Validation Test of Geant4 Simulation of Electron Backscattering

Sung Hun Kim, Maria Grazia Pia, Tullio Basaglia, Min Cheol Han, Gabriela Hoff, Chan Hyeong Kim, and Paolo Saracco

***Abstract*—Backscattering is a sensitive probe of the accuracy of electron scattering algorithms implemented in Monte Carlo codes. The capability of the Geant4 toolkit to describe realistically the fraction of electrons backscattered from a target volume is extensively and quantitatively evaluated in comparison with experimental data retrieved from the literature. The validation test covers the energy range between approximately 100 eV and 20 MeV, and concerns a wide set of target elements. Multiple and single electron scattering models implemented in Geant4, as well as preassembled selections of physics models distributed within Geant4, are analyzed with statistical methods. The evaluations concern Geant4 versions from 9.1 to 10.1. Significant evolutions are observed over the range of Geant4 versions, not always in the direction of better compatibility with experiment. Goodness-of-fit tests complemented by categorical analysis tests identify a configuration based on Geant4 Urban multiple scattering model in Geant4 version 9.1 and a configuration based on single Coulomb scattering in Geant4 10.0 as the physics options best reproducing experimental data above a few tens of keV. At lower energies only single scattering demonstrates some capability to reproduce data down to a few keV. Recommended preassembled physics configurations appear incapable of describing electron backscattering compatible with experiment. With the support of statistical methods, a correlation is established between the validation of Geant4-based simulation of backscattering and of energy deposition.**

***Index Terms*—Electrons, Geant4, Monte Carlo, simulation.**

I. Introduction

THE simulation of backscattering is a sensitive playground to evaluate the capability of a Monte Carlo transport code to describe electron multiple scattering accurately. Multiple scattering modeling affects not only directly associated observables, such as the fraction of electrons impinging on a target that are backscattered, their energy spectrum and angular distribution, but also the simulated energy deposition in the target volume.

Quantitative evaluations [1], [2] of the capability of Geant4 [3], [4] to reproduce high precision measurements of the energy deposited by low energy electrons in various targets [5], [6] hint at a significant contribution of multiple scattering implementations to determine the accuracy of the simulated energy deposition.

The study documented in this paper analyzes quantitatively the simulation of electron backscattering based on Geant4. It evaluates the performance of several Geant4 physics configuration options, in an extended series of Geant4 versions, with respect to a large sample of experimental data collected from the literature, which cover the energy range from 100 eV to 22 MeV approximately. Compatibility with experiment is established by means of goodness-of-fit statistical methods, while the different ability of Geant4 physics modeling options to reproduce experimental data is quantified by the statistical analysis of categorical data derived from the outcome of goodness-of-fit tests. Finally, the results of this validation process are correlated with the outcome of the validation of electron energy deposition in [1], and the significance of this correlation is quantified.

The scope of the paper is limited to testing the electron backscattering fraction simulated by Geant4. Apart from the considerable amount of material needed to document this subject alone, the results reported in Section VI suggest that validation tests of more complex observables, such as the spectrum and angular distribution of backscattered electrons, would be more meaningful once the Geant4 multiple scattering domain has benefited from the opportunities for improvement highlighted in this paper for future versions of the toolkit.

The results of this validation test provide guidance to optimize the selection and configuration of electron scattering models in experimental application scenarios based on several Geant4 versions examined in this paper. They also provide a quantitative ground for future improvement of the physics modeling and the software development process in Geant4 electromagnetic physics domain.

II. Electron Backscattering

Backscattered electrons play an important role in various experimental applications, such as scanning electron microscopy, electron probe microanalysis, Auger electron spectroscopy, as well as in radiobiology. More in general, backscattering affects the spatial pattern of the energy deposited by electrons in the medium they interact with.

Manuscript received December 12, 2014; revised February 03, 2015; accepted February 04, 2015. Date of publication March 23, 2015; date of current version April 10, 2015. This research was supported in part by the Brazilian Project (161/2012).

S. H. Kim, M. C. Han, and C. H. Kim are with the Department of Nuclear Engineering, Hanyang University, Seoul 133-791, Korea (e-mail: ksh4249@hanyang.ac.kr; mchan@hanyang.ac.kr; chkim@hanyang.ac.kr).

M. G. Pia and P. Saracco are with INFN Sezione di Genova, I-16146 Genova, Italy (e-mail: MariaGrazia.Pia@ge.infn.it; Paolo.Saracco@ge.infn.it).

T. Basaglia is with CERN, CH-1211 Genève 23, Switzerland (e-mail: Tullio.Basaglia@cern.ch).

G. Hoff is with the CAPES Foundation, Ministry of Education of Brazil, Braslia - DF 70040-020, Brazil (e-mail: ghoff.gesic@gmail.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNS.2015.2401055

Electron backscattering has been the subject of experimental and theoretical interest for several decades. Recent papers [7]–[9] attest the attention to this issue in current research. A review on this topic is outside the scope of this paper; the brief overview in this section has the purpose of summarizing information relevant to the simulation validation test documented in this paper.

A. Experimental Data

A large number of experiments have measured various features related to electron backscattering: the fraction of backscattered particles from targets of various thickness and material composition, the angular distribution of backscattered electrons and their energy spectrum, surface effects on solid targets and other effects associated with material properties. These experimental observables provide information to investigate underlying physical effects.

Typical experiment arrangements involve an electron beam of defined energy, high purity targets and a detector to measure the electron current in the backward hemisphere with respect to the target. A hemispherical grid is often used to discriminate electrons with energy above a preset threshold, with the intent of reducing the contamination from secondary electrons in the detected electron sample. The fraction of backscattered electrons is determined as the ratio between the number of electrons reaching the detector and the total number of primary electrons. Experimental uncertainties are associated with the electron beam energy, beam current and current leakage. The detected electron sample can be contaminated by particles scattered from components of the experimental apparatus other than the target, or by secondary electrons. Detailed descriptions of measurement setups can be found in [10] and [11].

The present study collected more than 3000 experimental data points [10]–[66] from the literature, concerning the measurement of the electron backscattering fraction from infinite or semi-infinite targets (i.e., of size exceeding the electron range in the target material, or half of its value). The experimental sample encompasses 48 target elements, which span the range of atomic numbers from beryllium to uranium. Electron energies vary from 78 eV to 22.2 MeV. The analysis reported in the following sections concerns measurements performed with a normally incident beam.

