Modelling the primary damage in Fe and W: influence of the short-range interactions on the cascade properties: Part 2 – multivariate multiple linear regression analysis of displacement cascades

Andrée De Backer[1,2,3], Charlotte S. Becquart[1,2], Pär Olsson[4], Christophe Domain[5,2]

[1] Univ. Lille, CNRS, INRAE, Centrale Lille, UMR 8207 - UMET - Unité Matériaux et Transformations, F-59000 Lille, France

[2] EM2VM, Joint laboratory Study and Modeling of the Microstructure for Ageing of Materials

[3] CCFE, Culham Centre for Fusion Energy, Abingdon, Oxon, United Kingdom

[4] KTH Royal Institute of Technology, Nuclear Engineering, Roslagstullsbacken 21, SE-10691 Stockholm, Sweden

[5] EDF-R&D, Département Matériaux et Mécanique des Composants, Les Renardières, F-77250 Moret sur Loing, France

Corresponding author: Charlotte S. Becquart charlotte.becquart@univ-lille.fr

Abstract

Neutrons and high energy ions create displacement cascades in materials, which have been simulated using Molecular Dynamics, for many decades now. The breakthrough of this work is to explore in a large statistics of more than 7000 cascades the relation between early cascade morphology and the final primary damage using a multivariate multiple regression analysis.

For two energies in Fe and W, the total number of defects, the number of SIA and vacancy clusters and their full size distributions have been characterized using a multivariate multiple linear regression analysis based on 7 descriptors of the primary damage and 3 morphology descriptors. We find that the combination of the volume and the sphericity is significant. This analysis highlights several cascade properties, among them, that the large and spherical cascades create less defects and in particular, less mono defects than small and fragmented ones.

13 interatomic potentials differing either by their equilibrium part or the way they were hardened have been included in this study and the multivariate analysis shows that the choice of potential has a limited influence on the total number of defects but a large one on the number of mono vacancies. On average, soft potentials create cascades of larger volume, smaller sphericity and producing more defects than hard potentials.

Finally, the formation of vacancy clusters is different in Fe than in W. In Fe, the fraction of vacancies in clusters is larger than that of SIAs and larger vacancy clusters are created than SIA clusters. In W, it is the opposite. The reasons are the differences of stopping power and threshold displacement energies, which result in different spatial distributions of open volumes that form during the expansion stage of the cascade.

Key words

Radiation damage, Primary damage, Displacement cascades, Primary knock on atom, Metallic alloys, Empirical potentials, Multivariate multiple regression analysis.

Highlights

- Multivariate multiple regression analysis on a large cascade database

- The cascade early morphology has a significant impact on the primary damage

- The choice of potential has a larger effect on the number of mono vacancies than on the total number of defects

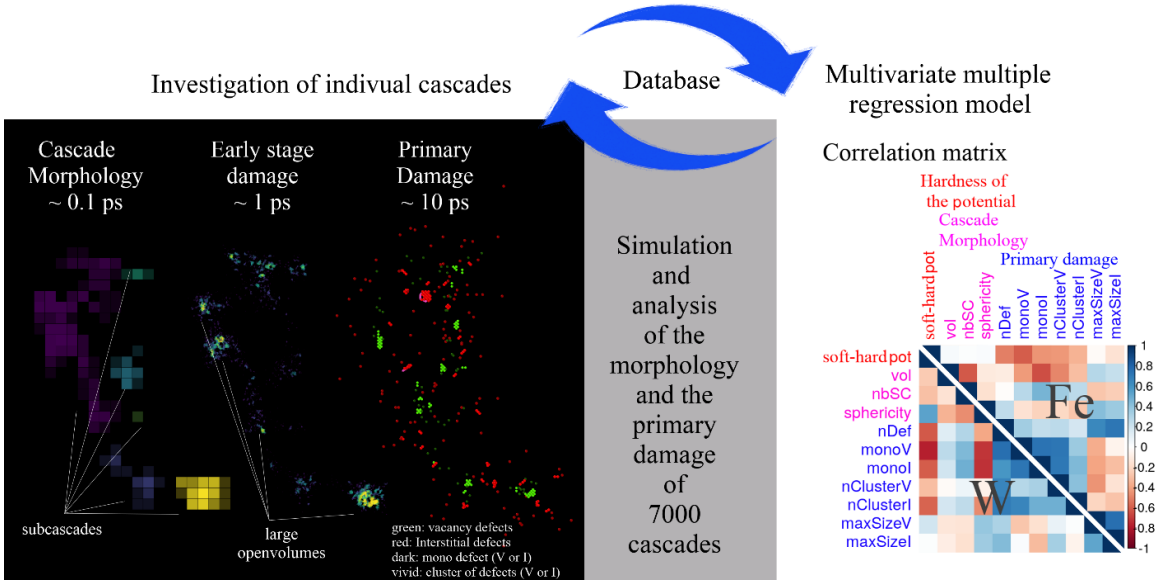
- The number of defects predicted by the potentials for each element agrees with one another within a 50% margin.

- All potentials predict similar SIA cluster sizes and vacancy cluster sizes that differ by a factor two at the most.

- Potentials for Fe predict a larger fraction of vacancies in clusters, whereas potentials for W predict the opposite

- Soft potentials produce cascades which are less spherical, with more defects as well as more clusters.

Graphical abstract



Introduction

The study of radiation damage, initiated by displacement cascade events has motivated and still motivates significant efforts to identify the mechanisms of damage generation and to predict the evolution of mechanical properties of materials under irradiation [1]. In the companion article to this one [2], the elementary energy transfer properties for primary damage generation in iron and tungsten has been studied using atomic scale modelling techniques, such as the binary collision approximation (BCA) and molecular dynamics (MD), as well as density functional theory (DFT). These atomistic techniques enable detailed investigation of fundamental interactions, as well as dynamics and complex many-body behavior.

For the dynamic studies, there are several possible levels of sophistication, realistically starting from dynamic DFT, to semi-empirical MD, to BCA. Some fundamental properties can be addressed at the higher level of sophistication and cost but in order to perform larger dynamic studies, such as that of energetic displacement cascades, it is necessary to use less computationally costly methods, but to first validate and understand their range of applicability. Empirical interaction potentials (EIP) here provide an optimal middle ground in terms of cost, reliability and physical accuracy. These are usually fitted to equilibrium and near-equilibrium properties of the intended material of interest.

During displacement cascades the local compression conditions are often such that the atoms approach one another until they are much closer than what can be considered near-equilibrium. It is therefore necessary to extend, or harden, the potentials to include appropriate physics for very small distances. For the very short range, the Ziegler-Biersack-Littmark [3] (ZBL) model for the screened Coulomb interaction is well motivated and highly used, and it is spliced to the fitted equilibrium potential using some interpolation scheme. The choice of interpolation scheme is, however, usually much less well-motivated with often arbitrarily chosen knot points. Historically, a central idea for how to perform the hardening has been to make sure that the EIP would predict threshold displacement energies (TDE) in agreement with available experimental values. In the companion article to this one [2], there is an overview of the issues related to the choice of hardening methodology.
Nowadays, the number of EIP available for Fe and W, including potentials with the same equilibrium part which have been hardened differently [2] allows important progress in the comprehension of the primary damage formation, compared to the first studies [4] of this kind. Furthermore, thanks to the increase of computing resources larger statistics are possible, allowing the production of meaningful distributions. The statistical analysis done on BCA cascades almost 20 years ago with 5000 to 15000 cascades [5] is nowadays accessible, thanks to high performance computing resources, with MD and thus with a much better description of the recombination phase of the cascade.

In this work, we compare a large number of EIPs for damage production in tungsten and iron and continue the companion article [2] by expanding the energy range and complexity of phenomena studied. Based on a large statistical data base, we evaluate the characteristics of displacement cascades and the relation between the cascade morphology and the defect production. For this purpose, two categories of descriptors have been chosen:
- three morphological descriptors for the cascades (the volume, the number of subcascades and the sphericity)
- seven primary damage descriptors (the total number of defects, the number of mono vacancies and SIAs, the number of vacancy and SIA clusters and the maximum size of the SIA and vacancy clusters).
Using these descriptors, the fraction of vacancies / SIAs in clusters and the mean vacancy / SIA cluster size can be easily obtained.

The first section presents briefly the different numerical methods used. In the second section, the means and the variance of the descriptors, i.e. the characteristics of the cascades are presented. In the third section, we describe the main outcome of the multivariate multiple regression analysis.

I. Methods

1) Modelling the displacement cascades

Displacement cascades were simulated using the MD code DYMOKA [6] at 100 K for the cascade energies 80 keV for Fe and 50 keV for W using all the potentials in the micro canonical ensemble with periodic boundary conditions. The time-step was adjusted during the simulations and no electronic losses were considered, as is the case in most MD studies. Studies of the impact of taking into account the electron-phonon coupling have shown that strong electron-phonon coupling leads to the reduction of the number of defects as well as of the size of point defect clusters [7]. The purpose of this work is to compare potentials and show how properties correlate and thus neglecting electronic losses and electron phonon coupling is not an issue here. It means however that cascade energy is not the same as what is traditionally called the PKA energy. Indeed, in real materials, the PKA energy is partitioned into energy loss due to electronic stopping and elastic processes. The latter is called the damage energy and is the fraction of the PKA energy that produces defects [8]. In our MD simulations, the cascade energy is thus more like the damage energy than the PKA energy and we thus use the term damage energy to characterize the kinetic energy that is given to the PKA in our simulations.

The choice of the simulation box size depends upon the energy of the PKA and was chosen to be large enough (respectively 120 × 120 × 120 and 96 × 96 × 96 bcc unit cells for Fe and W) to avoid the displacement cascade to interfere with itself by periodic overlap. Before initiating the displacement cascade, the system of particles was equilibrated for 3 ps at the chosen temperature. Variability was introduced by changing the initial distribution of the velocities (which follows a Maxwell distribution) as well as the PKA direction. The PKA directions were chosen to be representative of an average behavior. Stoller examined the influence of the PKA direction and showed that the ⟨135⟩ cascades should provide a reasonable representation of average behavior [8]. We have observed similar trends for ⟨253⟩ directions. The OVITO software is used for the 3D visualization [9].

