



PAPER

OPEN ACCESS

RECEIVED

29 September 2022

REVISED

21 December 2022

ACCEPTED FOR PUBLICATION

9 January 2023

PUBLISHED

24 January 2023

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



MINAS TIRITH: a new tool for simulating radiation-induced DNA damage at the cell population level

Y Thibaut¹ , G Gonon¹, J S Martinez¹, M Petit¹ , A Vaurijoux¹, G Gruel¹, C Villagrasa¹, S Incerti² and Y Perrot^{1,*}

¹ Institut de Radioprotection et de Sécurité Nucléaire (IRSN), PSE-SANTE/SDOS/LDRI, PSE-SANTE/SERAMED/LRAcc, PSE-SANTE/SDOS/LMDN, BP 17, F-92262 Fontenay-aux-Roses, France

² Université de Bordeaux, CNRS/IN2P3, LP2i, UMR 5797, F-33170 Gradignan, France

* Author to whom any correspondence should be addressed.

E-mail: yann.perrot@irsn.fr

Keywords: microdosimetry, nanodosimetry, DNA damage, geant4-DNA, cell population

Abstract

Objective. The mechanisms of radiation-induced DNA damage can be understood via the fundamental acquisition of knowledge through a combination of experiments and modeling. Currently, most biological experiments are performed by irradiating an entire cell population, whereas modeling of radiation-induced effects is usually performed via Monte Carlo simulations with track structure codes coupled to realistic DNA geometries of a single-cell nucleus. However, the difference in scale between the two methods hinders a direct comparison because the dose distribution in the cell population is not necessarily uniform owing to the stochastic nature of the energy deposition. Thus, this study proposed the MINAS TIRITH tool to model the distribution of radiation-induced DNA damage in a cell population. **Approach.** The proposed method is based on precomputed databases of microdosimetric parameters and DNA damage distributions generated using the Geant4-DNA Monte Carlo Toolkit. First, a specific energy z was assigned to each cell of an irradiated population for a particular absorbed dose D_{abs} , following microdosimetric formalism. Then, each cell was assigned a realistic number of DNA damage events according to the specific energy z , respecting the stochastic character of its occurrence. **Main results.** This study validated the MINAS TIRITH tool by comparing its results with those obtained using the Geant4-DNA track structure code and a Geant4-DNA based simulation chain for DNA damage calculation. The different elements of comparison indicated consistency between MINAS TIRITH and the Monte Carlo simulation in case of the dose distribution in the population and the calculation of the amount of DNA damage. **Significance.** MINAS TIRITH is a new approach for the calculation of radiation-induced DNA damage at the cell population level that facilitates reasonable simulation times compared to those obtained with track structure codes. Moreover, this tool enables a more direct comparison between modeling and biological experimentation.

1. State of the art

Ionizing radiation induces both direct and indirect interactions with the environment through which it passes. In living organisms, these interactions can affect all organelles. Among these radiation-induced damages, initial DNA damage is of importance as its evolution can result in genetic alterations. Consequently, cell dysfunctions may be induced, which can lead to deterministic or stochastic biological effects at the macroscopic scale. Various forms of DNA damage may be generated, with the most dangerous being DNA double-strand break (DSB) (Khanna and Jackson 2001, Rothkamm and Lobrich 2002). In addition, biological studies have shown that radiation quality plays a crucial role (Gonon *et al* 2019). In most cases, these biological experiments are performed at the scale of a cell population, and their results are therefore subject to the stochastic character of the

radiation-matter interaction (Gruel *et al* 2016). In parallel, following modeling using Monte Carlo track structure codes, the simulations were compared to the experimental results and new knowledge on the initial damage formation were obtained (Uehara *et al* 1993, Friedland *et al* 2003, Meylan *et al* 2016, Lampe *et al* 2018, Gonon *et al* 2019). Experimentation and modeling are two complementary tools used to understand the link between energy deposition and initial radiation-induced damage to predict the biological effect of radiation quality. Nevertheless, comparison of the results obtained using these two tools is complex. Although the scale difference between the two is identified as a necessary element (Chatzipapas *et al* 2020), it induces a bias according to the variability of the energy allocated to each cell.

At IRSN, a nanodosimetric simulation chain was developed to model the topology of radiation-induced DNA damage in a single-cell nucleus produced by a given radiation quality (Meylan *et al* 2016, Tang *et al* 2019a). This chain is based on the Geant4-DNA toolkit (Incerti *et al* 2010a, 2010b, Bernal *et al* 2015, Incerti *et al* 2018), with the aim of simulating the physical stage of irradiation from which direct damage occurs, as well as the chemical stage, which induces indirect damage. The modeling of these two steps, coupled with a realistic nuclear DNA geometry, is essential for radiation-induced damage calculation (double-strand breaks, single-strand breaks, and base damages) (Meylan *et al* 2017, Tang *et al* 2019b, Thibaut *et al* 2022). However, the computation times for these types of simulations are considerably long, and the geometries used to characterize the DNA molecule are highly memory-consuming. Therefore, this approach cannot be extended to include the DNA geometries of several cells without compromising the level of detail or accuracy of the results.

However, in contrast to modeling, biological experiments are most often performed at the cell population scale (Chatzipapas *et al* 2020). Therefore, the validation of Monte Carlo codes from these experimental data can only be performed based on mean results (McNamara *et al* 2017), occulting in the stochastic nature of DNA damage induction. Thus, the variability of intercellular responses was not represented by these simulation codes. This variability is because of the response of each cell and can also be correlated to the distribution of the energy imparted to each cell of the same cell population because of the stochastic nature of the interaction between radiation and matter (Gruel *et al* 2016). To address this problem, several solutions have been proposed. For example, cell survival models have been proposed based directly on the microdosimetric characterization of irradiation (Sato and Furusawa 2012, Manganaro *et al* 2017, Bellinzona *et al* 2021, Cordoni *et al* 2022). However, these models did not simulate a detailed description of the induction of DNA damage. (Douglass *et al* 2015) suggested the computation of a tumor-scale dose distribution using the Livermore model included in Geant4, coupled with a damage calculation at the cellular scale using the Geant4-DNA tool. However, no damage calculations were performed from DNA geometry, rather they were based on thresholding the clustered ionization energies according to their position in a water-equivalent nucleus. Baiocco *et al* (2016) offered a similar method with a damage calculation based on analytical formulae determined considering the PARTRAC track structure simulation code, which involves complex DNA geometries. Similarly, the CPOP tool (Maigne, *et al* 2021) allows the construction of realistic spheroid geometries coupled with Geant4-DNA. Although it has been designed to model dose distribution in *in vitro* systems, nanodosimetric calculations in such biological structures have not yet been implemented. In all cases, the solutions resort to Monte Carlo modeling of the irradiation configuration, which incurs high costs in terms of computing time, memory consumption, and geometry design.

This study proposed the Microdosimetry and Nanodosimetry to simulate the heterogeneity of the initial radiation-induced damage topology (MINAS TIRITH) tool, which allows the calculation of a realistic DNA damage distribution in a cell population according to a particular radiation quality. It can handle complex radiation qualities, combinations of different irradiation angles, and different types of particles, for example, over a large energy range: 1 keV–20 MeV for electrons and 10 keV–20 MeV for protons and α particles. However, MINAS TIRITH only considers the variability of cellular responses related to the distribution of energy deposition and not to cell nucleus shape variability (geometry or volume) within the cell population or other types of individual variability. Thus, to quantify this energy deposition variability within the same cell population, a microdosimetry formalism was applied (Kellerer and Chmelevsky 1975a, 1975b, Rossi and Zaider 1996). The method implemented in MINAS TIRITH is based on the combination and sampling of the spectra. In brief, the first set of spectra was reconstructed from a database calculated through microdosimetric simulations using Geant4 and Geant4-DNA, whereas the second set was calculated based on the results of the IRSN radiation-induced damage simulation chain. Thus, within acceptable simulation times, the MINAS TIRITH tool provides DNA damage distributions that can be easily compared with biological irradiation experiments on cell populations. MINAS TIRITH is currently limited to a single-nucleus geometry, representing the endothelial cell type. This type of cell nucleus was chosen because the development of MINAS TIRITH is a part of IRSN's largest project aimed at better understanding the side effects of hadrontherapy in healthy tissues, particularly the inflammatory processes wherein endothelial cells are a key target.