Apparent inconsistencies are visible in the experimental data sample, regarding measurements performed by different experimental groups in similar nominal configurations of target composition and primary electron energy: they hint at the presence of systematic effects. Several references do not report any experimental uncertainties, nor evaluate possible sources of systematic errors affecting the measurements. In some cases experiments performed with similar techniques report largely different uncertainties, which hint to possibly underestimated errors.

The experimental data are shown in the figures of this paper with error bars corresponding to the uncertainties reported in the related references. The apparent absence of error bars associated with some data points reflects either the omission of experimental uncertainties in their published source or values smaller than the size of the data point markers in the plots.

TABLE I
GEANT4 DEFAULT MULTIPLE SCATTERING SETTINGS

Geant4 Version	Multiple Scattering Model
9.1	G4UrbanMscModel
9.2	G4UrbanMscModel2
9.3	G4UrbanMscModel92
9.4	G4UrbanMscModel93
9.6	G4UrbanMscModel95
10.0	G4UrbanMscModel
10.1	G4UrbanMscModel

B. Simulation with General Purpose Monte Carlo Codes

The simulation of electron backscattering has been a subject of interest for general purpose Monte Carlo codes, such as EGS [67], [68], FLUKA [69], Geant4 [3], [4], ITS [70], MCNP [71] and Penelope [72], as it is a sensitive instrument to demonstrate the capabilities of their electron transport models, namely the treatment of multiple and single electron scattering.

The models implemented in these codes are based on a relatively limited set of theoretical approaches [73], developed by Goudsmit and Saunderson [74], [75], Molière [76], Lewis [77], and Wentzel [78], complemented by algorithms specifically tailored to the application of the theory to the Monte Carlo particle transport environment. Some of these algorithms are documented in the literature (e.g., [79], [80]), other are only described in the software documentation of the Monte Carlo systems; nevertheless, limited information is usually available about their theoretical grounds, methods of approximation and implementation details, and the evolution of the software is often the source of inconsistencies between the description of the algorithms in the literature or the software documentation and their actual implementation.

Comparisons of electron backscattering simulations with experimental data are reported in the literature: some examples are [81] and [82], concerning EGS5 [67] and EGSnrc, respectively, [83], concerning MCNP4B [84] and EGS4 [85], [86], concerning FLUKA. The comparisons available in the literature are limited to a single source or a small collection of experimental data: this limitation exposes them to the risk of biased conclusions, if the reference data are affected by unidentified systematic effects, or are not adequately representative of the variety of scenarios encountered in particle transport applications regarding electron energies and target composition. They rest on a qualitative appraisal of the compatibility between simulation and experiment.

To the best of our knowledge, this paper is the first quantitative report, based on rigorous statistical methods, of the comparison of electron backscattering simulation based on a major Monte Carlo system with respect to an extensive collection of experimental data.

III. ELECTRON MULTIPLE SCATTERING SIMULATION IN GEANT4

The electromagnetic physics package of Geant4 encompasses processes and models, which deal with electron single and multiple scattering with the interacting medium. In addition, the physics_lists package of Geant4 collects a set of