6 (in Fe) and 7 (in W) empirical potentials have been used. For W, we have used four modified versions of the potential originally derived by Finnis and Sinclair [10]: one version modified by Mason et al. [11], that will be referred to as MN, one version modified by Juslin et al. [12], referred to as JW, one version modified by Derlet et al. [13] and hardened by Björkas et al. [14], referred to as DD and one derived by Ackland and Thetford [15] and hardened by Zhong et al. [16] referred to as AT. A new potential derived by C. Marinica et al. [17] and hardened in two different ways (softer and harder) by A. Sand [18] will be referred to as MS-s and MS-h. The potential derived by Mason et al. has also been hardened in two different ways also and the two versions will be referred to as MN-s and MN-h.

For Fe, the potentials investigated will be referred to as MA10A for the MO7 potential initially published in [19], MEND10 for the potential derived by Ackland et al. [20] and CO21013, CO21620 and CO30912 for potentials newly derived by Marinica et al. [21] and hardened in the companion paper [2].

2) Point defect and cluster characterization

The point defect population and their clusters were estimated using a classical lattice site analysis. A cluster is defined by all entities, vacancy and self-interstitial atoms (SIA), within a critical distance. We chose for both elements a point defect cluster (PDC) capture distance, i.e. the vacancy (or interstitial) cluster capture distance for point defects between 2nn (nearest neighbor) and 3nn; more precisely, 3 Å for Fe and 3.5 Å for W. The net defect sum determines the cluster size which is used for the statistics of defect clusters. Note that the amount of point defects and the proportion of defects in clusters depend on the choice of this critical distance, or PDC capture distance. This issue will be discussed in the result section as well as in the supplementary materials.

3) Cascade characteristics and subcascades

The overall cascade morphological characteristics, i.e. their shapes and the number of subcascades produced, have been analyzed following the method proposed in [22]. At different moments of the simulation, the position and the kinetic energy of the atoms whose kinetic energy is above 0.1 eV (this guarantees that they are not at equilibrium and have thus been involved in collisions) is recorded. The simulation box is divided in cubic cells and the average kinetic energy is calculated in each cell. Cells where the average kinetic energy is larger than a specific criterion are considered to be part of the cascade.

Subcascades are determined as regions separated from each other by cells where the average kinetic energy is smaller than the criterion. In [22], we showed that it is important to pick a value larger than 0 and smaller than the maximum energy that can be lost by the PKA in a small volume of material. A systematic study showed that the results do not vary much given that the energy criteria is close to the energy required to melt this volume and we chose 26 eV/nm^3. For the analysis in this work, the subcascade decomposition was taken at the maximum expansion of the cascade, the edges of the cubic cells were 1.5 nm for W and 1.0 nm for Fe. We chose to compare 50 keV cascades in W and 80 keV cascades in Fe in order to compare cascades of similar volumes, in two different damage energy regimes. The volumes were chosen to be large enough to contain clusters that are big enough for statistical analysis. 50 keV cascades in W correspond to the single cascade regime as the damage energy is lower than the average subcascade energy formation threshold (W: 75 keV [22]); whereas in Fe, 80 keV is well above the subcascade threshold (15 keV [22]) and subcascades are formed.

The cascade volumes and surfaces have been calculated using our decomposition method. The cascade volume V is the sum of all the cubic cells and its surface $\boldsymbol{\gamma}$ is the number of external facets of the cubes forming the subcascades. Note that, with this method, the volumes and surfaces obtained depend on the cube size chosen but further studies of this is beyond the scope of this paper and our purpose here is to compare cascades obtained with the different potentials. Instead of considering the surface, we calculated the cascade sphericity, defined as: $S = \pi^{\frac{1}{3}} \frac{(6V)^{\frac{2}{3}}}{\gamma}$.

II. Results

In what follows, we will characterize the primary damage using 7 descriptors for each cascade:
- nDef, the total number of defects
- monoV, the number of mono vacancies
- monoI, the number of mono SIA
- nClusterV, the number of vacancy clusters
- nClusterI, the number of SIA clusters
- maxSizeV, the maximum size of the vacancy clusters
- maxSizeI, the maximum size of the SIA clusters.

Using these descriptors, other characteristics of the primary damage can be obtained: for example, the fraction of vacancies in clusters is (nDef –monoV)/(nDef) and the mean vacancy cluster size is (nDef - monoV)/(nClustersV). Note that the variance of these combinations is unfortunately not that simple to obtain because of the covariances.

The cascade morphology is described by 3 descriptors:
- vol, the cascade volume
- nbSC, the number of subcascade
- sphericity, the cascade sphericity.

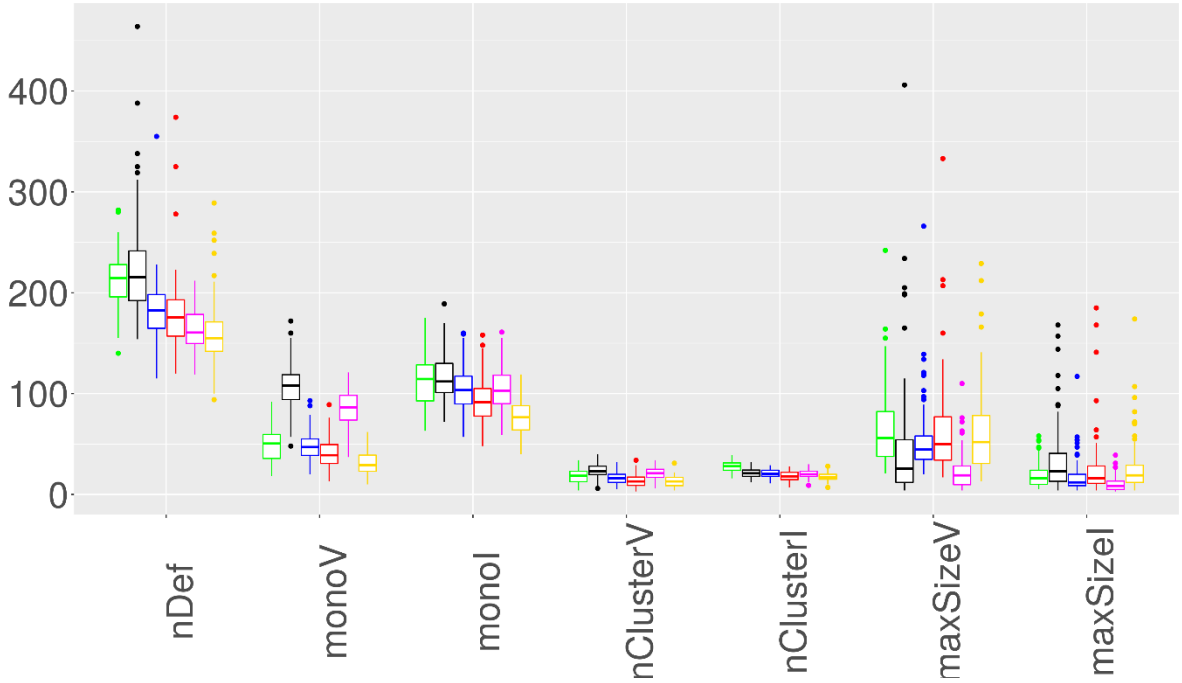
Thanks to the large number of cascades (see **Table S1** in the supplementary material) we obtained accurate means of the descriptors and, more important, we explored their distributions (see **Figure S1** in the supplementary). Interestingly, some distributions are neither symmetrical nor normal distributions, in particular for the maximum cluster sizes. For the sake of brevity, we will not enter into details here. To take into account the shape of the distributions on the plots, we use first and third quartiles and stress that the variance of the descriptors is the key point of the multivariate multiple regression analysis.

An important remark is the impact of the PDC capture distance chosen: here 0.3 nm for Fe and 0.35 nm for W. This point is discussed in detail in the supplementary material which presents the evolution of the mean of the primary damage descriptors as a function of the PDC capture distance varying from 0.3 nm to 1 nm, as well as a discussion of how it impacts differently each descriptor and each element (Fe or W). The sensitivity of the primary damage descriptors to the PDC capture distance is related to the spatial distribution of the defects.

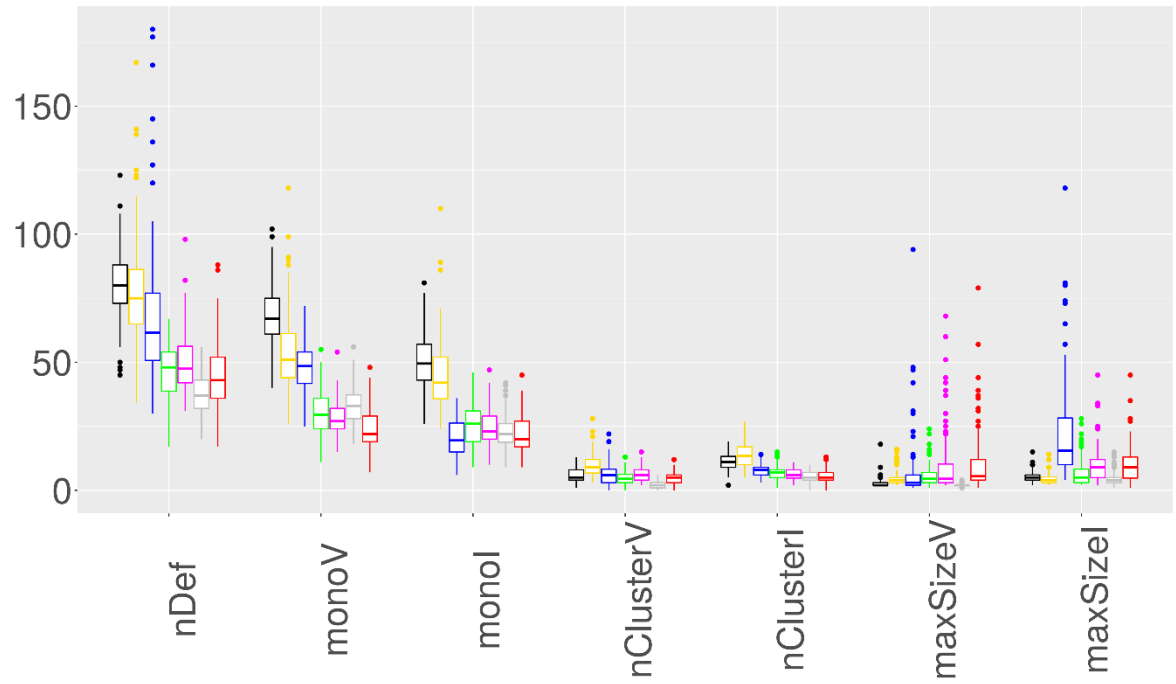
1) Primary damage

Figure 1 sketches the distributions of the primary damage descriptors for displacement cascades with 80 keV damage energy in Fe (a) and 50 keV damage energy in W (b). Box-and-whisker plots, a typical method of descriptive statistics, have been used. The median, the first and third quartiles, the left and right tails and the individual outlying cascades are represented. When the median is not in the middle, the asymmetry of the distributions is revealed, particularly for the cluster maximum sizes. The interquartile range is an indication of the standard deviation. Considering that the standard error of the mean which scales reciprocally to the square root of

the number of cascades (larger than 100), is by one order of magnitude smaller than the standard deviation, one can be convinced that the mean is accurately captured by our dataset. One also observes that the data sets have a relatively large variance.



