Šittner P (2016). Grain-resolved analysis of localized deformation in nickel-titanium wire under tensile load. Science (New York, N.Y.) 353:559–62.

Schabenberger O, Gotway CA (2001). Statistical Methods for Spatial Data Analysis. Chapman and Hall/CRC.

Székely GJ, Rizzo ML, Bakirov NK (2007). Measuring and testing dependence by correlation of distances. Ann Stat 35(6):2769–94.

Székely GJ, Rizzo ML (2009). Brownian distance covariance. Ann Appl Stat 3(4):1236–65.

Willot F, Brenner R, Trumel H (2020). Elastostatic field distributions in polycrystals and cracked media. Philos Mag 100(6):661–87.

Yamaji A, Sato K (2006). Distances for the solutions of stress tensor inversion in relation to misfit angles that accompany the solutions. Geophys J Int 167:933–42.

Zener CM (1948). Elasticity and Anelasticity of Metals. University of Chicago Press.

Zienkiewicz O, Taylor R, Zhu JZ (2013). The Finite Element Method: Its Basis and Fundamentals: Seventh Edition. Elsevier Ltd.