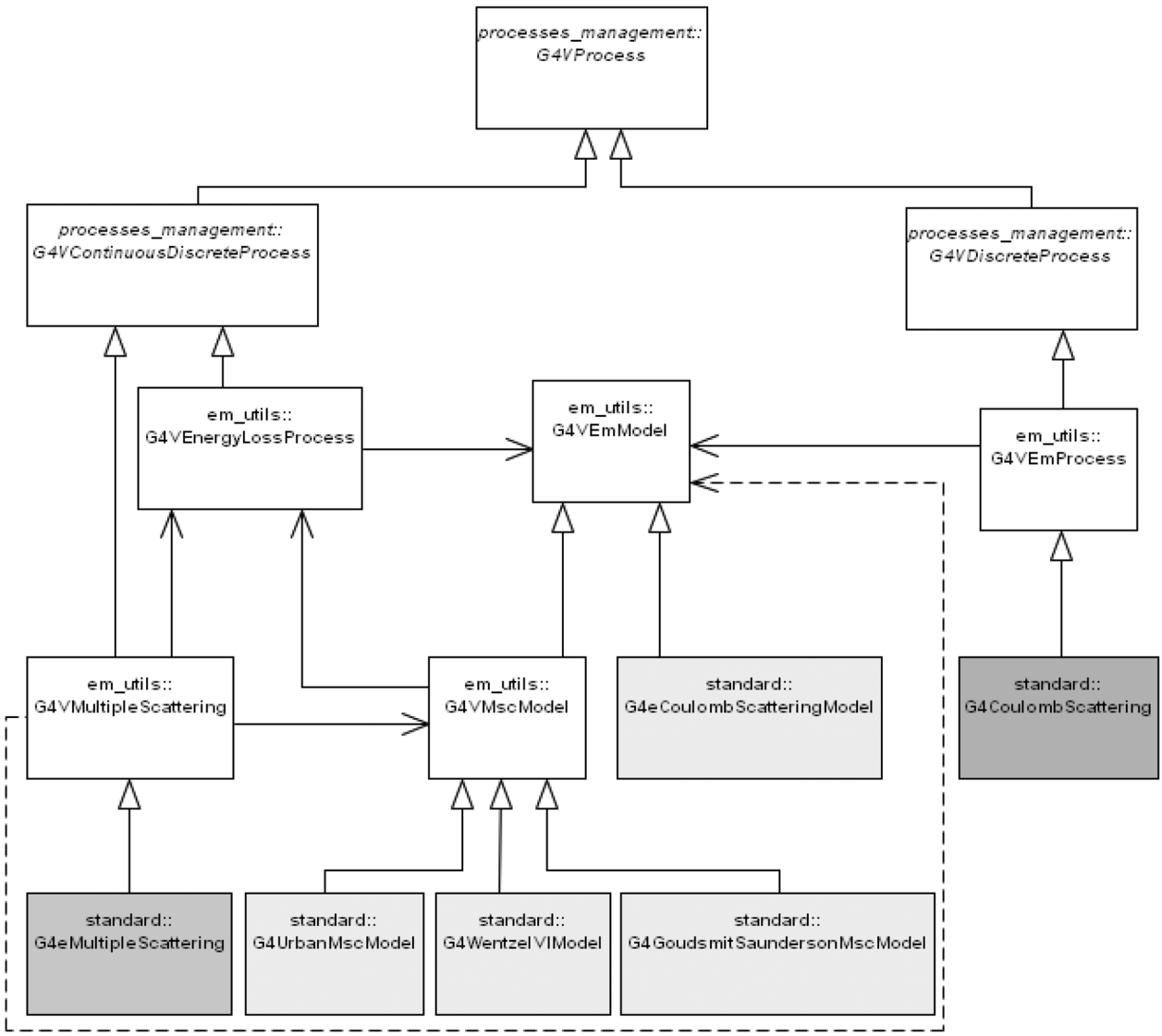

Fig. 1. UML (Unified Modeling Language) class diagram illustrating Geant4 multiple and single scattering processes (dark grey) and models (light grey) involved in this study, and their relationship with Geant4 abstract bases classes (in italic) interacting with the tracking kernel. The diagram is pertinent to Geant4 10.0 and 10.1 versions.

classes, which instantiate predefined selections of physics processes and models for several particle types: they can be directly employed in simulation applications by users not willing to develop physics selections specific to their own experimental scenario.

A. Processes and Models

Geant4 deals with the scattering of electrons with matter by means of two processes, multiple and single Coulomb scattering, which are specializations of the continuous-discrete and discrete processes handled by Geant4 tracking algorithm, respectively. Part of the functionality associated with these processes is delegated to multiple or single scattering models. A process can be configured with one or more associated models, which implement different algorithms and may cooperate to describe electron scattering over different energy ranges. The UML (Unified Modeling Language) class diagram in Fig. 1 illustrates the main features of multiple and single electron scattering in Geant4 10.0 and 10.1.

The Urban multiple scattering model [87]–[89], intended to be based on Lewis' theory [77], is instantiated by default in Geant4 electron multiple scattering process. Several variants of this model have been released in the course of the evolution of Geant4: G4MscModel, G4UrbanMscModel, G4MscModel71, G4UrbanMscModel90, G4UrbanMscModel2, G4UrbanMscModel92, G4UrbanMscModel93, G4UrbanMscModel95 and G4UrbanMscModel96. Table I lists the default model associated with the Geant4 versions considered in this paper. Some relevant parameters characterizing the multiple scattering algorithm are described in [90].

Simulation of single electron scattering is implemented in Geant4 G4CoulombScattering process and G4CoulombScatteringModel model [91]. This model is based on Wentzel calculations [78]; it has been available in Geant4 since version 9.0.

An algorithm combining multiple and single scattering is implemented in the G4WentzelVIModel class [92].

A multiple scattering model based on Goudsmit-Saunderson calculations [74], [75] has been available in Geant4 for the simulation of electron multiple scattering since version 9.3 [93]. It is implemented in the G4GoudsmitSaundersonModel class.

B. Multiple Scattering in Geant4 Prepackaged PhysicsLists

Geant4 software documentation [94] recommends the use of predefined classes in simulation applications, which implement selections of electromagnetic processes and models for several particle types. These classes, derived from G4VPhysicsConstructor, are supplied within the physics_lists package of the Geant4 toolkit; they are used in the prepackaged PhysicsLists hosted in the same package or may be instantiated in user defined PhysicsLists. According to [95], they are "intensively validated".

Geant4 versions 10.0 and 10.1 encompass six generic electromagnetic PhysicsConstructors, which are briefly described below, as well as a few specialized ones (e.g., for optical physics and for the interactions of polarized particles), which are outside the scope of this paper.

The selection of electromagnetic processes and models implemented in the G4EmStandardPhysics PhysicsConstructor is described in the user documentation of Geant4 10.0 [94] as "default electromagnetic physics". From an inspection of the source code it appears that for electrons the Urban multiple scattering model is activated up to 100 MeV, while the WentzelVI model is selected above that energy threshold. Single Coulomb scattering is activated in association with the WentzelVI model. A polar angle threshold is set at 180 degrees. Step limit and range factor settings are kept unchanged with respect to the default implementations in the multiple scattering classes it instantiates.

According to Geant4 user documentation [94], G4EmStandardPhysics_option1 provides "fast but less accurate electron transport" due to the choice of the *Simple* method of step limitation by multiple scattering, reduced step limitation by the ionisation process and enabled *ApplyCuts* option. It uses G4UrbanMscModel for multiple scattering of electrons and positrons. From an inspection of the source code it appears that the model selections, energy and angle thresholds concerning multiple scattering are the same as in G4EmStandardPhysics.

G4EmStandardPhysics_option2 is defined in [94] as "experimental electromagnetic physics with disabled *ApplyCuts* option". From an inspection of the source code it appears that the model selections, energy and angle thresholds concerning electron scattering are the same as in G4EmStandardPhysics.

Geant4 user documentation [94] states that G4EmStandardPhysics_option3 selects electromagnetic physics for simulation with high accuracy due to *UseDistanceToBoundary* multiple scattering step limitation, reduced *finalRange* parameter of stepping function optimized per particle type, alternative model G4KleinNishinaModel for Compton scattering, Rayleigh scattering, and G4IonParameterisedLossModel for ion ionisation. From an inspection of the source code it appears that for electrons the default configuration of the G4eMultipleScattering process is selected, which in turn instantiates the Urban multiple scattering model.

According to Geant4 user documentation [94], the combination of "best electromagnetic models for simulation with high accuracy" includes *UseDistanceToBoundary* multiple scattering step limitation, reduced *finalRange* parameter of stepping function optimized per particle type, low-energy sub-library models G4LivermorePhotoElectricModel, G4LowEPComptonModel below 20 MeV, G4PenelopeGammaConversionModel below 1 GeV, G4PenelopeIonisationModel below 100 keV, and G4IonParameterisedLossModel for ion ionisation. This combination is implemented in G4EmStandardPhysics_option4. From an inspection of the source code it appears that the model selections, energy and angle thresholds concerning multiple scattering are the same as in G4EmStandardPhysics. Some documentation of the performance of G4EmStandardPhysics_option4 with respect to experimental data is available in [96]. Despite our best efforts, we could not retrieve references in the literature supporting quantitatively the statement that G4EmStandardPhysics_option4 corresponds to "best electromagnetic models for simulation with high accuracy".