(a) Fe



(b) W

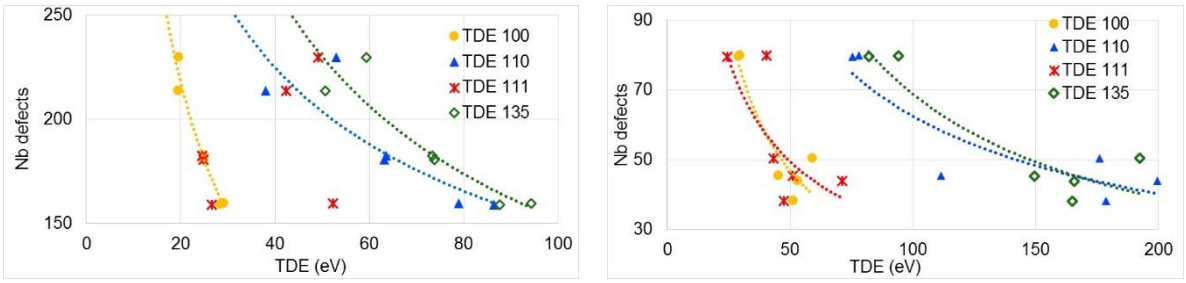
Figure 1: Primary damage created in a (a) 80 keV cascade in Fe and (b) a 50 keV cascade in W: total number of point defects, total number of mono defects, number of defect clusters and defect clusters maximum sizes. Distributions are compactly displayed by box-and-whisker plots with five summary statistics: the median, the first and third quartiles (hinges), ~1.5 interquartile range

from the median (whiskers which include95 % of the points) and "outlying" points shown individually.

As could be expected, we find that the potentials with the lowest TDEs (CO21620, MEND10 for Fe and MN-s and MS-s for W, see the companion paper [2]), i.e. the softest potentials predict more residual defects in agreement with [18] [23]. Overall, though, the number of defects predicted by each potential agrees with one another within a 50% margin.
Our data also show that, in Fe, the fraction of vacancies in clusters is higher than the fraction of SIA in clusters (i.e., there are more mono SIAs than mono vacancies). More than 70% of vacancies are in clusters except for MEND10 and MA10A. In W, it is the opposite: there are more mono vacancies than mono SIAs.

Figure 1 also underlines that displacement cascades in Fe and W create more SIA clusters than vacancy clusters, except for MEND10 and MA10A and that the soft potentials for W predict more clusters than the other potentials.

A clear correlation between the average number of point defects (for a PKA of given energy) and the mean TDE values along the ⟨100⟩, ⟨110⟩ and ⟨135⟩ directions can be seen in **Figure 2** where the lines are regression fitted power laws (guide for the eyes) for the mean number of defects created versus the mean TDE value. The lower the TDE, the higher the number of defects. The correlation is even much better when the data corresponding to the DD potential for W is removed. Note that there are no reasons to believe that the behavior of this potential, which exhibits several energy barriers in the RCS along ⟨111⟩ (**Figure S2** of the supplementary material of paper I [2]), is realistic.



(a) Fe (b) W

Figure 2: Mean number of defects created by a 80 keV cascade in Fe (a) and 50 keV cascade in W (b) versus mean TDE energy. The lines are regression fitted power laws. The data obtained with DD potential has been omitted.

Along the ⟨111⟩ direction, no such correlation is found, most probably because of the fact that this direction being close packed, the energy transmission is mostly 1D and thus does not represent well what happens in cascades.

The average point defect cluster size distributions are represented in **Figure 3**. The distributions show the typical power law which vanishes at the maximum cluster size [24] with a bump which

is a due to the distribution of the largest defect in each cascade [25]. These data clearly show that soft W potentials predict the smallest clusters.

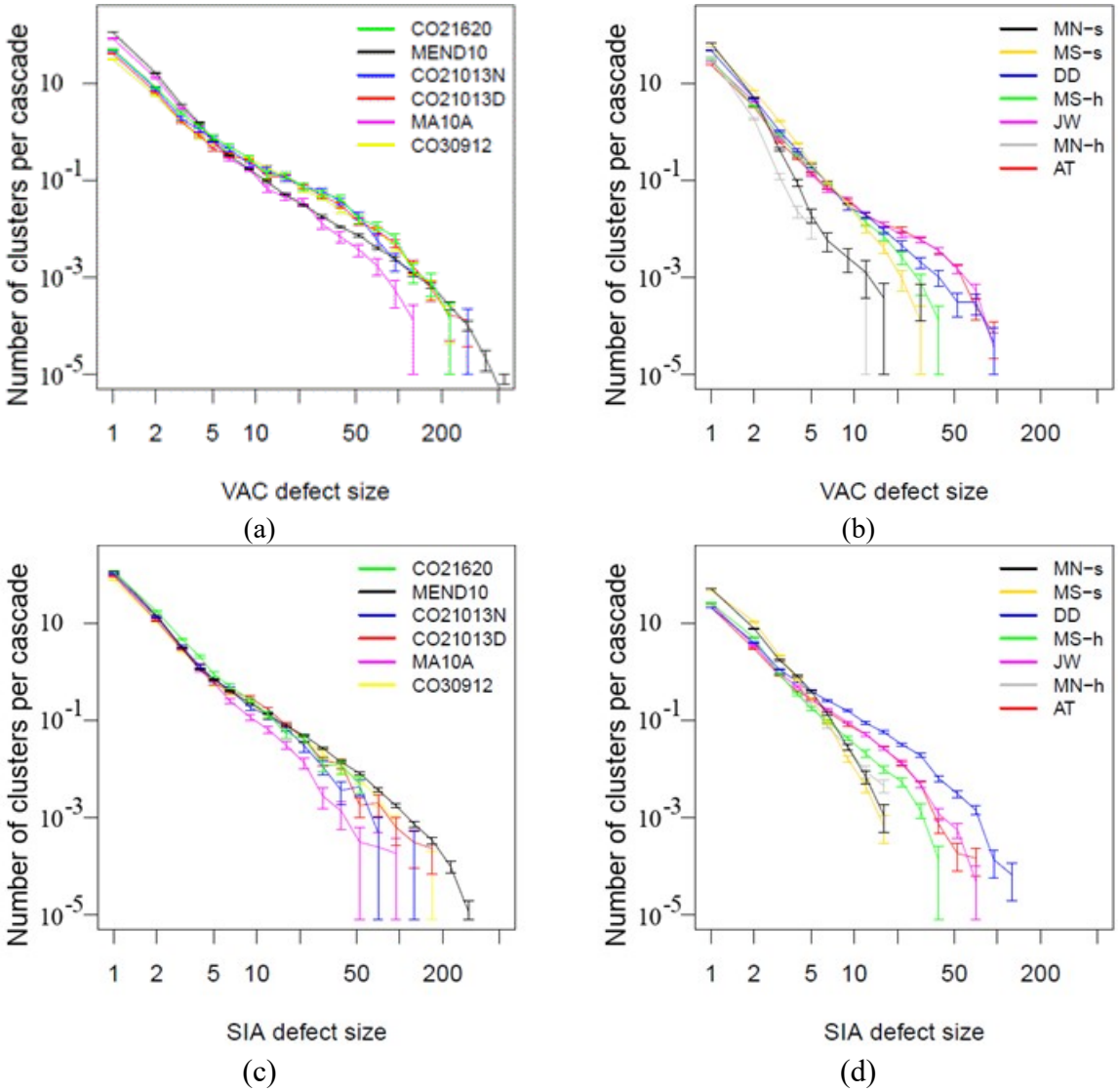
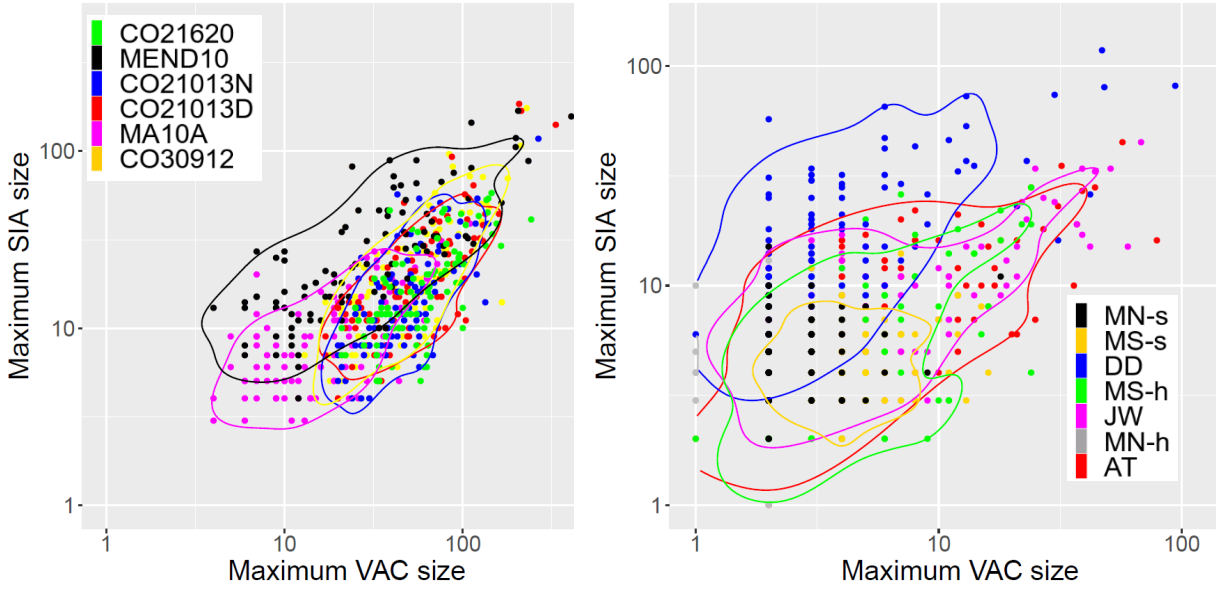


Figure 3: point defect cluster size distributions in 80 keV displacement cascades created in Fe (a) and (c) and 50 keV displacement cascades created in W (b) and (d). (a) and (b): vacancy cluster, (c) and (d): SIA cluster. The bin width increases exponentially and each data point corresponds to the sum of all the clusters having a size within the bin, divided by the bin width and the number of cascades. The vertical bars correspond to the standard error.