This study employed a method that was used in MINAS TIRITH to distribute damage in a cell population. Further, it was verified that each step of the method provided results similar to those obtained using the damage

simulation chain. The purpose of this validation was the evaluation of the degree of confidence that can be assigned to the results produced by MINAS TIRITH compared to those produced by Monte Carlo simulation methods using track structure codes. This can aid in future comparisons with experimental biological results obtained from cell populations.

2. Materials and methods

2.1. Construction of the databases

2.1.1. The microdosimetric database

In the first step, a database of microdosimetric quantities was built for monoenergetic particles: electrons from 1 keV–21 MeV, protons from 10 keV to 20 MeV, and alpha rays from 10 keV to 21 MeV. Geant4 (Agostinelli *et al* 2003, Allison *et al* 2006, 2016) and Geant4-DNA tools (Incerti *et al* 2010a, 2010b, Bernal *et al* 2015, Incerti *et al* 2018), version 10.6, were used to build this database. For each sampled primary energy and for each particle type, a simulation was performed including a target volume corresponding to the mean geometry of a human endothelial cell nucleus that was represented by an elliptical cylinder measuring 2 μm in height, with half major axis of 9.5 μm and half minor axis of 5.1 μm . The emission points of the primary particles were sampled with a uniform probability over the entire surface of the elliptical cylinder in the isotropic direction. Simulations were performed using the default option of the Geant4-DNA physics constructor (option 2). However, electrons with energies greater than 1 MeV were not supported by Geant4-DNA and thus the standard electromagnetic physics constructor option 4 of Geant4 was used after checking the continuity of the results between the two configurations. The number of initial particles in each simulation was fixed to have sufficient statistical power while maintaining a reasonable simulation time, considering that described in the literature (Parisi *et al* 2022): 2×10^7 primaries for electrons and 2×10^6 primaries for protons and alphas. In each simulation, the energy imparted (ε_i) per event for each particle and its secondaries in the elliptical cylinder representing the cell nucleus, as well as the chord length of the primary particle (CL_i), were recorded. Chord length is defined as the distance between the entrance and exit points, or the stopped point, of the primary particle. The truncated lineal energy y_{CL_i} of each particle was then calculated according to:

$$y_{CL_i} = \frac{\varepsilon_i}{CL_i}.$$

This definition of lineal energy does not correspond to the standard definition because it introduces chord length and not mean chord length. Indeed, this definition decorrelates energetic stochasticity to volumetric stochasticity. This definition of lineal energy is possible in case of simulations; however, it is more difficult in case of experimental measurements. In other studies, the quantity y_T , obtained by calculating the ratio between the imparted energy ε_i and track length of each particle, has been identified as a good predictor of biological effectiveness (Vassiliev *et al* 2020). Moreover, it can be experimentally measured (Missiaggia, *et al* 2021, 2022). However, the y_T definition does not correspond to the use for which the spectra are obtained because the track length cannot be reconstructed geometrically. This is more relevant in case of electrons. Indeed, they were intended to find the energy deposit ε_i through the multiplication with a geometrical chord length calculated from the position and direction of entry of the particle in the nuclear volume.

Thus, two subpopulations were formed among the y_{CL_i} to differentiate the particles that stop within the cell nucleus from the others. The first population was composed of the y_{CL_i} , for which the primary particle has exited the elliptical cylinder without having deposited all its energy; the y_i of this population is denoted $^{exit}y_{CL_i}$. The y_{CL_i} of the second population, for which the primary particles stopped in the elliptical cylinder, is denoted $^{stop}y_{CL_i}$. The population $^{stop}y_{CL_i}$ was considered non-negligible (proportion greater than 5%) for electrons between 1 keV and 16 keV, for protons from 10 to 800 keV, and for alphas from 10 keV to 3.5 MeV. From these two populations, the $f(^{exit}y_{CL})$ and $f(^{stop}y_{CL})$ spectra were constructed for each sampled energy and particle type. These spectra were then integrated, spline-smoothed, and scaled into 1909 points with ordinates between 0 and 1 to constitute the spectra databases $F(^{exit}y_{CL})$ and $F(^{stop}y_{CL})$. The number of points was owing to a dynamic sampling step that was necessary to ensure a correct reconstruction by interpolation of sigmoid-shaped curves. Its value was obtained after several trials to determine the best compromise for the reconstruction.

Simultaneously, over the energy and particle range where the population of stopped particles in the elliptical cylinder is non-negligible, a database of $F(^{stop}CL)$ was obtained for the chord lengths using the same method as described for $F(^{stop}y_{CL})$.

2.1.2. The damage data base

In parallel, a database of damage was built for different mono-energies sampled between 10 keV and 20 MeV for protons, 10 keV–21 MeV for alpha rays, and 1 keV–1 MeV for electrons. To construct this database, a

nanodosimetric simulation chain developed previously was used (Meylan, *et al* 2016), coupled to the Isochore cell nucleus geometry (Thibaut *et al* 2022). This simulation chain facilitated the calculation of the topology of radiation-induced damage at the nuclear scale through the coupling of realistic DNA geometries to nanodosimetric simulations, using the track structure code Geant4-DNA, version 10.6. The number of simulated particles for each sampled energy was chosen to have sufficient statistical power: 25 000 for electrons, 1500 for alphas, and between 1500 and 10 000 for protons depending on the LET. Regarding the irradiation geometry, in case of the microdosimetric database, the particles are emitted from the surface of the nucleus, in an isotropic manner. In these simulations, the number of radiation-induced damages ($nDSB_i$, $nSSB_i$) was determined for each track of each projectile type. The chord length (CL_i) and imparted energy (ϵ_i) of each track in the nucleus were extracted.

Based on the data extracted from these simulations, spectra in terms of $nDSB/\epsilon$ and $nSSB/\epsilon$ (denoted $F(nDSB/\epsilon)$ and $F(nSSB/\epsilon)$) were constructed for each sampled energy. In addition, the cumulative probability spectra of DSB and SSB occurrence as a function of chord length (denoted $F_{DSB}(CL)$ and $F_{SSB}(CL)$) were established.

3. Minas tirth tool

3.1. Distribution of damages in the cell population

3.1.1. Distribution of tracks per cell

The MINAS TIRITH tool calculated the damage topology within a cell population associated with a given irradiation. The inputs required for this calculation were as follows:

- Number of cells in the population (N_{cells}).
- Macroscopic absorbed dose delivered to the population (D_{abs}).
- Energy spectrum for each type of directly ionizing particle (p) entering the cell nuclei ($f_p(E)$).
- Angle spectrum for each type of directly ionizing particle (p) entering the cell nuclei ($f_p(\theta)$).
- Weight (W_p) of each type of ionizing particle (p) in terms of the proportion of the number of particles entering.