G4EmLivermorePhysics selects models for electrons and photons based on the EEDL [97] and EPDL [98] data libraries. From an inspection of the source code it appears that model selections, energy and angle thresholds, step limitation and *RangeFactor* settings concerning multiple scattering are the same as in G4EmStandardPhysics_option4.

Two PhysicsConstructors specific to simulations using single scattering and G4WentzelVIModel, named G4EmStandardPhysics_SS and G4EmStandardPhysics_WVI respectively, are released for the first time in Geant4 10.1.

IV. Simulation

A. Simulation Application

A Geant4-based application was developed to simulate the electron backscattering experiments involved in the validation test. The application models the experimental scenarios, encodes a selection of physics configurations corresponding to the options documented in Section III, drives the simulation execution and assembles a set of significant observables produced by the simulation into for further analysis.

The geometry configuration reproduces the relevant features of the experimental setups described in [10]–[66]: a semi-infinite or infinite target of pure elemental composition and a hemispherical sensitive volume representing the electron detector. The target and the detector are surrounded by a very low density material (equivalent to galactic vacuum) to avoid contamination of the test results from spurious interactions. An electron beam impinges on the target. A sketch of the geometry configuration is illustrated in Fig. 2. Relevant parameters, such as the target shape and thickness, the angular acceptance of the detector, the energy and angular spread of the beam are retrieved from the experimental references [10]–[66] whenever available. Electrons entering the sensitive volume are considered detected, when their energy exceeds a preset threshold.

As most references do not report the experimental configuration in detail, some assumptions are made in the simulation, when the corresponding parameters are not explicitly doc-

TABLE II
MULTIPLE AND SINGLE SCATTERING CONFIGURATIONS EVALUATED IN THE ELECTRON BACKSCATTERING VALIDATION TEST

Configuration	Description	Process class	Model class	StepLimitType	RangeFactor
Urban	Urban model, user step limit	G4eMultipleScattering	G4UrbanMscModel	default	default
UrbanB	Urban model, user step limit	G4eMultipleScattering	G4UrbanMscModel	DistanceToBoundary	default
UrbanBRF	Urban model	G4eMultipleScattering	G4UrbanMscModel	DistanceToBoundary	0.01
GSBRF	Goudsmit-Saunderson	G4eMultipleScattering	G4GoudsmitSaundersonModel	DistanceToBoundary	0.01
WentzelBRF	WentzelVI model	G4eMultipleScattering	G4WentzelVIModel	DistanceToBoundary	0.01
WentzelBRFP	WentzelVI, angle limit	G4eMultipleScattering	G4WentzelVIModel	DistanceToBoundary	0.01
Coulomb	Single scattering, default	G4CoulombScattering	G4eCoulombScatteringModel	default	default
CoulombP	Single scattering, with angle limit *(10.1)*	G4CoulombScattering	G4eCoulombScatteringModel	default	default
		PhysicsConstructor class			
EmLivermore		G4EmLivermorePhysics		DistanceToBoundary	0.01
EmStd	Predefined	G4EmStandardPhysics		default	default
EmOpt1	electromagnetic	G4EmStandardPhysics_option1		default	default
EmOpt2	physics	G4EmStandardPhysics_option2		default	default
EmOpt3	selections	G4EmStandardPhysics_option3		DistanceToBoundary	default
EmOpt4		G4EmStandardPhysics_option4		DistanceToBoundary *(10.0)*	0.01 *(10.0)*
				SafetyPlus *(10.1)*	0.02 *(10.1)*
EmSS		G4EmStandardPhysics_SS *(10.1)*		default	default
EmWVI		G4EmStandardPhysics_WVI *(10.1)*		default	default

umented: the target is assumed to be a cylinder of thickness and diameter larger than the incident electron range retrieved from the ESTAR [99] database, the sensitive detector volume is assumed to cover the whole backward hemisphere, a detection threshold of 50 eV is applied, consistent with common experimental practice to exclude secondary electrons [8], the detection efficiency above this threshold is assumed to be 100%, the electron beam is assumed to be monochromatic and incident at 90°. These assumptions correspond to modeling an infinite target, whose geometrical characteristics would not affect the resulting backscattering measurement, and to presuming that the backscattering values reported in the experimental references were previously corrected for geometrical acceptance and detection efficiency as appropriate.

The backscattering fraction is calculated as the ratio between the number of events with at least one detected electron and the number of primary electrons. Detected electrons may be backscattered primary electrons or secondary electrons with energy above the preset threshold. Events with more than one detected electron contributed to the backscattering fraction as a single count.

The physics configurations concerning electron scattering evaluated in this paper are summarized in Table II. The class names reported in this table appear as in Geant4 10.0; different settings in Geant4 10.0 and 10.1 are identified by the respective version number in parentheses. The UML (Unified Modeling Language) class diagram in Fig. 3 illustrates the PhysicsConstructors involved in the application and their relationship with classes in Geant4 kernel responsible for the configuration of physics selections.

B. Simulation Production

Simulations are produced with six Geant4 versions released between December 2007 and December 2014. The selection reflects the study of Geant4 energy deposited by electrons documented in [1], with the addition of Geant4 version 10.0 and 10.1, which were first released after the publication of that paper, and the exclusion of Geant4 version 9.5, which exhibited a very similar behaviour to version 9.6 in [1]. Correction patches to these

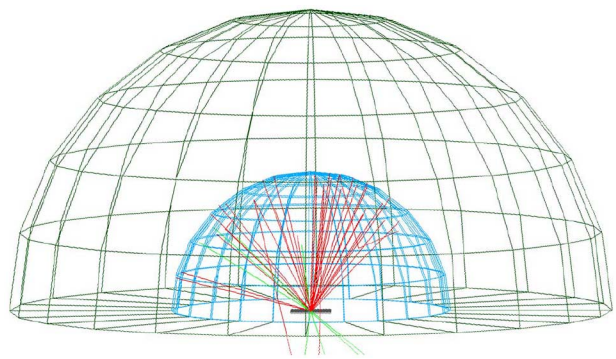

Fig. 2. A sketch of the geometrical setup implemented in the simulation, produced by Geant4 visualization package. Red lines represent backscattered electrons; green lines correspond to photons escaping from the target.

versions released by October 2014 are applied on top of the original versions. For convenience, the Geant4 versions evaluated in this study are identified through their original version number; the corresponding patched versions used in the validation test are listed in Table III.