The primary damage descriptors also describe the cluster size distribution: the number of mono defects is the value for the smallest size equal to 1 (low limit), the number of clusters is the area below the distribution, the total number of defects is the integral of the distribution times the cluster size and finally, the maximum size of the clusters is related to the distribution tail (large limit). The value for which the distribution drops (see **Figure 3**) is not only related to the mean value but also the variance of the maximum size, which is particularly large (see **Figure 1**).

In agreement with the box-and-whisker plots of maxSizeI and maxSizeV, we can see that in Fe, the size for which the size distribution vanishes is larger for vacancy clusters than for SIA clusters. In W, however, it depends on the potentials.

There are interesting correlations between the VAC and SIA clusters, between monoV and monoI but also between the maxSizeV and maxSizeI as can be seen in **Figure 4**. This indicates that cascades which create large SIA clusters, also create large vacancy clusters and vice versa, as previously pointed out by Calder et al.[26].



(a) 80 keV cascades in Fe (b) 50 keV cascades in W

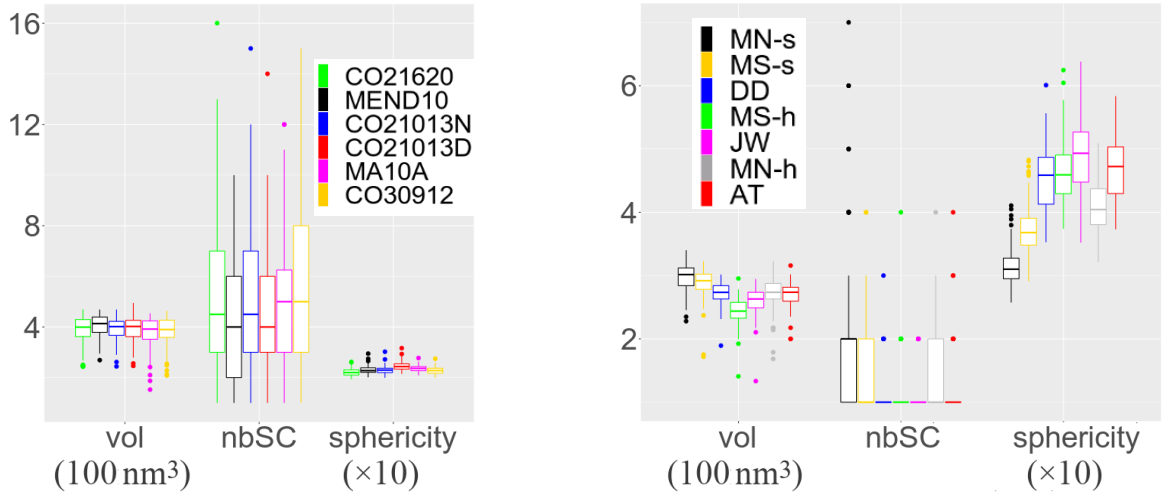
Figure 4: Maximum SIA cluster size vs maximum vacancy cluster size for all the potentials. Each point corresponds to one cascade, the solid curves are the 95% isocontour of the kernel density estimation, which estimates the region containing 95% of the cascade and the highest density of points. The points which lie outside the solid lines are cascades with extreme values.

2) Cascade morphology: cascade volume, sphericity and number of subcascades

Figure 5 represents the means and the first and third quartiles of the volumes, number of subcascades and sphericities, at the maximum cascade expansion. As expected from our choice of damage energy for this study, we find that the great majority of 50 keV cascades in W form one compact core, whereas for 80 keV Fe cascades, the number of subcascades is much larger. In W the energy chosen in this work is below but close to the fragmentation energy. Here, most of the cascades consist of one main subcascade decorated by one small subcascade. See [22] for a description of the subcascade volume distributions and the spatial correlation distance between subcascades.

On average for all the potentials studied, the number of subcascades is 1.48 for 50 keV W and 7.34 for 80 keV in Fe. Considering that there is a shift due to the absence of electronic losses in this work, these numbers are in agreement with our work done on BCA cascades which gave the fragmentation energy equal to 15 keV, resp. 75 keV for Fe, resp. W [22]. The fragmentation is

also captured by the sphericity: a volume v distributed in n identical spheres decreases as $n^{-1/3}$. In Fe, the cascades are more fragmented and the sphericity is lower than in W.



(a) Fe (b) W

Figure 5: Morphology descriptors for 80 keV cascades in Fe (a) and 50 keV cascades in W (b) at the maximum expansion of the cascades. Distributions are compactly displayed by box-and-whisker plots, with five summary statistics: the median, two hinges (corresponding to first and third quartiles), two whiskers (not further than 1.5 interquartile range from the median and which include 95 % of the points) and "outlying" points shown individually.

In W, the soft potentials, MS-s and MN-s, produce cascades with the largest average volume but their sphericity is the smallest because they are fragmented (indicated by nbSC larger than one). Our data also show that the fragmentation energy depends on the hardness of the potential. An explanation for this behavior may be found in the RCS analysis along 〈111〉. Soft potentials tend to transfer kinetic energy perpendicular to the RCS direction, at lower energies than the other potentials, i.e. the energy transfer regime change takes place at lower energies (**Table 3** and **Figure 6** of the companion paper [2]), this indicates that given a distribution of energies in the cascade, for the soft W potentials, the threshold to channel away energy at long distances through RCS or focuson is much lower, so there should be more subcascades. For the Fe potentials, on the other hand, no real difference regarding the subcascade sizes can be observed, however, the softest potentials in Fe are not as soft as the softest potentials in W.

3) Correlation between cascade properties and morphologies

Having such a database at hand, it was tempting to look for correlations between cascade properties. In **Figure 6** are correlation plots of the number of vacancy clusters (nClusterV) as well as the volume (vol), with the number of mono vacancies (monoV), in Fe; and of the sphericity as well as the number of defects (nDef), with the number of mono vacancies (monoV) in W. These plots show that, in Fe, small cascades containing the largest amount of mono vacancies also have the largest number of vacancy clusters. In W, this correlation is observed for cascades of high sphericity rather than large volume, and the trends are more potential dependent (i.e. the clouds corresponding to each potential do not have the same slope).

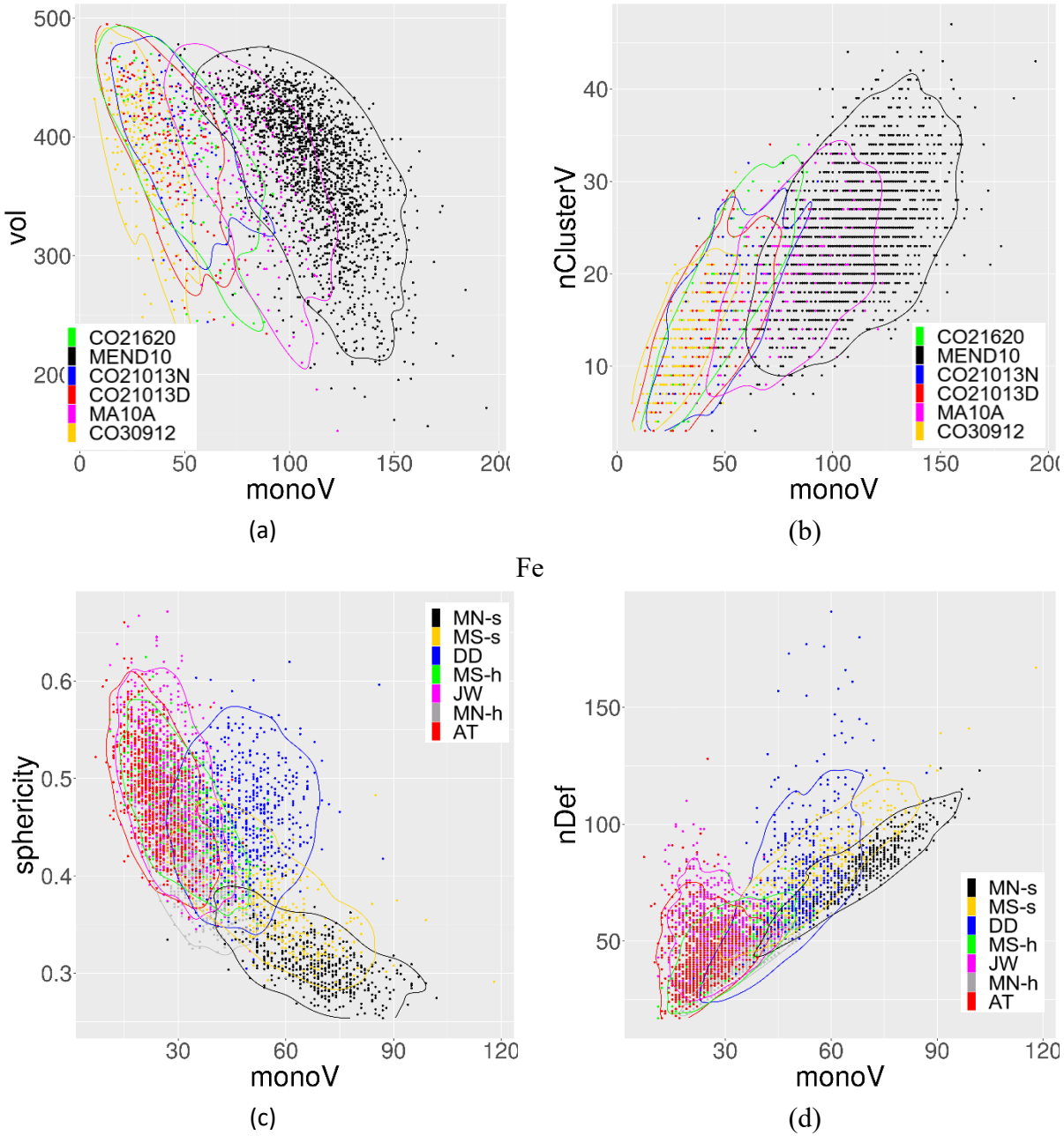


Figure 6: correlation between descriptors: (a) volume versus number of mono vacancies in Fe (b) number of vacancy clusters versus number of mono vacancies in Fe. (c) sphericity versus number of mono vacancies in W (d) number of defects versus the number of mono vacancies in W. The solid curve is the 95% isocontour of the kernel density estimation which estimates the region containing 95% of the cascade and the highest density of points. The points which lie outside the solid lines are cascades with extreme values.