Therefore, the MINAS TIRITH tool requires prior calculation of the phase space at the interface between the cell nuclei and culture medium. Because of these elements and based on the microdosimetric database, MINAS TIRITH realistically distributes the number of tracks for each type of particle p in the phase space observed by each cell nucleus in the population according to the following method:

- The geometric chord length spectrum $f_p(CL^{geo})$, of the irradiation was calculated from the dimensions of the endothelial cell nucleus model and the $f_p(\theta)$ spectrum. This geometric chord length corresponds to the distance between the entrance and exit points in the direction of the particle.
- For each energy constituting the $f_p(E)$ spectrum, the $F_{p,E_l}^{(exit)}(\widehat{\gamma_{CL}})$ spectra were reconstructed from the database of $F^{(exit)}(\gamma_{CL})$ spectra. This is performed through the spline interpolation of the value $\widehat{\gamma_{CLk}}^{exit}$ from the values γ_k^{exit} for the K ordinates ($K = 1909$) of the database.
- Step ii was repeated to reconstruct the $F_{p,E_l}^{(stop)}(\widehat{\gamma_{CL}})$ and $F_{p,E_l}^{(stop)}(\widehat{CL})$ spectra from the $F^{(stop)}(\gamma_{CL})$ and $F^{(stop)}(CL)$ spectra in the database.
- The mean values of each spectrum, $F_{p,E_l}^{(exit)}(\widehat{\gamma_{CL}})$ and $F_{p,E_l}^{(stop)}(\widehat{\gamma_{CL}})$, which are denoted by $\overline{\widehat{\gamma_{CLf}}^{exit}}_{p,E_l}$ and $\overline{\widehat{\gamma_{CLf}}^{stop}}_{p,E_l}$, respectively, were calculated.
- The $\overline{\gamma_f}$ value was approximated using the following formula

$$\overline{\gamma_f} = \sum_{p=1}^P \sum_{l=1}^L \sum_{j=1}^J W_p \times f_p(E_l) \times f_p(CL_j^{geo}) \times \frac{\overline{\widehat{\gamma_{CLf}}^{exit}}_{p,E_l} \times CL_j^{geo} \times (1 - F_{p,E_l}^{(stop)}(\widehat{CL_j^{geo}})) + \overline{\widehat{\gamma_{CLf}}^{stop}}_{p,E_l} \times CL_j^{geo} \times F_{p,E_l}^{(stop)}(\widehat{CL_j^{geo}})}{m_{nucleus}},$$

where $m_{nucleus}$ is the mass of the endothelial cell nucleus, $(F(\widehat{CL_j^{geo}}))$ is the chance for the particle to stop in the cell nucleus geometry, and thus $(1 - F(\widehat{CL_j^{geo}}))$ is the chance for an exit. The summation over j allows the coverage of all possible chord lengths in the density of probability $f_p(CL_j^{geo})$.

vi. A number of tracks were assigned to each cell nucleus in the population by sampling the Poisson law of parameter $\lambda = \frac{D_{abs}}{\bar{z}_f}$.

3.1.2. Determination of the number of damages assigned to each track

At this stage, for a given dose D_{abs} , each cell nucleus of the population was assigned a number of tracks. The next step involves the assigning of a number of damages to each track using the following method:

- i. A type of particle denoted p was sampled from the cumulative histogram of W_p .
- ii. The energy E_p and angle θ_p were sampled from the $f_p(E)$ and $f_p(\theta)$ spectra.
- iii. The track entry point on the surface of the nucleus was sampled.
- iv. The geometric chord length CL_p^{geo} of the track was calculated from the entry point and angle θ_p .
- v. The $F(\widehat{CL_p^{geo}})$, $F(\widehat{CL_p^R})$, $F\left(\frac{nDamage}{\varepsilon}\right)$, and $F_{Damage}(p, E_p, \widehat{CL})$ spectra were reconstructed using the method described earlier. Here $F_{Damage}(p, E_p, \widehat{CL})$ represents the probability spectrum for a track of p type and E_p energy to damage the function of the \widehat{CL} .
- vi. A random number R_{CL} was sampled between 0 and 1, and the sampled chord length CL_p^R was computed such that $F(\widehat{CL_p^R}) = R_{CL}$.
- vii. If $CL_p^{geo} > CL_p^R$, the particle was considered to have stopped in the nucleus; that is, $\varepsilon_p = E_p$ and $CL_p = CL_p^R$. Otherwise, y_p was sampled in the $F(\widehat{y_{CL}})$ spectrum and ε_p was computed such that $\varepsilon_p = y_p \times CL_p^{geo}$ as $CL_p = CL_p^{geo}$.
- viii. Two random numbers, R_{DSB} and R_{SSB} were sampled between 0 and 1. These numbers are referred to as R_{Damage} .
- ix. For each type of damage:

If $R_{Damage} < F_{Damage}(p, E_p, \widehat{CL_p})$, then a value $\frac{nDamage}{\varepsilon}$ is sampled from the $F\left(\frac{nDamage}{\varepsilon}\right)$ spectrum.

From this, $nDamage_p = \frac{nDamage}{\varepsilon} \times \varepsilon_p$ was estimated. Otherwise, $nDamage_p = 0$.

Using this method, once the characteristics of each track were determined, a realistic damage distribution in the cell population was obtained.

3.2. Approximations and assumptions made by the method

The method of damage number and type distribution in the cell population proposed here has several limitations and approximations:

- The cell nucleus geometry is fixed because the database must be simulated for a particular geometry. Currently, only the geometry representing the mean size of an endothelial cell nucleus has been considered; however, other databases can be calculated in the future. However, the variability of the nuclei size within the same population was not considered by the MINAS TIRITH tool. However, the method proposed here is applicable to other cell nuclei geometries.
- The tool decorrelated the energy component to the angular component of the phase space. This parameter is important because it influences the dose deposition. Similarly, the relative positions of certain secondary particles originating from the same primary particle in the phase space were not considered.
- To calculate the number of tracks per cell, the z value was approximated in the geometry by the value $y_{CL} * CL$. Thus, the calculation of z was directly related to the calculated value of CL . For electrons between 1 keV and 16 keV, a significant difference between the geometrically estimated chord length and real chord

length was observed. This was primarily owing to the chaotic trajectory of the electrons at this energy. Thus, a random multiplier uniformly sampled between 0 and 1 was applied to the geometrically calculated chord length. Although this correction method has no physical basis, it allowed the recovery of a more realistic chord length population. In turn, this factor provided a calculated average z similar to the value calculated by the Monte Carlo simulation in the proposed geometry.

- Currently, the number of DSB is not correlated with the number of SSB in a given track. In the proposed method, the number of each type of damage assigned to each track was sampled independent of the spectra specific to each damage type. In addition, evaluating the correlation between the number of each type of damage in the simulation using track structure code is challenging because it is sensitive to geometry and interaction types.

Although these approximations and biases have been identified, it is difficult to quantify their impact on the damage distribution at the cell population scale.

4. Validation of MINAS TIRITH tool

To verify the MINAS TIRITH tool, several tests were conducted at different key stages of the method.

4.1. Evaluation of the reconstruction of spectra by interpolation

First, it is necessary to ensure that the interpolation reconstruction from the spectra calculated using the Geant4 and Geant4-DNA codes is accurate. To assess the reliability of the reconstruction method over the entire energy and particle range of the database, the following method was established for each energy E_l and particle p :

- The $F_{(p,E_l)}(x)$ spectra were removed from the database. Here, x represents the interpolated observable type (γ_{CL} , $nDSB/\varepsilon$ or $nSSB/\varepsilon$).
- The $F_{(p,E_l)}(\hat{x})$ spectra were calculated through interpolation from the other spectra in the database.
- The mean absolute percentage error ($MAPE$) between the $N_{p,E_l}(\hat{x}_n)$ and $p_{p,E_l}(\hat{x}_n)$ was assessed using