The simulations were executed on workstations running the Scientific Linux 6 operating system with the gcc version 4.4.7 compiler and on computers running MacOS 10.8.5 (Mountain Lion) operating system with clang 3.3. The same Geant4 versions and application code were installed on all machines.

The simulation configurations corresponding to the experimental test cases were encoded in "macro" files handled by Geant4 user interface. Configurations corresponding to different physics options for the same experimental scenario were automatically generated based on a master configuration file to minimize the possibility of accidental encoding errors, which could be the source of systematic effects in the validation results. All simulation configuration files were kept under version control to ensure the reproducibility of results.

The physics configurations produced in the Geant4 versions subject to evaluation are listed in Table IV. Simulations using the family of Urban configurations and the default single Coulomb scattering configuration were produced in all versions, while those using the WentzelVI and Goudsmit-Saunderson

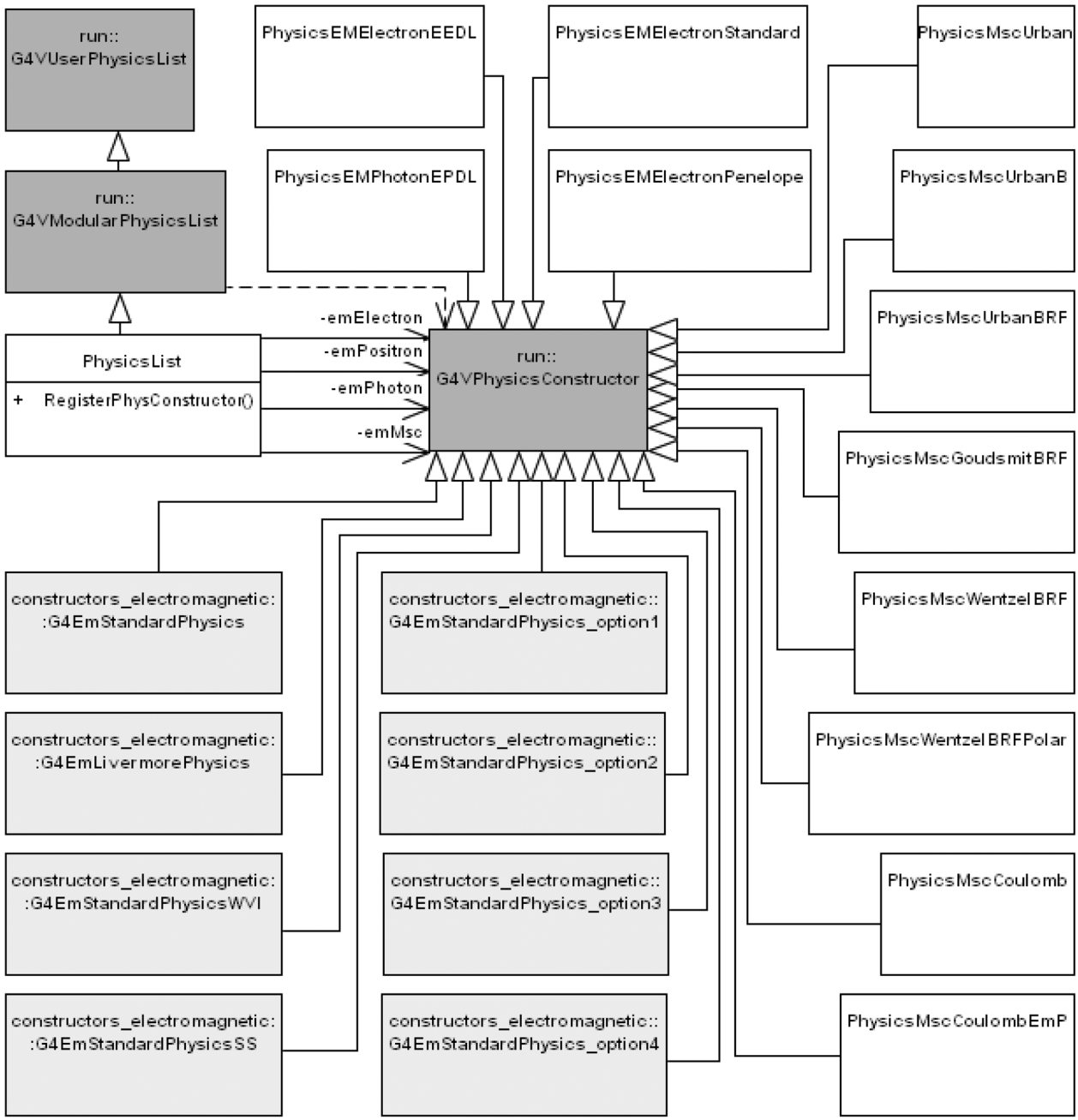

Fig. 3. UML (Unified Modeling Language) class diagram illustrating the physics configuration in the simulation: simulation application classes (white), Geant4 kernel classes (dark grey) and PhysicsConstructor classes released in the Geant4 physics_lists package (light grey).

TABLE III
GEANT4 VERSIONS SUBJECT TO TEST

Version Identifier	Patched Geant4 Version
9.1	9.1p03
9.2	9.2p04
9.3	9.3p02
9.4	9.4p04
9.6	9.6p03
10.0	10.0p03
10.1	not pertinent

models were produced with Geant4 version 9.3 and later, consistent with their first release. Simulations with Coulomb scattering settings different from the default ones (namely with the "θ limit" parameter set to zero) were produced with Geant4 version 10.1.

Simulations using predefined electromagnetic PhysicsConstructors were produced with Geant4 versions 9.6, 10.0 and 10.1, which reflect the recommendations to experimental users issued in the most recent versions of Geant4 at the time of writing this paper. Simulations using the G4EmStandardPhysics_WVI and G4EmStandardPhysics_SS predefined electromagnetic PhysicsConstructors were limited to Geant4 version 10.1, where these classes have been first released.