4) Models and MANOVA

After having looked at correlations, for all potentials, between our 10 descriptors (and others not presented here), we tested linear and non-linear models to explain the primary damage properties with an increasing number of explanatory variables. We analyzed the p-values which indicate

the confidence one can have in the model (the smaller the value, the larger the confidence) and R-squared, R^2, that evaluates (in %) how much of the variance of the response variable is explained by the explanatory variable. The results have been analyzed in details, potential by potential, as well as averaged on all potentials and variables. **Table S4** in the supplementary provides examples of the models we tested, and the fraction of experiments (one experiment being one potential and one variable) whose p-value is below 0.01, n0 and R^2. Interestingly, in the supplementary are the radar plots of R^2 which indicate that a linear model including volume and sphericity is the best option.

The main outcomes of our investigations are:

- linear models are the best compromise between simplicity and prediction capacity,

- in general, the morphological descriptor sphericity performs better than the number of subcascades,

- the combination volume + sphericity, is better than each taken individually.

- monoV and monoI have the largest R^2, hence a large sensitivity to the cascade morphology at the maximum expansion

In order to rationalize our findings, we present, in what follows, a multivariate multiple linear regression analysis. This multivariate analysis of variance (MANOVA) considers:

- the potential: one categorical variable **pot** with 6 levels for Fe and 7 levels for W

- seven continuous variables of the primary damage, the total number of defects, **nDef**, the number of mono vacancies, **monoV**, the number of mono SIAs, **monoI**, the number of vacancy clusters, **nClusterV**, the number of SIA clusters, **nClusterI**, the vacancy cluster maximum size, **maxSizeV**, the SIA cluster maximum size, **maxSizeI**.

- two continuous variables describing the morphology, the volume, **vol** and the sphericity, **sphericity**.

- one discrete variable describing the morphology, the number of subcascades, **nbSC**.

We first consider the MANOVA based on **pot** alone. The global p-value is very low which proves the effect of the potential on the primary damage. Indeed, the p-value gives the probability of the null hypothesis, i.e. the probability that the difference between the means of two sets of data, has been found by chance. The probability is large if 1/ the means are similar, 2/ the variance is large and 3/ the number of data is too small, which indicates that no conclusion can be drawn.

The impact of the potentials on the primary damage is confirmed for all variables. This means that one pair of two potentials does give different means for the response variables. Note that the conclusion would be different if we decreased the number of cascades in the data set. For example, in Fe, if one takes only 20 cascades per potential, the confidence of an effect of the potential on the mean of one descriptor decreases drastically: the p-value = 98% for the volume (because the mean values are close to each other) and the p-value = 3% for the maxSizeV and maxSizeI (because of their large variance). We now look for the percentage of the full dataset variance explained by the potentials. **Table 1** compares the variance for the full data set containing all the cascades obtained with all the potentials (2733 cascades for Fe and 4283

cascades for W), the variance of the residuals with the **pot** analysis and the R^2, the % of the full dataset variance explained by the potentials.

Table 1: variance of the primary damage descriptors for the full data set (Var_full), variance of the residuals with the **pot** analysis (Var_pot), $R^2 = 1 - \frac{Var_{pot}}{Var_{full}}$, % of the variance explained by **pot**. The color indicates if potentials are strongly or not responsible for the variance of the full data set.

	nDef	**monoV**	**monoI**	**nClusterV**	**nClusterI**	**maxSizeV**	**maxSizeI**
Fe							
Var_full	1595.3	1009.2	617.0	53.5	31.3	1893.9	536.5
Var_pot	1043.4	270.0	447.5	38.5	20.1	1667.2	482.2
R^2	35	73	27	28	36	12	10
W							
Var_full	553.3	325.7	212.9	14.4	18.5	81.3	110.1
Var_pot	284.7	101.0	80.7	9.7	9.3	72.3	75.8
R^2	49	69	62	33	50	11	31

strong impact of the potential
medium impact of the potential
small impact of the potential

Interestingly, Var_pot can be seen as an intrinsic variance due to the variation from one cascade to another within each potential dataset. R^2 **in Table 1** compares the variation due to the potentials and the intrinsic variation, and its values indicate that there are 3 categories of variables:

- monoV (and monoI for W) which has a larger variance due to the potential than the intrinsic variance.

- nDef, (monoI for Fe), nClusterV, nClusterI (and maxSizeI in W) which exhibit a variance due to the potential similar to the intrinsic variance.

- maxSizeV (and maxSizeI for Fe) which has a larger intrinsic variance within the full database than the one due to the potentials.

These findings indicate that all potentials agree on the total number of defects and it furthermore draws attention to the vacancies. We previously showed that the fraction of vacancies in clusters depends on the material. The analysis here adds clearly that this property is highly potential dependent.

The large intrinsic variations of maxSizeV and maxSizeI agree with their broad distributions. One immediate consequence is that the number of cascades necessary to get accurate results is large; the other one is that it can hide the impact of the potentials on these two properties, hence the need to use statistical tools.

In the next step, we increase the number of terms in the function that is used in the analysis. Now the reference is the analysis with 'pot' alone and we investigate the effect of the volume (analysis

"vol +pot" compared to "pot"), the effect of the sphericity (analysis "sphericity+pot" compared to "pot") and the effect the number of subcascades (analysis "nbSC + pot" compared to "pot").

Table 2 presents the variance of the residuals for the three analyses. By comparing the new variances of the residuals to the ones obtained with the analysis "pot", we obtain the % of variance explained by one aspect of the morphology. The results are given in **Table 2.**

For both Fe and W, we find that the volume and the sphericity have a significant explainability however, volume performs better in Fe and sphericity in W.

Table 2: Variance of the residuals and R^2 for the "vol+pot" analysis compared to the "pot" analysis, Var_vol+pot, $R^2 = 1 - \frac{Var\_vol+pot}{Var_{pot}}$; for the "sphericity+pot" analysis, Var_sphericity+pot, $R^2 = 1 - \frac{Var_{sphericity}+pot}{Var_{pot}}$ and for the "nbSC+vol" analysis, Var_nbSC+pot, $R^2 = 1 - Va\frac{r_{-nbSC+pot}}{Var_{pot}}$. The cells colored in red indicate the morphological descriptor which explains the best the variables.

	nDef	**monoV**	**monoI**	**nClusterV**	**nClusterI**	**maxSizeV**	**maxSizeI**
Fe							
Var_vol+pot	990.6	177.5	244.1	24.9	17.9	1531.8	457.2
R^2	**5**	**34**	**45**	**35**	11	8	5
Var_sphericity+pot	1031.3	216.0	361.9	32.1	17.5	1338.4	405.8
R^2	1	20	19	17	**13**	**20**	**16**
Var_nbSC+pot	1009.6	205.0	307.6	29.3	18.0	1556.7	462.5
R^2	3	24	31	24	11	7	4
W							
Var_vol+pot	257.9	91.7	72.0	9.1	8.8	72.1	75.6
R^2	**9**	9	11	**6**	5	0	0
Var_sphericity+pot	277.9	80.5	61.4	9.5	8.6	67.5	70.2
R^2	2	**20**	**24**	1	**8**	**7**	**7**
Var_nbSC+pot	268.8	90.8	71.5	9.4	8.7	72.3	75.7
R^2	6	10	11	3	7	0	0

Finally, we combined 2 morphological descriptors simultaneously and for completeness, all three descriptors together. The reference remains the 'pot' analysis. The new R^2 are presented in **Table 3**. In general, they are larger than in **Table 2** and in fact they are equal to the sum of the corresponding R^2 in **Table 2**, despite the fact that we found some correlation between the three morphological predictors. Note that taking the three descriptors together does not increase the R^2.

Table 3: variance of the residuals and R^2 for the "vol+sphericity+pot" analysis, Var_vol+sphericity+pot, $R^2 = 1 - \frac{Var\ \ vol+sphericity+pot}{Var_{pot}}$; for the "vol+sphericity+nbSC+pot" analysis, Var_vol+sphericity+nbSC+pot, $R^2 = 1 - \frac{Var\ \ vol+sphericity+nbSC+pot}{Var_{pot}}$. The cells coloured in red indicates which combination of morphological descriptors explains the best the variance of the primary damage descriptors.

	nDef	**monoV**	**monoI**	**nClusterV**	**nClusterI**	**maxSizeV**	**maxSizeI**
Fe							
Var_vol + sphericity+pot	977.8	125.4	162.3	18.8	15.3	1209.1	382.1
R^2	**6**	**54**	**64**	**51**	**24**	**27**	**21**
Var_vol + sphericity+nbSC+pot	974.6	125.2	161.5	18.8	15.3	1209.1	381.9
R^2	7	54	64	51	24	27	21
W							
Var_vol+sphericity+pot	252.3	72.3	53.8	9.0	8.1	67.2	69.8
R^2	**11**	**28**	**33**	**7**	**12**	**7**	**8**
Var_vol+sphericity+nbSC+pot	251.3	71.6	53.3	9.0	8.0	67.2	69.8
R^2	12	29	34	7	14	7	8

One interesting result is that the number of subcascades, nbSC, has no impact or very little impact on R^2. The number of subcascades is a descriptor which comes naturally to mind as it is easy to imagine that fragmented cascades do not produce the same sort of primary damage as compact ones. Furthermore, one would expect a multiple core cascade to not cool down the same way as a compact core cascade and that possibly, subcascade interaction would take place. However, our results show that it is not a good descriptor. Possible reasons are its integer nature which in our case implies a limited variation, especially for W and the fact that it does not capture the shape of the subcascades. These results are consistent with the discussion in [28] where the authors investigate the impact of subcascade formation. On the contrary, the sphericity, as we calculate it, is sensitive to the number of subcascades, has a large range of possible values and its variance is similar to the variance of the volume (6 - 15%).