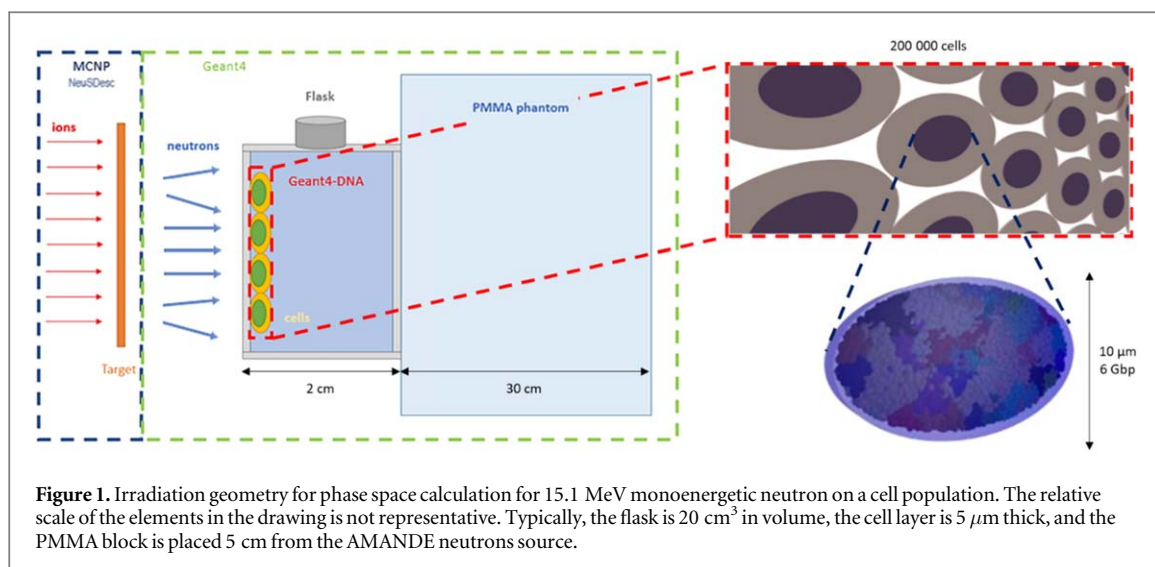
$$MAPE_{p,E_l} = \frac{1}{N} \sum_{n=1}^N \frac{|p_{p,E_l}(\hat{x}_n) - p_{p,E_l}(x_n)|}{p_{p,E_l}(x_n)}.$$

The $MAPE$ indicator was chosen to quantify the reliability of the interpolation reconstruction (De Myttenaere *et al* 2016). An interpolated reconstruction was considered reliable when the $MAPE$ was less than 15%. The fact that this indicator is expressed in terms of percentage assigns the same importance to all points in our spectrum. Moreover, it is important to note that this method overestimates the error. For the calculation of each $MAPE_{p,E_l}$, the worst case was considered because the $F_{(p,E_l)}(x)$ spectrum was removed from the database. During the normal use of MINAS TIRITH, a point (p, E_l) has never been estimated far away from its neighbors during this test.

4.2. Distribution of the specific energy in the cell population

In the proposed method, the distribution of the number of tracks directly depends on the mean value $\overline{z_f}$ of the $f_1(z)$ spectrum. To verify that the MINAS TIRITH allows the correct number of tracks to be determined for a particular absorbed dose, a comparison with a Monte Carlo simulation was performed. To test MINAS TIRITH on all its dynamics in terms of energy and particle type, irradiation of an adherent monolayer endothelial cell population using a monoenergetic neutron beam of 15.1 MeV was performed. This irradiation was performed within the framework of the ModELL project, which aims to develop simulation tools for radiobiology, at the IRSN's AMANDE facility (Gressier *et al* 2004). Indeed, the secondary particle spectrum of this type of irradiation was mostly composed of protons and electrons (from 0.15 MeV) and also alphas (from 0 to 8 MeV), with various angles of incidence depending on the kinematics of the reaction. The phase space of the secondary particles in this irradiation campaign was obtained as follows:

- The geometry of the AMANDE neutron source was created using MCNP (Goorley *et al* 2012), and the neutron creation probability was defined using NeuSDesc software (Birgersson and Lövestam 2009) (figure 1). In the first phase space, primary neutrons were recorded from the MCNP PTRAC file with their energies and directions.
- In the second simulation, using Geant4, these neutrons were transported through a volume corresponding to the dimensions of the culture medium of the irradiated cell population inside a polystyrene flask (figure 1)



placed at the surface of a cubic PMMA phantom with sides measuring 15 cm. Consequently, the positions, directions, and energies of the secondary particles were obtained through this simulation.

iii. In the third step, tracks of the secondary particles were simulated with Geant4-DNA in the culture medium, with a scoring geometry including 200 000 elliptical cylinders representing the cell nuclei of the irradiated population (figure 1). The energy deposited by these particles in the scoring volume was used to compute the $f_1(z)$ spectrum. In addition, the direction, energy, and type of secondary particles penetrating the cell nuclei were recorded as inputs to the MINAS TIRITH tool.

By comparing the $f_1(z)$ spectra acquired through the Monte Carlo simulation and MINAS TIRITH, the track type distributed in MINAS TIRITH can also be interpreted, considering the variety of energy deposition depending on the secondary particle type. Furthermore, comparison of the $zf_1(z)$ spectra acquired by Monte Carlo simulation and MINAS TIRITH provides information regarding the weight of each track. Consequently, the contribution of each particle type in terms of specific energy is obtained.

To provide an idea of the possible outputs of the MINAS TIRITH tool and illustrate the importance of modeling the variability of the dose distribution and the damage production of several cells, an application to a concrete example is presented. The irradiation described above was used at a macroscopic dose of 120 mGy delivered to a population of 2×10^5 cells. To characterize the results obtained, the spectrum $f(z, D)$ of the irradiation and the proportion of cells with a particular amount of damage are presented.

4.3. Evaluation of the number of damages by track sampling method

In the third step, it must be verified that the method for distributing the number of damages for each track yields results that are consistent with those of the simulation chain (Meylan *et al* 2016). Therefore, the mean number of damages per track obtained using the MINAS TIRITH tool was compared with that obtained using the simulation chain developed at IRSN for one cell nucleus. Thus, the Monte Carlo simulations that facilitated the creation of the damage spectrum database were reproduced using the MINAS TIRITH tool. Subsequently, the mean number of damages per track was compared for each simulation at each energy. In addition, this validation step enabled the evaluation of the time-saving factor (CPU 16, Mem 31 Go) allowed by MINAS TIRITH compared to the Monte Carlo simulation, according to the type and initial energy of the simulated particles.

5. Results

5.1. Evaluation of the reconstruction of spectra by interpolation

The interpolation reconstruction errors of the $F_{p,E_l}^{(exit)}(\widehat{\gamma_{CL}})$ and $F_{p,E_l}^{(stop)}(\widehat{\gamma_{CL}})$ spectra shown in figure 2 were evaluated using the MAPE indicator. For electrons, the indicator was less than 1% over the entire E_l range, except for the $F_{e-,3\text{ keV}}^{(exit)}(\widehat{\gamma_{CL}})$, $F_{e-,3\text{ keV}}^{(stop)}(\widehat{\gamma_{CL}})$, and $F_{e-,1.5\text{ MeV}}^{(stop)}(\widehat{\gamma_{CL}})$ spectra, wherein it was approximately 3%. For protons, the indicator was less than 2% over the entire energy range, and in most cases, it was less than 1%. Similarly, for alphas, indicators lower than 1% were obtained throughout the range, except for certain E_l where it reached 3.5%. It is important to remember that the method used to evaluate the quality of the reconstruction

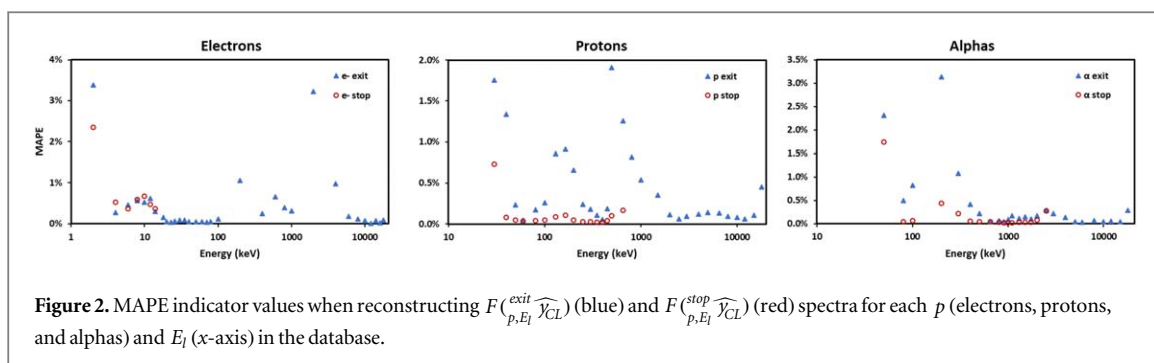


Figure 2. MAPE indicator values when reconstructing $F_{p,E_l}^{(exit)}(\widehat{\gamma}_{CL})$ (blue) and $F_{p,E_l}^{(stop)}(\widehat{\gamma}_{CL})$ (red) spectra for each p (electrons, protons, and alphas) and E_l (x-axis) in the database.

by interpolation of the spectra represents a conservative error value for the entire energy range, as explained in the Materials and Methods section.