The secondary production threshold was set to 1 μm (expressed in terms of particle range) in all simulations, with the

TABLE IV
PHYSICS CONFIGURATIONS IN THE SIMULATION PRODUCTION

Versions	Scattering	Electrons	Photons
9.1-10.1	Urban	EEDL, Standard, Penelope	EPDL
	UrbanB	EEDL	
	UrbanBRF	EEDL	
	Coulomb	EEDL	
9.3-10.1	GSBRF	EEDL	EPDL
	WentzelBRF		
	WentzelBRFP		
9.6-10.1	EmLivermore	Embedded	Embedded
	EmStd		
	EmOpt1		
	EmOpt2		
	EmOpt3		
	EmOpt4		
10.1	EmSS	Embedded	Embedded
	EmWVI		
	CoulombP	EEDL	EPDL

exception of test cases with primary electron beam energy above 1 MeV, for which it was raised to 10 μm to reduce the execution time compatible with the available computational resources.

A user-defined step limitation was applied in the simulations corresponding to the Urban and UrbanB configurations; the step limit was set to the same value as the secondary production threshold setting.

For each simulation run, the number of generated events was determined to ensure that the statistical error on the simulated backscattering fraction was smaller or at most comparable to the experimental uncertainty reported in the corresponding reference paper or in similar experimental configurations, when the associated reference did not document uncertainties. Based on these considerations, the number of events in a single run varied between 10000 and 500000, depending on the corresponding experimental configuration.

V. DATA ANALYSIS

The data analysis addresses various issues related to the validation of Geant4-based simulation of electron backscattering: the evaluation of the capability of Geant4 to produce results consistent with experiment, the comparison of the simulation accuracy achievable with different Geant4 multiple and single scattering configurations in the user application, the correlation between the validity of simulated electron backscattering and energy deposition, and the evolution of the results over a series of Geant4 versions. Appropriate statistical methods are applied to analyze the data pertinent to each problem.

A. Analysis Method

The experimental data sample retrieved from the literature is dominated by low energy measurements (mostly below a few tens keV): this intrinsic characteristic of the data set would affect the ability to discern the energy dependence of the capabilities of Geant4 models from an analysis of their compatibility with experiment over the whole data sample, as the the outcome of the statistical tests would be mainly determined by the weight of the low energy component. Therefore, the experimental data sample was partitioned into three energy ranges, identified in the following as “low”, “intermediate” and “high” energy, respectively, to evaluate the capabilities of the examined models as a function of the electron energy.

The grouping of data in three energy intervals is empirically based on the behaviour observed in the plots reporting simulation results and experimental data. To mitigate the risk of biased conclusions due to the empirical nature of its definition, it was verified that the general conclusions of the validation test reported in the following sections are stable with respect to small variations of the definition of low, intermediate and high energies (e.g., 10 keV instead of 20 keV, or 75 keV instead of 100 keV). This verification was done by performing the whole analysis procedure over four different definitions of the three energy groups: although the numerical values of the test statistics were slightly different for different partitioning of the data sample, the global conclusions of the validation tests were insensitive to how the three energy ranges were defined. The results reported in this paper data correspond to grouping the data between 1 and 20 keV, between 20 and 100 keV and above 100 keV. Data below 1 keV are excluded from the analysis, since all the examined models appear incapable of reproducing the experimental measurements below this threshold. The apparent absence of any backscattering occurrence below this energy, concerning some models, hints at an intrinsic limitation of their functionality.

In the context of the data analysis a test case is defined by a set of experimental data, associated with a given target composition, within a given energy range, performed by a given research group. The previously mentioned energy intervals encompass 137, 112 and 57 test cases, respectively.

The statistical analysis of compatibility between simulation and experiment is articulated over two stages.

The first stage of the statistical analysis consists of evaluating the compatibility between simulation and experiment in single test cases. The compatibility between simulated and experimental data in a single test case is determined by two-sample goodness-of-fit tests. The details of this component of the analysis are described in Section V-B.

In the second stage of the analysis categorical statistical tests determine whether the differences in compatibility with experiment observed across the various categories of Geant4 models can be explained only by chance, or should be interpreted as deriving from intrinsic behavioural characteristics.

The approach adopted in the categorical analysis is connected to the physics configuration that characterizes how the data samples were produced, namely whether the data subject to test are unrelated or to some extent related. Simulations using different multiple scattering models produce unrelated samples, as the different modeling options available in Geant4 intend to implement distinct conceptual alternatives in the treatment of multiple scattering. Data samples deriving from simulations that differ only for a secondary option (e.g., a parameter setting), but are based on the same multiple scattering model, are to some extent related: this is the case, for instance, of the three configurations based on the Urban model and the two configurations based on the WentzelVI model listed in Table II. It is also the case of simulations based on single Coulomb scattering and the WentzelVI multiple scattering model, since the latter incorporates the former. Statistical tests pertinent to independent or re-

lated data samples are applied as appropriate in this stage of the analysis.

Finally, a statistical analysis is performed to evaluate whether a correlation can be established between the compatibility with experiment of electron backscattering and energy deposition simulated by Geant4.

The significance level for the rejection of the null hypothesis is set at 0.01 for all tests, unless otherwise specified.

The statistical data analysis reported in the following sections exploits the Statistical Toolkit [100], [101] for goodness-of-fit testing and R [102] for the analysis of contingency tables and of correlation.

B. Evaluation of Individual Test Cases

This stage of the analysis determines whether the experimental and simulated data distributions associated with each test case are consistent with deriving from the same parent distribution.

The generally poor documentation of experimental errors in the experimental references, discussed in Section II-A, prevents the use of the χ^2 test [103], which involves measurement uncertainties explicitly. Therefore goodness-of-fit tests based on the empirical distribution function, where uncertainties do not appear directly in the calculation of the test statistic, were used in the analysis. Four independent goodness-of-fit tests were performed for each test case to mitigate the risk of introducing systematic effects in the validation process due to peculiarities of the mathematical formulation of the test statistic: the Anderson-Darling (AD) [104], [105], Cramer-von Mises (CvM) [106], [107], Kolmogorov-Smirnov (KS) [108], [109] and Watson [110] tests.

The null hypothesis for each goodness-of-fit test is that the simulated and experimental distributions derive from the same parent distribution. The outcome of the test is classified as "fail", if the null hypothesis is rejected, as "pass" otherwise.