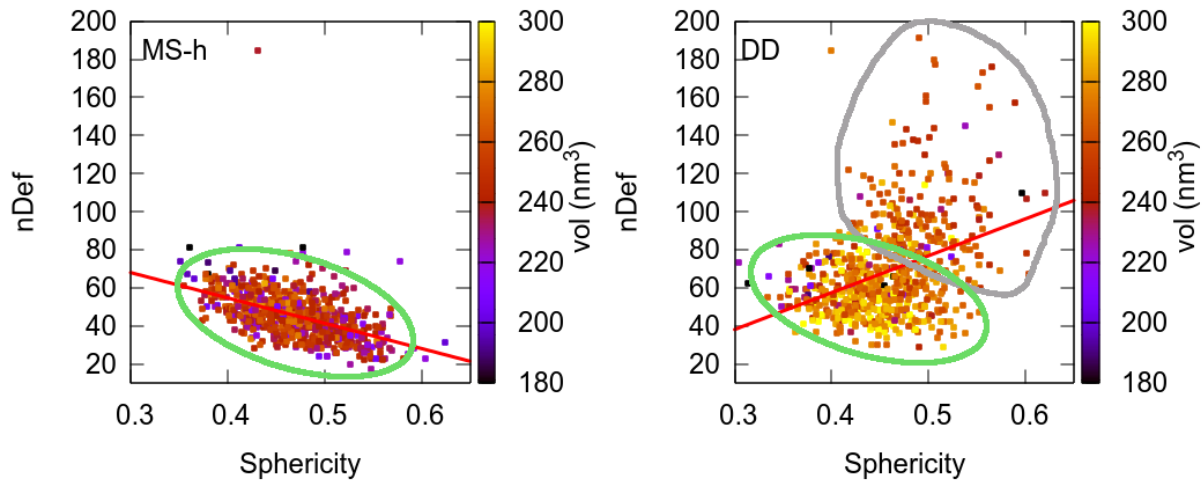
In the multivariate multiple linear regression analysis presented here, the interpretation of the regression coefficients is not trivial. On the one hand, our variables are not normalized and on the other hand, there is a non-negligible covariance between volume and sphericity. However, we observed that the coefficients of models with two morphological variables are similar to the ones in models with one morphological variable, indicating that the effect of volume and sphericity simply adds up. The sign of the coefficients, however, provides information on the correlations: a positive value means that there is a correlation between the data, and a negative value, indicates anti correlation.

The signs of the coefficients are presented in **Table 4**. Our results show that the larger the cascade and the more spherical, the lower the number of defects, the number of mono defects and the number of clusters.

Table 4: sign of the coefficients for the "vol" and "sphericity" variable in the multivariate multiple linear regression analysis.

	nDef	monoV	monoI	nClusterV	nClusterI	maxSizeV	maxSizeI
Fe							
Sign of the coefficient for the « vol » variable.	–	–	–	–	–	+	+
Sign of the coefficient for the «sphericity» variable	+	–	–	–	–	+	+
W							
Sign of the coefficient for the « vol » variable	–	–	–	–	–	–	–
Sign of the coefficient for the «sphericity» variable	–	–	–	–	–	+	+

The analysis of the sign shows thus that in Fe and W, the large cascades create less defects, less mono defects and less clusters. A difference can be noticed between Fe and W about the effect of the sphericity on the number of defects. This comes from one of the W potentials, DD, for which, the most spherical cascades give more defects, more mono vacancies (monoV) but less mono SIAs (monoI) and more clusters with larger maximum sizes (maxSizeV and maxSizeV) in contrast with the correlations revealed by the other potentials. The clouds of cascades for MS-h and DD potentials are compared in **Figure 7**. The possibility of atypical cascades is not excluded as it seems that there is one cloud of cascades with a slight negative correlation and a less dense cloud of spherical cascades creating a particularly large number of defects and thus reversing the sign of the correlation. These two clusters of data are clearly visible on **Figure 7.** A simple linear regression cannot capture this aspect of the data which requires the combination with a clustering algorithm.



(a) MS-h (b) DD

Figure 7: correlation plots between the number of defects (nDef) and the sphericity for MS-h (a) and DD (b). The red lines correspond to simple linear regression between the two variables.

For DD, two clusters of points are visible; the points which are within the grey lines correspond to atypical cascades and are responsible for the positive correlation with a simple linear model.

An interesting result is that, overall, R^2 are better for Fe than for W which indicates that the cascade morphology has a stronger impact in Fe than in W. Another result is the R^2 for the total number of defects, which is lower than the ones of the mono defects (particularly in Fe). This latter point is also probably a consequence of the occurrence of "atypical" cascades. The weight of atypical cascades probably depends on each variable. A dedicated study with a clustering algorithm is required. The R^2 for the maxSize are not very good either, however this can be related to their large variance.

The difference in R^2 between Fe and W does not arise from the fact that we chose different effective energies and thus from the fact that the two materials were observed in two different fragmented states. Indeed, fragmentation is taken into account into the model, since one of our descriptors is the sphericity and another one the number of subcascades. Furthermore, we did a similar R^2 analysis on cascade energies below the fragmentation limit in Fe with the MEND potential, and the predictability was similar to the results obtained at 80 keV despite the fact (and this is important) that the point defect clusters were small and thus the conditions very different from our present analysis. This, we believe, shows that it is not the fragmentation level which degrades the R^2.

The difference between 100% and the obtained R^2, indicates that there are some aspects of randomness of the cascades that are not accounted for in our analysis. It may be not related to the cascade morphology. However, one possibility is still that the maximum expansion of the cascade may not be the only crucial moment of the cascade development and it could be thus interesting to analyze other moments of the development. Furthermore, we did not include a descriptor of the subcascade spatial correlation in our models.

III. Discussion

1) Comparison with literature

Some of the results presented here seem at odds with most of what has been previously reported for iron from MD simulations using similar potentials. Indeed, it has generally been observed, in Fe, that the fraction of vacancies in clusters is much lower than that of interstitials and that larger interstitial clusters are formed than vacancy clusters. Our results indicate that for all the Fe potentials, the fraction of vacancies in clusters is higher (more than 70 % of vacancies are in clusters except for MEND10 and MA10A) than the fraction of SIAs in clusters (**Figure 1**) and that the largest clusters are always vacancy clusters in Fe (**Figure 3**).

To start the discussion, it is necessary to first point out that the large majority of results for Fe in the literature has been obtained with other and older Fe potentials. The only Fe potential we can compare our results to is the Mendelev potential (MEND10) for which we predict that the amount of vacancy in clusters is close to that of SIAs for reasonable PDC capture distances, i.e. capture distances lower than 0.5 nm.

We believe that there are many reasons behind the discrepancy regarding whichever point defect tends to cluster the most and form the largest clusters in Fe:

- The statistics: the data presented in our paper have been obtained with statistics on 1960 80 keV cascades in Fe (for MEND10). Aliaga et al. [29] presented the results of 14 cascades in bulk Fe, Byggmastar [23] database contained 50 cascades of 100 keV and Terentyev [30] had 10 cascades per damage energy.
- The PDC capture distance: the data presented in our paper have been obtained with a similar PDC capture distance for SIA and vacancies, taken to be between 2nn and 3nn. Aliaga et al. [29] presented the results of 14 cascades in bulk Fe with a PDC capture distance between 1nn and 2nn for both SIA clusters and vacancy clusters, whereas Byggmästar and Sand used (2nn+3nn)/2 for vacancies and (3nn and 4nn)/2 for SIAs [23], i.e. they used a larger PDC capture distance for SIAs than for vacancies. Terentyev [30] used a visual inspection method to extract the number of SIA clusters as "it has been observed that automated procedures tend to provide somewhat smaller fractions of SIA in clusters".
- The data analysis: if one represents our data, using exactly the same bins as Byggmastar, we also find that SIA clusters are larger and more numerous than vacancy clusters. However, if one, because of the statistics, has access to longer distribution tails, the results are different.

To illustrate this, we picked 50 randomly chosen 100 keV cascades done with the Mendelev potential and used the representation in bins chosen by Byggmastar et al. [23]. The results are presented in **Figure S3** in the supplementary and show that, we, also, find that there are more SIA clusters than vacancy clusters.

Another possible difference is the fact that our calculations do not take into account the electronic stopping power (ES); however, Sand et al. [31] clearly showed that, at least in W, the ES did not change much the ratio of the fraction of vacancy in clusters to that in SIA in clusters which was close to 0.3 (Table 1 in [31]) for the DD potential. Another difference between the results obtained by Byggmästar and ours is the fact that their simulations were run at 0 K, extracting the heat, whereas our cascades were simulated at 100 K, without any removal of the heat.
Finally, let us add that one of the first displacement cascade simulation in Fe, done by Beeler [32] using a potential derived by Erginsoy and co-workers [33] showed that, for damage energies ranging from 0.5 keV to 20 keV, close to 66 % of the vacancies were in clusters.

Our work thus shows that the methodology used until now by many groups to identify two defects as belonging to the same cluster, which relies only on the distance between the defects, is not sufficient to properly characterize the cluster. Indeed, beyond the number of point defects, the shape (loops, C15 and all kinds of imperfect structures) is also very important, raising the question of the classification of defects as shown for instance in [34,35]. Here, we describe in detail our methodology to be able to compare with the work reported in the literature, but the purpose here is not to propose a new methodology which would require a significant amount of statistics and which is out of the scope of the present study.