For SSBs, for electrons, an MAPE indicator of approximately 5% was observed between 1 and 10 keV, and then lower over the rest of the energy range. For DSBs. Further, an MAPE indicator between 5% and 10% was observed between 1 and 100 keV, and between 10% and 15% for the range of 100–1 MeV (up to 20% for the $F_{(e^-,500\text{ keV})}$ spectrum). For protons, the behaviors of DSBs and SSBs were similar. An MAPE indicator of 1%–6% was obtained over the entire energy range, except for the $F_{(p^+,80\text{ keV})}$ and $F_{(p^+,100\text{ keV})}$ spectra, for which the indicator was in the range of 11%–13%, respectively. Finally, for alphas, the behavior of DSBs and SSBs was also similar, with an MAPE indicator between 2% and 8% over the entire energy range. As highlighted for protons, the $F_{(\alpha,50\text{ keV})}$ and $F_{(\alpha,100\text{ keV})}$ spectra were not in this range, with indicator values of 10% and 12%, respectively.

5.2. Distribution of the specific energy in the cell population

Figure 4 presents microdosimetric spectra resulting from the simulation of the secondary particle transport of the 15.1 MeV mono-energetic neutron irradiation of the endothelial cell population. The $f_1(z)$ and $z f_1(z)$ spectra obtained by Monte Carlo simulation of the secondary particle transport with the Geant4/Geant4-DNA code and those obtained with the MINAS TIRITH tool were compared and found to be consistent. However, certain differences were observed. For instance, an over-representation of tracks with a z between 0.15–0.3 MeV μg^{-1} and between 4–8 MeV μg^{-1} with the MINAS TIRITH tool was observed. Moreover, a very slight under-representation of tracks with a z occurred between 20 and 100 MeV μg^{-1} . Simultaneously, the major difference between the $z f_1(z)$ spectra from the two simulation methods was related to the tracks with a z between 20 and 100 MeV μg^{-1} . The mean value $\overline{z_f}$ obtained by Monte Carlo simulation was $43.83 \pm 0.20\text{ MeV } \mu\text{g}^{-1}$, while that obtained with the MINAS TIRITH tool was $41.00 \pm 0.24\text{ MeV } \mu\text{g}^{-1}$. Thus, there was an underestimation of 6.5% for the $\overline{z_f}$ value when using the MINAS TIRITH tool in the case presented here.

5.3. Evaluation of the mean number of damages by track sampling method

Figure 5 shows a comparison of the mean number of DSBs and SSBs obtained using the MINAS TIRITH tool with those obtained using the Geant4-DNA based Monte Carlo simulation chain. First, a difference of $\pm 20\%$ in the mean number of DSB/track was observed for electrons over most of the energy range. Moreover, the four critical points (20, 65, 500, and 800 keV) differed by approximately $\pm 35\%$. For electrons, the mean number of SSB/track exhibited a difference of $\pm 10\%$ with the Monte Carlo simulation over the 1–50 keV range. For the remaining range, an almost linear increase of up to +42% at 1 MeV was observed. For protons, a difference in the mean DSB/track number was observed between -15% and $+20\%$ over the entire range, with three critical points (3, 12, and 20 MeV) for which the differences were +30, +24, and +38%, respectively. For the mean SSB/track numbers of protons, the deviation was $\pm 10\%$ over the entire energy range. Finally, for alphas, the differences in the mean Damage/track numbers for both DSBs and SSBs, were in the $\pm 10\%$ range. However, a trend was observed for the alphas: the mean number of Damage/track in the peak at approximately 2 MeV was always slightly underestimated by the MINAS TIRITH tool.

5.4. Use of MINAS TIRITH to model a concrete irradiation case

Figure 6 shows the dose variability observed in cells of the same population during neutron irradiation. The specific energies ranged from 10 mGy to 1 Gy in a cell for an average macroscopic absorbed dose of 120 mGy delivered to the cell population. Moreover, although the dose variability is important, the distribution of the dose was close to an asymmetric Gaussian distribution around the macroscopic absorbed dose. However, the distribution of damage in the population was not Poissonian, as it was directly correlated to the specific energy

distribution. Moreover, this correlation does not induce a deterministic relationship because of the variability in DSB induction.

5.5. Time saving by MINAS TIRITH in comparison to MC calculation

Figure 7 shows the time-saving factor for the damage calculation for a single track with the MINAS TIRITH tool compared to the simulation times with the nanodosimetric simulation chain developed at IRSN. The simulations compared here provided the results plotted in figure 5. MINAS TIRITH offers a time-saving factor of approximately 1000 for electrons in most of the energy range. For protons and alphas, this time-saving factor is higher than 10,000 (approximately 100 000 for alphas at 2 MeV) over almost the entire energy range of the tool. Moreover, the highest time-saving factors for each type of particle were in case of the energies where the ionization densities are the highest: approximately 10 keV for electrons, approximately 500 keV for protons, and approximately 2 MeV for alphas.