For convenience, the "efficiency" of a Geant4 simulation configuration is defined as the fraction of test cases in which a goodness-of-fit test does not reject the null hypothesis at 0.01 level of significance. This variable quantifies the capability of that simulation configuration to produce results statistically consistent with experiment over the whole set of test cases pertinent to one of the energy ranges defined in Section V-A.

Despite some qualitatively visible inconsistencies in the experimental data, no attempt was made to exclude the concerned data sets from the analysis, nor to remove single outliers, since the poor documentation of experimental errors prevents a proper evaluation of the consistency of the various measurements. The possible leftover presence of experimental data affected by systematic errors may artificially lower the efficiency associated with a given Geant4 model: therefore, caution should be applied to interpreting the efficiency values reported in this papers as absolute estimates of Geant4 simulation capabilities. Nevertheless, since all Geant4 physics configurations are exposed to the same test cases, the possible inclusion of experimental data samples exhibiting systematic effects would marginally affect the conclusions regarding the comparison of their compatibility with experiment.

TABLE V
CONFIGURATION OF CONTINGENCY TABLES FOR THE EVALUATION OF UNRELATED SIMULATION CONFIGURATIONS

	Category A	**Category B**
Pass	N_{passA}	N_{passB}
Fail	N_{failA}	N_{failB}

C. Evaluation of Unrelated Data Samples

This component of the analysis evaluates the difference in compatibility with experiment across independent categories of data, e.g. associated with two different multiple scattering models. It receives its input from the results of goodness-of-fit tests described in Section V-B.

The differences in the behavior of the two categories are quantified by means of contingency tables, which are built by counting the number of test cases identified as "fail" or "pass" in Section V-B. Their configuration is illustrated in Table V.

In the analysis of contingency tables the null hypothesis is that there is no relationship between the two categories; in physical terms it means that the two simulation configurations under examination are equivalent regarding the compatibility with experiment of their respective outcome.

A variety of tests is applied to determine the statistical significance of the difference between the two categories of data subject to evaluation: Pearson's χ^2 test [111] (when the number of entries in each cell of the table is sufficiently large to justify its applicability), Fisher's exact test [112] and a variety of unconditional exact tests. The use of different tests mitigates the risk of introducing systematic effects in the validation results due to peculiarities in the mathematical formulation of the test statistic.

Fisher's exact test is widely used in the analysis of contingency tables; although it is based on a model in which both the row and column sums are fixed in advance, which seldom occurs in experimental practice, it remains valid when the row or column totals, or both, are not fixed, but in these cases it tends to be conservative, yielding a larger p-value than the true significance of the test [113].

Unconditional tests, such as Barnard's test [114], Boschloo's test [115] and Suissa and Shuster's [116] calculation of a Z-pooled statistic, are deemed more powerful than Fisher's exact test in some configurations of 2×2 contingency tables [117], [118], but they are computationally more intensive. They yield consistent results in the context of the validation process described in this paper; the p-values of two of such tests are reported among the results of the validation analysis in Section VI.

D. Evaluation of Related Data Samples

This component of the analysis evaluates the difference in compatibility with experiment across related categories of data. It receives its input from the results of goodness-of-fit tests.

This analysis method is applied to two situations: when each subject (e.g., a physics configuration in the simulation) serves in both categories being evaluated (e.g., two Geant4 versions),

TABLE VI
CONFIGURATION OF CONTINGENCY TABLES FOR THE EVALUATION OF RELATED SIMULATION CONFIGURATIONS

	Category A: Pass	**Category A: Fail**
Category B: Pass	$N_{pass\,A,\,pass\,B}$	$N_{fail\,A,\,pass\,B}$
Category B: Fail	$N_{pass\,A,\,fail\,B}$	$N_{fail\,A,\,fail\,B}$

and when one examines two closely related subjects (e.g., simulations differing for a secondary feature of a multiple scattering model, while all the other physics configuration settings are identical).

Appropriate 2×2 contingency tables are built for this purpose, based on the results of the goodness-of-fit tests documented in Section V-B: for each goodness-of-fit test, they report on one diagonal the number of test cases where both categories subject to evaluation "pass" or "fail" the test, and on the other diagonal the number of test cases where one category "passes" the χ^2 test, while the other one"fails". An example of their configuration is shown in Table VI.

McNemar's test [119] is applied to the analysis of related data samples. This test focuses on the significance of the discordant results, i.e., the number of test cases where one category "passes" a goodness-of-fit test and the other one "fails". The null hypothesis for McNemar's test is that the proportion of discordant results is the same in the two cells corresponding to "pass-fail" or "fail-pass" associated with the two categories subject to test.

The calculation of McNemar's test is performed using either the χ^2 asymptotic distribution or the binomial distribution [120]: the latter is also known as "McNemar exact test". Yates' [121] continuity correction may be applied to the calculation of the χ^2 statistic to account for cells with a small number of entries. According to [122], McNemar's test uncorrected for continuity is more powerful than the exact test, and performs well even when the number of discordant pairs is as low as 6, while both the exact test and the corrected McNemar's test are conservative. The results reported in this paper concern McNemar exact test, as most of the analyzed tables involved a small number of entries in some cells, which prevented the calculation of the χ^2 statistic.

VI. RESULTS

The analysis of electron backscattering simulation is focused on the quantification of a few salient features, based on the considerations discussed in the previous sections:

- the quantification of the capability of each physics configuration to simulate electron backscattering compatible with experiment;
- its evolution with Geant4 versions, with emphasis on quantifying the capabilities of the latest versions at the time of writing this paper, to which the interest of experimental users is more generally directed;
- the identification of the physics configurations and Geant4 versions achieving the highest compatibility with experiment;
- the role played by single and multiple scattering in contributing to the accuracy and the computational speed of the simulation.

The following sections document first an overview of the main features of the statistical results, followed by detailed discussion of each Geant4 model configuration subject to evaluation.

The large number of experimental test cases, Geant4 physics modeling options and versions evaluated in the validation tests makes a full graphical illustration of the results impractical within the scope of a journal publication. A sample of plots, which span a wide range of electron beam energies, target atomic numbers, Geant4 models and versions, complement the outcome of the statistical analysis with a qualitative illustration of the results.