2) Comparison between damage prediction in W and in Fe

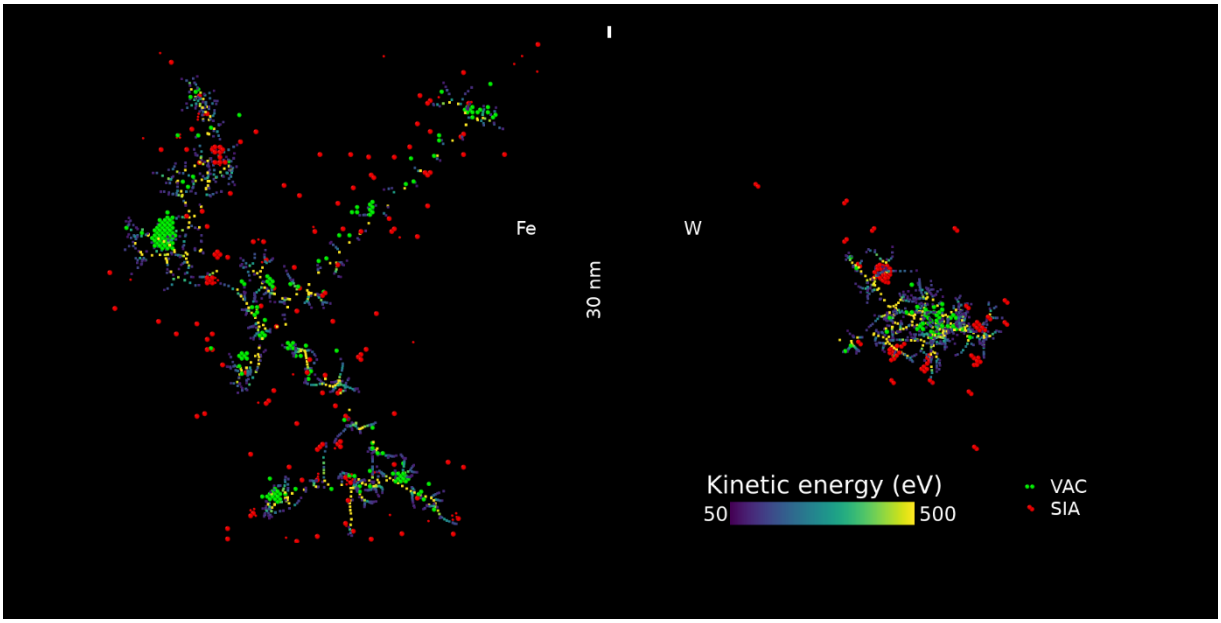
We have compared the primary damage obtained in W and Fe predicted by 6 and 7 potentials, respectively. We compare 50 keV cascades in W and 80 keV cascades in Fe. As mentioned previously, with this choice we compare cascades of similar volume but in two different damage energy regimes. 50 keV cascades in W correspond to the single cascade regime more or less, whereas in Fe, 80 keV is well above the subcascade threshold (15 keV [22]) and subcascades are formed. It is far from straight forward to compare Fe and W in the same fragmentation conditions because it would imply very different subcascade volumes. Furthermore, W and Fe have a range of more significant differences (atomic mass, melting temperature, ...). We are aware that some questions will thus remain unanswered until we report more comparisons with, for instance, other cascade energies.

We find that the number of defects predicted by all the potentials is three times larger in Fe than in W in agreement with the TDE and the damage energy differences between the two materials.

Regarding the fraction of defects in clusters, we find that in Fe, more vacancies than SIAs are in clusters (i.e., there are close to twice more mono SIAs than mono vacancies) in agreement with the picture that vacancies are created in the cascade core and thus they lie closer to one another. In W, it is the opposite: there are more mono vacancies than mono SIAs, i.e. the fraction of clustered SIAs is higher than the fraction of clustered vacancies. This later result is in agreement with what Fikar et al. [36] observed using the DD and AT potentials. Sand et al., [31] in a study of the impact of electronic stopping power in W, using the DD potential also found that, whatever the stopping power value they chose, the clustered fraction of SIAs in W was higher than that of vacancies. Setyawan, with three different W interatomic potentials [37], also observed a larger clustered fraction of SIAs but a smaller clustered fraction of vacancies.

The difference in clustering behavior is not due to temperature. Indeed, as W and Fe have very different melting temperatures, one can argue that our simulations were not done at the same temperature. However, according to Gao et al. [38], Stoller [39] and Setyawan et al. [37] higher temperature tends to increase the clustering fraction of SIAs. Furthermore, according to Setyawan, the impact of temperature in the clustering behavior is asymmetric: the SIA clustering increases with temperature, while the vacancy clustering decreases. Therefore, in Fe, where the cascades were simulated at a ‘higher relative temperature” than W, there should be more SIA clustering and less vacancy clustering. Temperature is therefore not the reason behind the difference in clustering behavior.
Examination of the evolution of the number of clusters during the cascade development shows that they exist quickly (before the first ps). For both Fe and W, the fraction of mono SIAs decreases during the full cascade process, because SIA clusters continuously form by agglomeration of the very mobile SIAs which also eliminate vacancies. However, the evolution of the fraction of mono vacancies is not similar for Fe and W. In W, it increases continuously, whereas in Fe, except for MEND10, which behaves similarly to W in that regard, it decreases in a manner similar to the population of SIAs. Thus, the fact that the fraction of vacancies in clusters is larger than the fraction of SIA in clusters in Fe (and the opposite in W) is the consequence of a continuous process initiated at the early stage of the cascade.

We looked at the possible location of the early created vacancies. A close examination of the cascade morphology shows that cascades in W are similar to Baobab trees, i.e. a thick trunk and almost no branches and leaves, whereas cascades in Fe look more like Joshua trees with thin trunks and pockets of leaves at the end of the trunks. At the end of the cascades, in W, the vacancy clusters form in the trunk and thus there are less of them, whereas in the Joshua tree like cascades, the vacancy clusters are mostly created in the "leaves" and thus are more numerous. This appears clearly on **Figure 8** where the green spheres indicate vacancies. The difference in the cascade morphology arises from, on the one hand, the stopping power which is three times lower in Fe than in W: high energy Fe atoms will create less vacancies, and the TDEs which are close to twice lower in Fe than in W: low energy atoms (i.e. the atoms in the "leaves of our trees") will create more defects in Fe than in W. The stopping power does have an impact on the fragmentation state of the cascade, but as mentioned above, the Joshua tree like versus Baobab tree like morphologies observed do not come from the fact that the cascade energies have been chosen in a different subcascade regime for Fe and W but from the two reasons exposed above.



Fe Ma10A W MS-h

Figure 8: typical cascade morphologies: red spheres represent the SIAs, green spheres represent vacancies. For these pictures, 30 frames have been selected between 0 and 0.2 ps. The cubes represent positions where one atom, in any of the frames, has a kinetic energy higher than 40 eV in Fe and 80 eV in W. The 30 frames are superimposed here, so as to underline the trajectories of atoms of energy large enough to create one vacancy.

In Fe, vacancy clusters are created right from the beginning of the cascades and remain there during the cooling phase, whereas in W, vacancy clusters form, disappear and reappear elsewhere as the cores cool down. Despite the fact that the overall volumes of the cascades are similar, the cores in W are much bigger than those in Fe. There are thus pockets of vacancy clusters that appear and disappear during the cooling of the cascade cores. This can only happen in large and dense cascade cores and this mechanism is not consistent with the typical picture of a recrystallization front propagating inward from the outside of the cascade as proposed for instance by Rubia and coworkers [40].

This study also shows that within a specific element (Fe or W), the number of defects predicted by each potential agrees with one another within a 50% margin (**Figure 1**). Furthermore, all potentials predict similar SIA cluster sizes and vacancy cluster sizes that differ by a factor two at the most and the sizes of the largest clusters in each distribution also differ by a factor two at the most. The multivariate multiple analysis actually suggests that a consensus based on all potentials can be found to build statistical models of the main features of the primary damage. However, another outcome of this study is the impact of atypical cascades, and more generally, the clustering of the data which needs to be accounted for.

Conclusions

We have characterized a large panel of EAM potentials for W and Fe, differing either by their equilibrium part or the way they were hardened. In the companion paper to this one, we have found correlations between TDE, QSD and RCS along the 〈110〉 direction. Here, after analyzing a very large cascade database, we show that another correlation exists between the average number of point defects created by a PKA of given energy and the mean TDE values along the 〈100〉, 〈110〉 and 〈135〉 directions: as could be expected, the lower the TDE, the more defects are created.

The analysis of a large statistics for 80 keV cascades in Fe and 50 keV cascades in W for all potentials, representing a total of about 7000 displacement cascades allowed to characterize qualitatively and quantitatively the role of the cascade morphology on the primary damage. The statistical analysis of the cascade database indicates that the sphericity of the cascades is a good predictor of the primary damage, in particular in W. The best model is the combination of the volume and the sphericity for both Fe and W. Among the properties investigated, we showed that soft potentials produce larger cascades, which are less spherical, more subcascades, more defects and more defect clusters. They also predict smaller maximum cluster sizes than the other potentials. Another outcome of this study is the impact of atypical cascades, and more generally, the clustering of the data.

Comparing cascade debris in W and Fe, we find that Fe potentials predict a larger fraction of vacancies than SIAs in clusters, and they also predict larger vacancy clusters than SIA clusters. W potentials predict larger SIA clusters than vacancy clusters or clusters of similar sizes and a larger fraction of SIA than vacancy in clusters. The difference in the clustering behavior between the two metals arises from the difference in their mass and their TDEs as well as stopping power which lead to very distinct cascade morphologies. Comparing cascades with trees, W cascades look like Baobabs, with thick trunks, where the vacancy clusters form, whereas Fe cascades look more like Joshua trees, with thinner trunks and small but numerous branches where the vacancy clusters form.

This study also shows that within a specific element (Fe or W), the number of defects predicted by each potential agrees with one another up to 50% margin; all potentials predict similar SIA

cluster sizes and vacancy cluster sizes that differ by a factor two at the most and the same applies for the size of the largest clusters in each distribution.

Another important conclusion that can be drawn from our study is that a methodology that identifies two defects as belonging to the same cluster which relies only on the distance between the defects is not sufficient and a new methodology should be defined.

Beyond the detailed description of the difference between potentials, our multivariate multiple analysis performed on one energy per material of a large cascade database is a new step of understanding of the primary damage, which actually suggests that a consensus based on all potentials can be found to build statistical models of the main features of the primary damage to model microstructure evolution.

Acknowledgements

The authors wish to thank R. E. Stoller for the numerous discussions we had regarding this work which contributed to improve a great deal the manuscript. This work has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. This work was also partly supported within the European project SOTERIA (661913) and contributes to the Joint Programme on Nuclear Materials (JPNM) of the European Energy Research Alliance (EERA).

Data availability

The complete raw/processed data required to reproduce these findings cannot be shared at this time due to technical or time limitations but are available upon request.
Some of the cascades (final atomic positions at the end of the displacement cascade simulation) are available on the IAEA cascade database (https://cascadesdb.org/).

References

[1] C.S. Becquart, C. Domain, Modeling Microstructure and Irradiation Effects, Metall. Mater. Trans. A. 42 (2011) 852–870. https://doi.org/10.1007/s11661-010-0460-7.
[2] C.S. Becquart, A. De Backer, P. Olsson, C. Domain, Modelling the primary damage in Fe and W: influence of the short range interactions on the cascade properties: part 1 – energy transfer., accepted. (n.d.).
[3] J.F. Ziegler, J.P. Biersack, The Stopping and Range of Ions in Matter, in: D.A. Bromley (Ed.), Treatise Heavy-Ion Sci., Springer US, Boston, MA, 1985: pp. 93–129. https://doi.org/10.1007/978-1-4615-8103-1_3.