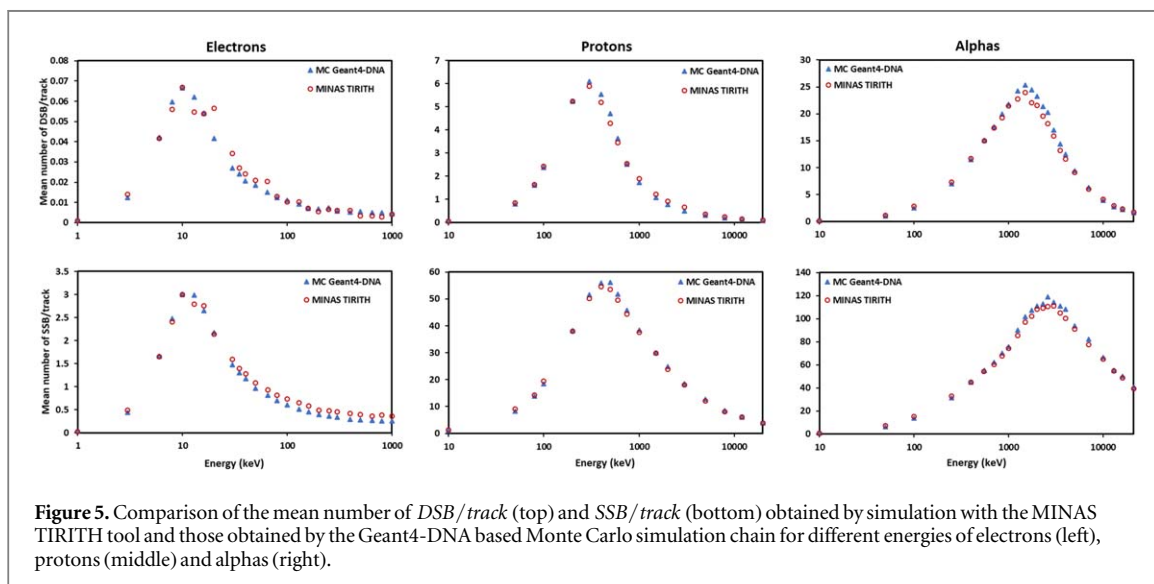
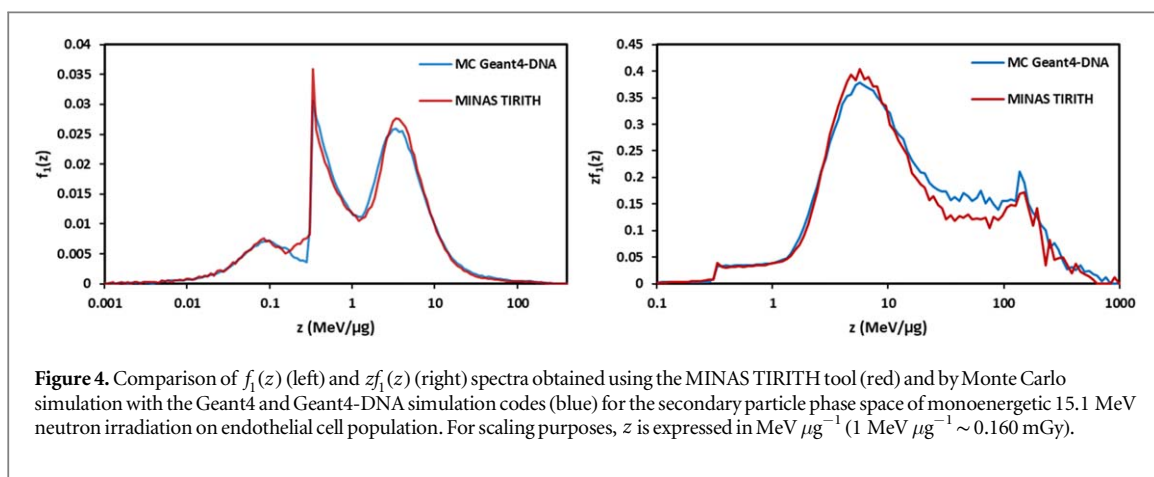
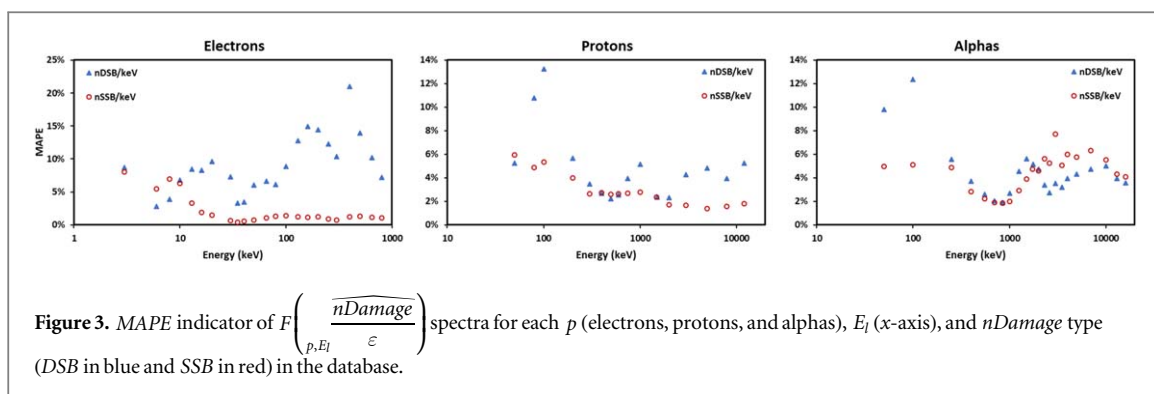
6. Discussion

The MINAS TIRITH tool is a solution for distributing radiation-induced DNA damage in a cell population while respecting the stochastic character of energy deposition. This type of modeling typically requires the use of a Monte Carlo track-structure code coupled with realistic DNA geometries. However, these simulations are expensive in terms of time and computational resources and are not applicable at the scale of a cell population (typically 10^5 cells per part of a flask of endothelial cells). Therefore, it is impossible to verify the results provided by the MINAS TIRITH tool through direct comparisons with those provided by the Monte Carlo methods. Thus, to verify the MINAS TIRITH tool, a step-by-step validation of the method was required.

The MINAS TIRITH tool distributes the tracks in the cell population with respect to the stochastic character of the energy deposition from the reconstructed spectra. Direct validation of the correct distribution of the tracks is possible through comparisons with the Monte Carlo simulation. In figure 4, two characteristic microdosimetric spectra obtained by Monte Carlo simulation and MINAS TIRITH were compared for a phase space of secondary particles set in motion by a 15.1 MeV mono-energetic neutron beam. This phase space was chosen to test the MINAS TIRITH tool for all of its dynamics. First, the \overline{zf} value calculated by MINAS TIRITH exhibited relatively good consistency with that obtained by the Monte Carlo simulation. This implies that the MINAS TIRITH tool distributed several tracks consistent with those calculated by the Monte Carlo simulation. Further, the correspondence of the track type can be appreciated through comparisons of the $f_1(z)$ and $zf_1(z)$ spectra of the two modeling methods. First, a nearly perfect match of the $f_1(z)$ spectra was noticeable, with some discrepancies observed for the z spectra between 0.15–0.3 MeV μg^{-1} and between 4–8 MeV μg^{-1} . Simultaneously, the $zf_1(z)$ spectra were similar for the two methods. It can be concluded from these observations that MINAS TIRITH faithfully reproduced the distribution of tracks in the cell population, respecting the dosimetric characteristics for each particle type.

The MINAS TIRITH tool is based on a spectral sampling approach for the distribution of tracks and the associated damage. The spectra involved in this method were reconstructed via interpolation from a database. The reliability validation of this reconstruction facilitated the validation of this step of the method. Figure 2 shows that the $F(\widehat{\gamma_{CL}})$ spectra were faithfully reconstructed with respect to the values of the MAPE indicator over the entire energy range and for every light-charged particle type. For the $F\left(\frac{n\text{Damage}}{\varepsilon}\right)$ spectra, whose reconstruction error is presented in figure 3, the reconstruction was less faithful but still acceptable if a reference was made to the uncertainties usually found in experimental radiobiology studies (Frankenberg *et al* 1999, Belli *et al* 2000, 2001, Campa *et al* 2005). For the SSBs of electrons, protons, and alphas, the MAPE indicator increased but remained below 10%, which is broadly acceptable for this type of observable. Nevertheless, in case of DSBs, certain points were above this threshold, reaching up to 20% for the $F(e^-, 500 \text{ keV})$ spectrum. Because the statistical power of the Monte Carlo simulations that allowed the construction of these spectra was weaker, DSBs were less frequent events than SSBs. This is illustrated in figure 5, where a factor of 4 to 9 in the occurrence of DSBs compared with SSBs was reached. However, the MAPE indicator value presented here increased and most MAPE for $F\left(\frac{n\text{DSB}}{\varepsilon}\right)$ spectra remained below 10%. Therefore, it can be concluded that the reconstruction of the spectra by interpolation was accurate. Indeed, the MAPE indicator increased but remained below 10%, except for a few critical points, which offers a significant level of confidence in the spectral reconstruction.

Consequently, to fully verify the method, evidence must be provided that the MINAS TIRITH tool is indeed capable of assigning a number for the damages corresponding to its characteristics to each track. With regard to the simulation of radiation-induced DNA damage by the Monte Carlo method, the data that are generally considered for comparisons between simulation codes is the mean number of damages per Gray and Gbp or per



track. Nevertheless, these data do not exhibit a consensus between the main codes, and the experimental data are quite rare and scattered, as illustrated in figure 8. The figure indicates the uncertainty that remains around this value, particularly because it presents experimental data on fibroblasts, with experimental data on endothelial nuclei being even rarer. Therefore, the results obtained with the simulation chain based on Geant4-DNA were used as a reference to verify those obtained by MINAS TIRITH. Figure 5 shows the differences in the mean number of damages per track over the entire dynamic range of the MINAS TIRITH tool. As evident, the two methods exhibited relative consistency (less than 10% error), despite certain critical points (10%–42% error). Considering the uncertainty of the experimental data and the gap in this type of data between the simulation codes based on Monte Carlo methods, it can be concluded that MINAS TIRITH correctly satisfies the