[4] C.S. Becquart, C. Domain, A. Legris, J.C. Van Duysen, Influence of the interatomic potentials on molecular dynamics simulations of displacement cascades, J. Nucl. Mater. 280 (2000) 73–85. https://doi.org/10.1016/S0022-3115(00)00029-5.
[5] A. Souidi, M. Hou, C.S. Becquart, C. Domain, Atomic displacement cascade distributions in iron, J. Nucl. Mater. 295 (2001) 179–188. https://doi.org/10.1016/S0022-3115(01)00556-6.
[6] C.S. Becquart, K.M. Decker, C. Domain, J. Ruste, Y. Souffez, J.C. Turbatte, J.C. Van Duysen, Massively parallel molecular dynamics simulations with EAM potentials, Radiat. Eff. Defects Solids. 142 (1997) 9–21. https://doi.org/10.1080/10420159708211592.
[7] E. Zarkadoula, D.M. Duffy, K. Nordlund, M.A. Seaton, I.T. Todorov, W.J. Weber, K. Trachenko, Electronic effects in high-energy radiation damage in tungsten, J. Phys. Condens. Matter. 27 (2015) 135401. https://doi.org/10.1088/0953-8984/27/13/135401.
[8] R.E. Stoller, G.R. Odette, B.D. Wirth, Primary damage formation in bcc iron, J. Nucl. Mater. 251 (1997) 49–60. https://doi.org/10.1016/S0022-3115(97)00256-0.
[9] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO–the Open Visualization Tool, Model. Simul. Mater. Sci. Eng. Httpovitoorg. 18 (2010) 015012. https://doi.org/10.1088/0965-0393/18/1/015012.
[10] M.W. Finnis, J.E. Sinclair, A simple empirical *N* -body potential for transition metals, Philos. Mag. A. 50 (1984) 45–55. https://doi.org/10.1080/01418618408244210.
[11] D.R. Mason, D. Nguyen-Manh, C.S. Becquart, An empirical potential for simulating vacancy clusters in tungsten, J. Phys. Condens. Matter. 29 (2017)505501. https://doi.org/10.1088/1361-648X/aa9776.
[12] N. Juslin, B.D. Wirth, Interatomic potentials for simulation of He bubble formation in W, J. Nucl. Mater. 432 (2013) 61–66. https://doi.org/10.1016/j.jnucmat.2012.07.023.
[13] P.M. Derlet, D. Nguyen-Manh, S.L. Dudarev, Multiscale modeling of crowdion and vacancy defects in body-centered-cubic transition metals, Phys. Rev. B. 76 (2007) 054107. https://doi.org/10.1103/PhysRevB.76.054107.
[14] C. Björkas, K. Nordlund, S. Dudarev, Modelling radiation effects using the ab-initio based tungsten and vanadium potentials, Nucl. Instrum. Methods Phys. Res. Sect. B Beam Interact. Mater. At. 267 (2009) 3204–3208. https://doi.org/10.1016/j.nimb.2009.06.123.
[15] G.J. Ackland, R. Thetford, An improved N-body semi-empirical model for body-centred cubic transition metals, Philos. Mag. A. 56 (1987) 15–30. https://doi.org/10.1080/01418618708204464.
[16] Y. Zhong, K. Nordlund, M. Ghaly, R.S. Averback, Defect production in tungsten: A comparison between field-ion microscopy and molecular-dynamics simulations, Phys. Rev. B. 58 (1998) 2361–2364. https://doi.org/10.1103/PhysRevB.58.2361.
[17] M.-C. Marinica, L. Ventelon, M.R. Gilbert, L. Proville, S.L. Dudarev, J. Marian, G. Bencteux, F. Willaime, Interatomic potentials for modelling radiation defects and dislocations in tungsten, J. Phys. Condens. Matter. 25 (2013) 395502. https://doi.org/10.1088/0953-8984/25/39/395502.
[18] A.E. Sand, J. Dequeker, C.S. Becquart, C. Domain, K. Nordlund, Non-equilibrium properties of interatomic potentials in cascade simulations in tungsten, J. Nucl. Mater. 470 (2016) 119–127. https://doi.org/10.1016/j.jnucmat.2015.12.012.
[19] L. Malerba, M.C. Marinica, N. Anento, C. Björkas, H. Nguyen, C. Domain, F. Djurabekova, P. Olsson, K. Nordlund, A. Serra, D. Terentyev, F. Willaime, C.S. Becquart, Comparison of empirical interatomic potentials for iron applied to radiation damage studies, J. Nucl. Mater. 406 (2010) 19–38. https://doi.org/10.1016/j.jnucmat.2010.05.017.

[20] G.J. Ackland, M.I. Mendelev, D.J. Srolovitz, S. Han, A.V. Barashev, Development of an interatomic potential for phosphorus impurities in -iron, J. Phys. Condens. Matter. 16 (2004) S2629–S2642. https://doi.org/10.1088/0953-8984/16/27/003.
[21] R. Alexander, L. Proville, C.S. Becquart, A.M. Goryeava, J. Dérès, C. Lapointe, M.-C. Marinica, Interatomic potentials for irradiation-induced defects in iron, J. Nucl. Mater. (2020) 152141. https://doi.org/10.1016/j.jnucmat.2020.152141.
[22] A. De Backer, C. Domain, C.S. Becquart, L. Luneville, D. Simeone, A.E. Sand, K. Nordlund, A model of defect cluster creation in fragmented cascades in metals based on morphological analysis, J. Phys. Condens. Matter. 30 (2018) 405701. https://doi.org/10.1088/1361-648X/aadb4e.
[23] J. Byggmästar, F. Granberg, K. Nordlund, Effects of the short-range repulsive potential on cascade damage in iron, J. Nucl. Mater. 508 (2018) 530–539. https://doi.org/10.1016/j.jnucmat.2018.06.005.
[24] A.E. Sand, D.R. Mason, A.D. Backer, X. Yi, S.L. Dudarev, K. Nordlund, Cascade fragmentation: deviation from power law in primary radiation damage, Mater. Res. Lett. 5 (2017) 357–363. https://doi.org/10.1080/21663831.2017.1294117.
[25] C. Domain, A. De Backer, C.S. Becquart, Statistical analysis of MD cascades, Prep. (2020).
[26] A.F. Calder, D.J. Bacon, A.V. Barashev, Y.N. Osetsky, Effect of mass of the primary knock-on atom on displacement cascade debris in α -iron, Philos. Mag. Lett. 88 (2008) 43–53. https://doi.org/10.1080/09500830701733004.
[27] A. De Backer, A.E. Sand, K. Nordlund, L. Luneville, D. Simeone, S.L. Dudarev, Subcascade formation and defect cluster size scaling in high-energy collision events in metals, EPL Europhys. Lett. 115 (2016) 26001. https://doi.org/10.1209/0295-5075/115/26001.
[28] R.E. Stoller, L.R. Greenwood, Subcascade formation in displacement cascade simulations: Implications for fusion reactor materials, J. Nucl. Mater. 271–272 (1999) 57–62. https://doi.org/10.1016/S0022-3115(98)00730-2.
[29] M.J. Aliaga, I. Dopico, I. Martin-Bragado, M.J. Caturla, Influence of free surfaces on microstructure evolution of radiation damage in Fe from molecular dynamics and object kinetic Monte Carlo calculations: Influence of free surfaces on microstructure evolution of radiation damage in Fe, Phys. Status Solidi A. 213 (2016) 2969–2973. https://doi.org/10.1002/pssa.201600158.
[30] D. Terentyev, C. Lagerstedt, P. Olsson, K. Nordlund, J. Wallenius, C.S. Becquart, L. Malerba, Effect of the interatomic potential on the features of displacement cascades in α-Fe: A molecular dynamics study, J. Nucl. Mater. 351 (2006) 65–77. https://doi.org/10.1016/j.jnucmat.2006.02.020.
[31] A.E. Sand, K. Nordlund, S.L. Dudarev, Radiation damage production in massive cascades initiated by fusion neutrons in tungsten, J. Nucl. Mater. 455 (2014) 207–211. https://doi.org/10.1016/j.jnucmat.2014.06.007.
[32] J.R. Beeler, Vacancy and Interstitial Cluster Production in Neutron-Irradiated α Iron, J. Appl. Phys. 37 (1966) 3000–3009. https://doi.org/10.1063/1.1703153.
[33] C. Erginsoy, G.H. Vineyard, A. Englert, Dynamics of Radiation Damage in a Body-Centered Cubic Lattice, Phys. Rev. 133 (1964) A595–A606. https://doi.org/10.1103/PhysRev.133.A595.
[34] U. von Toussaint, F.J. Domínguez-Gutiérrez, M. Compostella, M. Rampp, FaVAD: A software workflow for characterization and visualizing of defects in crystalline structures, Comput. Phys. Commun. 262 (2021) 107816. https://doi.org/10.1016/j.cpc.2020.107816.

[35] U. Bhardwaj, A.E. Sand, M. Warrier, Classification of clusters in collision cascades, Comput. Mater. Sci. 172 (2020) 109364. https://doi.org/10.1016/j.commatsci.2019.109364.

[36] J. Fikar, R. Schäublin, Molecular dynamics simulation of radiation damage in bcc tungsten, J. Nucl. Mater. 386–388 (2009) 97–101. https://doi.org/10.1016/j.jnucmat.2008.12.068.

[37] W. Setyawan, G. Nandipati, K.J. Roche, H.L. Heinisch, B.D. Wirth, R.J. Kurtz, Displacement cascades and defects annealing in tungsten, Part I: Defect database from molecular dynamics simulations, J. Nucl. Mater. 462 (2015) 329–337. https://doi.org/10.1016/j.jnucmat.2014.12.056.

[38] F. Gao, D.J. Bacon, P.E.J. Flewitt, T.A. Lewis, A molecular dynamics study of temperature effects on defect production by displacement cascades in α-iron, J. Nucl. Mater. 249 (1997) 77–86. https://doi.org/10.1016/S0022-3115(97)00178-5.

[39] R.E. Stoller, The role of cascade energy and temperature in primary defect formation in iron, J. Nucl. Mater. 276 (2000) 22–32. https://doi.org/10.1016/S0022-3115(99)00204-4.

[40] T.D. de la Rubia, K. Smalinskas, R.S. Averback, I.M. Robertson, H. Hseih, R. Benedek, Mechanisms of Cascade Collapse, MRS Online Proc. Libr. OPL. 138 (1988) 29-34. https://doi.org/10.1557/PROC-138-29.