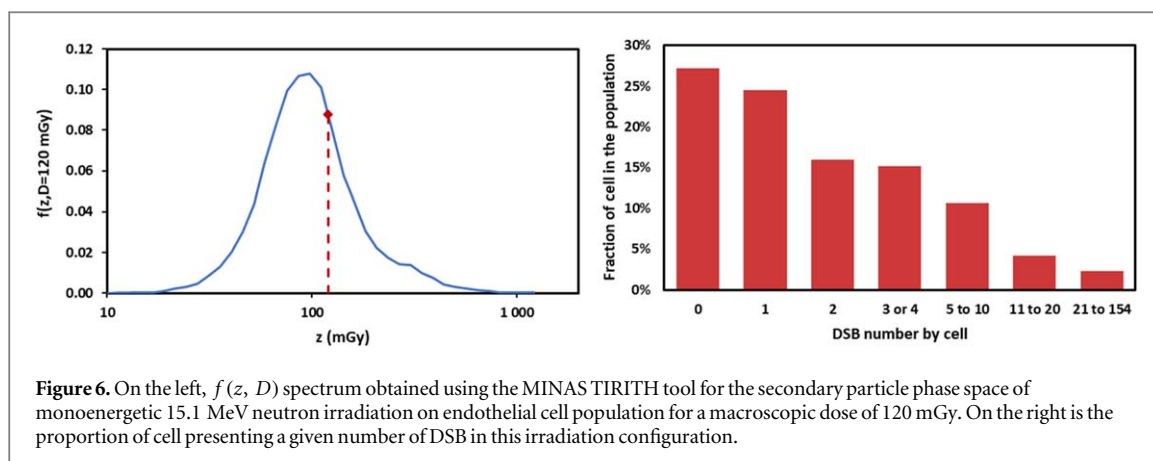


Figure 6. On the left, $f(z, D)$ spectrum obtained using the MINAS TIRITH tool for the secondary particle phase space of monoenergetic 15.1 MeV neutron irradiation on endothelial cell population for a macroscopic dose of 120 mGy. On the right is the proportion of cell presenting a given number of DSB in this irradiation configuration.

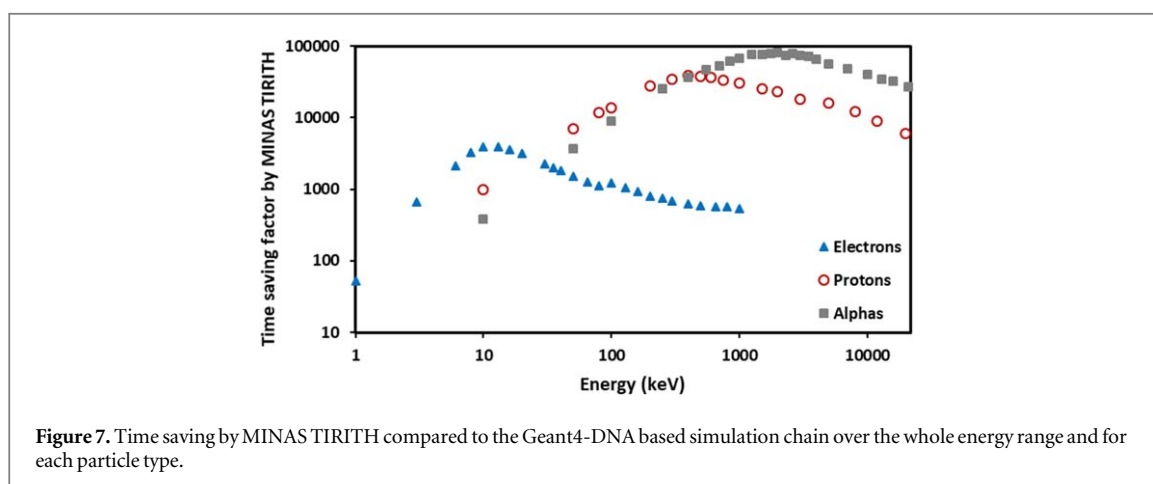


Figure 7. Time saving by MINAS TIRITH compared to the Geant4-DNA based simulation chain over the whole energy range and for each particle type.

requirements specified during its development and provides a damage distribution compatible with that obtained with the Monte Carlo simulation chain method.

Figure 7 highlights the time-saving potential of this new tool for the calculation of radiation-induced DNA damage. Computing time is one of the weaknesses of Monte Carlo simulations of track structure codes, and the time saving observed here provides access to a higher spatial scale in the simulation. Consequently, this renders the modeling of the result of irradiation at the scale of a cell population possible. This new scale opens the door to more direct comparisons between experiments and simulations by considering microdosimetric concepts in the stochastic nature of the dose deposit. Therefore, the needs identified in the literature are met (Chatzipapas *et al* 2020). To further model the fate of an irradiated cell population, classical repair models can be fed with the results from MINAS TIRITH. The characterization of the damage to each cell of the population results in the consideration of mechanistic repair models, applicable to each cell independently, or which take microdosimetric distributions as input (Sato and Furusawa 2012, Manganaro *et al* 2017, Bellinzona *et al* 2021, Cordoni *et al* 2022). Nevertheless, studies on the adaptability of parametric damage repair or cell survival models at the population level are planned for future studies.

Moreover, the biases induced by the method used by MINAS TIRITH can have consequences for dose distribution and damage calculations. For example, the decorrelation of directions and energies can be problematic, particularly for the modeling of irradiations where the Compton effect is predominant. To avoid this, an additional option was developed to allow for the correlation of angles and energies. However, this new option requires a more precise characterization of the phase space to determine this correlation. Furthermore, it is difficult to develop a method that considers the correlation between the number of DSB and SSB sampled within the same track because it is a multifactorial phenomenon (depending on irradiation geometry, interaction type, particle type, and energy) and is dependent on the classification of the damage as DSB or SSB. However, the ratio between the mean number of SSBs and the mean number of DSBs for a given type of track, which constitutes an observable for the validation of the simulation of radiation-induced damage (Nikjoo *et al* 2001), is preserved by the method used by MINAS TIRITH.

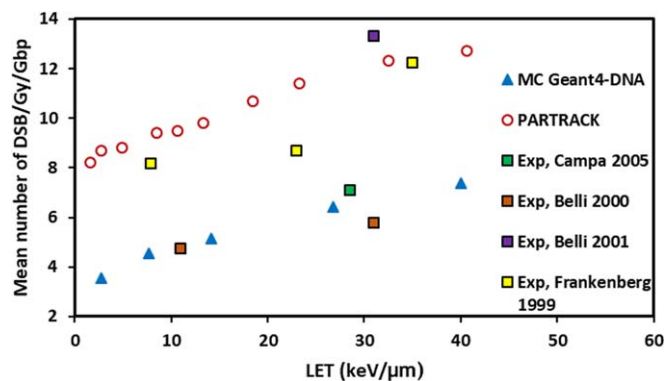


Figure 8. Mean number of DSB/Gy/Gbp obtained by Monte Carlo simulation, with the Geant4-DNA based simulation chain, with the PARTRAC code (Friedland *et al* 2003), and experimentally on human fibroblast nuclei (Frankenberg *et al* 1999, Belli *et al* 2000, 2001, Campa *et al* 2005).

7. Conclusion

This study discussed the development of the MINAS TIRITH tool and investigated its capability to allow the distribution of the tracks and their dosimetric characteristics in an irradiated cell population similar to that obtained by Monte Carlo simulation. Moreover, the MINAS TIRITH tool allowed assignment to each type of track of a number of radiation-induced DNA damages in a very acceptable range compared to that calculated by Monte Carlo simulation with the Geant4-DNA based simulation chain. The simulation times of this new tool were significantly shorter than those of the latter codes. Consequently, MINAS TIRITH paves the way for a more direct comparison between modeling and biological experiments. These comparisons can be performed at the scale of a cell population and are expected to facilitate a better understanding of the induction of radiation-induced damage and the influence of the latter on the fate of an irradiated cell population. Moreover, the results from MINAS TIRITH can feed repair models, in particular in case of the kinetic modeling of foci, to further approach the real experimental conditions. In the future, comparisons with experimental data obtained from cell populations, for which the phase space (mono-energetic neutrons of 15.1 MeV) was simulated, will allow the MINAS TIRITH tool to be validated on a larger scale though comparisons of the number of observed foci to the number of calculated DSBs.

Acknowledgments

This work was partially funded by the French Space Agency (CNES) as part of the ModEll project (Grant DAR No : 2021/4800001139) and by the European Spatial Agency (ESA) as part of the BioRad III project (Grant DAR No : 4000132935/21/NL/CRS).

ORCID iDs

Y Thibaut <https://orcid.org/0000-0002-4520-2728>

M Petit <https://orcid.org/0000-0003-0235-5107>

S Incerti <https://orcid.org/0000-0002-0619-2053>

Y Perrot <https://orcid.org/0000-0001-9043-5857>

References

- Agostinelli S *et al* 2003 Geant4—a simulation toolkit *Nucl. Instrum. Methods Phys. Res. A* **506** 250–303
- Allison J *et al* 2006 Geant4 developments and applications *IEEE Trans. Nucl. Sci.* **53** 270–8
- Allison J *et al* 2016 Recent developments in Geant4 *Nucl. Instrum. Methods Phys. Res. A* **835** 186–235
- Baiocco G *et al* 2016 The origin of neutron biological effectiveness as a function of energy *Sci. Rep.* **6** 34033
- Belli M *et al* 2000 Inactivation of human normal and tumour cells irradiated with low energy protons *Int. J. Rad. Biol.* **76** 831–9
- Belli M *et al* 2001 DNA fragmentation in mammalian cells exposed to various light ions *Adv. Space Res.* **27** 393–9
- Bellinzona V *et al* 2021 Linking microdosimetric measurements to biological effectiveness in ion beam therapy: a review of theoretical aspects of mkm and other models *Front. Phys.* **8** 578492
- Bernal M *et al* 2015 Track structure modeling in liquid water: a review of the Geant4-DNA very low energy extension of the Geant4 monte carlo simulation toolkit *Phys. Med.* **31** 861–74

- Birgersson E and Lövestam G 2009 NeuSDesc—neutron source description software manual. EUR 23794 EN. Luxembourg (luxembourg) OPOCE JRC51437 <https://publications.jrc.ec.europa.eu/repository/handle/JRC51437>
- Campa A et al 2005 DNA DSB induced in human cells by charged particles and gamma rays: experimental results and theoretical approaches *Int. J. Radiat. Biol.* **81** 841–54
- Chatzipapas K P et al 2020 Ionizing radiation and complex dna damage: quantifying the radiobiological damage using monte carlo simulations *Cancers* **12** 799
- Cordoni F et al 2022 Cell survival computation via the generalized stochastic microdosimetric model (GSM2): I. The theoretical framework *Rad. Res.* **197** 218–32
- De Myttenaere A et al 2016 Mean absolute percentage error for regression models *Neurocomputing* **192** 38–48
- Douglass M et al 2015 Development of a radiation track structure clustering algorithm for the prediction of DNA DSB yields and radiation induced cell death in Eukaryotic cells *Phys. Med. Biol.* **60** 3217–36
- Frankenberg D et al 1999 Induction of DNA double-strand breaks by ^1H and ^4He Ions in primary human skin fibroblasts in the LET Range of 8 to 124 keV μm^{-1} . *Rad. Res.* **151** 540–9
- Friedland W et al 2003 Simulation of DNA damage after proton irradiation *Rad. Res.* **159** 401–10
- Gonon G et al 2019 From energy deposition of ionizing radiation to cell damage signaling: benchmarking simulations by measured yields of initial dna damage after ion microbeam irradiation *Rad. Res.* **191** 566–84
- Goorley J T et al 2012 Initial MCNP6 release overview *Nucl. Technol.* **180** 298–315
- Gressier V et al 2004 AMANDE: a new facility for monoenergetic neutron fields production between 2 keV and 20 MeV *Radiat. Prot. Dosim.* **110** 49–52
- Gruel G et al 2016 Cell to cell variability of radiation-induced foci: relation between observed damage and energy deposition *PLoS One* **11** e0145786
- Incerti S et al 2010a The Geant4-DNA project *Int. J. Model. Simul. Sci. Comput.* **1** 157–78
- Incerti S et al 2010b Comparison of Geant4 very low energy cross section models with experimental data in water *Med. Phys.* **37** 4692–708
- Incerti S et al 2018 Geant4-DNA example applications for track structure simulations in liquid water: a report from the Geant4-DNA project *Med. Phys.* **45** e722–39
- Kellerer A M and Chmelevsky D 1975a Concepts of microdosimetry: I. Quantities *Rad. and Environm. Biophys.* **12** 61–9
- Kellerer A M and Chmelevsky D 1975b Concepts of microdosimetry: II. Probability distributions of the microdosimetric variables *Rad. and Environm. Biophys.* **12** 205–16
- Khanna K and Jackson S 2001 DNA double-strand breaks: signaling, repair and the cancer connection *Nat. Genet.* **27** 247–54
- Lampe N et al 2018 Mechanistic DNA damage simulations in Geant4-DNA: I. A parameter study in a simplified geometry *Phys. Med.* **48** 135–45
- Maigne L et al 2021 CPOP: an open source C++ cell POPulation modeler for radiation biology applications *Phys. Med.* **89** 41–50
- Manganaro L et al 2017 A Monte Carlo approach to the microdosimetric kinetic model to account for dose rate time structure effects in ion beam therapy with application in treatment planning simulations *Med. Phys.* **44** 1577–89
- McNamara A et al 2017 Validation of the radiobiology toolkit TOPAS-nBio in simple DNA geometries *Phys. Med.* **33** 207–15
- Meylan S et al 2016 Geant4-DNA simulations using complex DNA geometries generated by the DnaFabric tool *Comput. Phys. Commun.* **204** 159–69
- Meylan S et al 2017 Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA *Sci. Rep.* **7** 11923
- Missiaggia M et al 2021 A novel hybrid microdosimeter for radiation field characterization based on the tissue equivalent proportional counter detector and low gain avalanche detectors tracker: a feasibility study *Front. Phys.* **8** 578444
- Missiaggia M et al 2022 An exploratory study of machine learning techniques applied to therapeutic energies particle tracking in microdosimetry using the novel hybrid detector for microdosimetry (HDM) *Phys. Med. Biol.* **67** 185002
- Nikjoo H et al 2001 Computational approach for determining the spectrum of DNA damage induced by ionizing radiation *Radiat. Res.* **156** 577–83
- Parisi G et al 2022 A systematic study of the contribution of counting statistics to the final lineal energy uncertainty in microdosimetry *Phys. Med. Biol.* **67** 155002
- Rossi H and Zaider M 1996 *Microdosimetry and Its Application* (Berlin: Springer) 1st edn p 321
- Rothkamm K and Lobrich M 2002 Misrepair of radiation-induced DNA double-strand breaks and its relevance for tumorigenesis and cancer treatment *Int. J. Oncol.* **21** 433–40
- Sato T and Furusawa Y 2012 Cell survival fraction estimation based on the probability densities of domain and cell nucleus specific energies using improved microdosimetric kinetic models *Rad. Res.* **178** 341–56
- Tang N et al 2019a Assessment of radio-induced damage in endothelial cells irradiated with 40 kVp, 220 kVp, and 4 MV x-rays by means of micro and nanodosimetric calculations *Int. J. Mol. Sci.* **20** 6204
- Tang N et al 2019b Influence of chromatin compaction on simulated early radiation-induced DNA damage using Geant4-DNA *Med. Phys.* **46** 1501–11
- Thibaut Y et al 2022 Nanodosimetric calculations of radiation-induced DNA damage in a new nucleus geometrical model based on the isochore theory *Int. J. Mol. Sci.* **22** 3770
- Uehara S et al 1993 Cross-sections for water vapour for the Monte Carlo electron track structure code from 10 eV to the MeV region *Phys. Med. Biol.* **38** 1841–58
- Vassiliev O et al 2020 A simple model for calculating relative biological effectiveness of x-rays and gamma radiation in cell survival *Br. J. Radiol.* **93** 20190949