A. Model Compatibility with Experiment

The efficiencies resulting from goodness-of-fit tests are summarized in Table VII for all the physics configurations and Geant4 versions evaluated in this paper. One can observe that the four tests produce consistent outcomes regarding the rejection of the null hypothesis of equivalence between simulated and experimental distributions. Therefore, for convenience, the categorical analyses reported in the following sections are performed only based on the outcome of the Anderson-Darling test, unless otherwise specified.

The results listed in Table VII show that Geant4 multiple scattering models fail to reproduce experimental electron backscattering data in the lower energy end, while the single scattering approach adopted in the Coulomb and WentzelBRF configurations retains the capability to describe lower energy data, although with reduced efficiency with respect to the higher energy range. At energies above 100 keV the highest efficiency is achieved with the Urban configuration in Geant4 9.1 in the context of a condensed path scheme, and with the Coulomb scattering model in Geant4 10.0 in the context of simulating individual scattering occurrences.

Evaluations specific to each Geant4 model, or family of models, are detailed in the following sections.

B. Effect of Different Geant4 Electron Physics Models

The fraction of backscattered electrons produced in the simulation is mainly determined by multiple scattering; a small contribution to the number of electrons reaching the detector consists of secondary particles produced in the ionization of target atoms, which escape the target volume.

As Geant4 encompasses different models of electron ionization, their effect on the validation of the simulation of electron backscattering has been evaluated. For this purpose simulations where executed with the Urban multiple scattering configuration associated with different Geant4 electron physics models: those based on the EEDL [97] evaluated data library [123]–[125], those encompassed in the Geant4 Standard electromagnetic package [126] and those reengineered in Geant4 from the Penelope [72] Monte Carlo code.

The efficiency obtained with the three different configurations is listed in Table VIII for the Geant4 versions subject to evaluation: it derives from the results of the Anderson-Darling goodness-of-fit test comparing simulated and experimental backscattered electron distributions. One can observe that the results associated with different electron physics models are

TABLE VII
EFFICIENCY OF PHYSICS CONFIGURATIONS FOR GEANT4 VERSIONS 9.1 TO 10.1

Configuration	Geant4	1-20 keV				20-100 keV				$> 100keV$			
	version	AD	CvM	KS	Watson	AD	CvM	KS	Watson	AD	CvM	KS	Watson
Urban	9.1	< 0.01	< 0.01	< 0.01	< 0.01	0.10±0.03	0.10±0.03	0.10±0.03	0.10±0.03	0.79±0.05	0.79±0.05	0.77±0.05	0.79±0.05
Urban	9.2	< 0.01	< 0.01	< 0.01	< 0.01	0.03±0.02	0.03±0.02	0.03±0.02	0.03±0.02	0.79±0.05	0.77±0.05	0.82±0.05	0.88±0.04
Urban	9.3	< 0.01	< 0.01	< 0.01	< 0.01	0.09±0.03	0.09±0.03	0.09±0.03	0.09±0.03	0.74±0.06	0.74±0.06	0.72±0.06	0.74±0.06
Urban	9.4	< 0.01	< 0.01	< 0.01	< 0.01	0.10±0.03	0.10±0.03	0.10±0.03	0.10±0.03	0.56±0.06	0.56±0.06	0.56±0.06	0.61±0.06
Urban	9.6	< 0.01	< 0.01	< 0.01	< 0.01	0.17±0.04	0.17±0.04	0.17±0.04	0.17±0.04	0.68±0.06	0.68±0.06	0.68±0.06	0.82±0.05
Urban	10.0	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.11±0.04	0.11±0.04	0.11±0.04	0.09±0.04
Urban	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.09±0.04
UrbanB	9.1	< 0.01	< 0.01	< 0.01	< 0.01	0.10±0.02	0.10±0.02	0.10±0.02	0.10±0.02	0.79±0.05	0.79±0.05	0.77±0.05	0.79±0.05
UrbanB	9.2	< 0.01	< 0.01	< 0.01	< 0.01	0.03±0.02	0.03±0.02	0.03±0.02	0.03±0.02	0.79±0.05	0.77±0.05	0.82±0.05	0.88±0.04
UrbanB	9.3	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.11±0.04	0.11±0.04	0.11±0.04	0.11±0.04
UrbanB	9.4	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
UrbanB	9.6	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.05±0.03	0.05±0.03	0.05±0.03	0.07±0.04
UrbanB	10.0	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
UrbanB	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.09±0.04	0.07±0.04	0.09±0.04	0.11±0.04
UrbanBRF	9.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.05±0.03	0.05±0.03	0.05±0.03	0.07±0.04
UrbanBRF	9.2	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.02±0.02	0.02±0.02	0.02±0.02	0.04±0.03
UrbanBRF	9.3	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
UrbanBRF	9.4	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
UrbanBRF	9.6	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
UrbanBRF	10.0	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
UrbanBRF	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.09±0.04	0.09±0.04	0.09±0.04	0.09±0.04
Coulomb	9.1	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.16±0.03	0.16±0.03	0.17±0.04	0.18±0.04	0.60±0.06	0.60±0.06	0.60±0.06	0.68±0.06
Coulomb	9.2	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.04±0.02	0.04±0.02	0.04±0.02	0.04±0.02	0.61±0.06	0.61±0.06	0.65±0.06	0.65±0.06
Coulomb	9.3	0.19±0.03	0.18±0.03	0.18±0.03	0.20±0.03	0.08±0.03	0.08±0.03	0.08±0.03	0.09±0.03	0.63±0.06	0.63±0.06	0.63±0.06	0.67±0.06
Coulomb	9.4	0.03±0.02	0.03±0.02	0.04±0.02	0.04±0.02	0.22±0.04	0.21±0.04	0.21±0.04	0.21±0.04	0.68±0.06	0.68±0.06	0.72±0.06	0.74±0.06
Coulomb	9.6	0.48±0.04	0.48±0.04	0.48±0.04	0.46±0.04	0.42±0.05	0.42±0.05	0.42±0.05	0.41±0.05	0.79±0.05	0.77±0.05	0.72±0.06	0.81±0.05
Coulomb	10.0	0.49±0.04	0.49±0.04	0.48±0.04	0.48±0.04	0.40±0.05	0.39±0.05	0.39±0.05	0.38±0.05	0.79±0.05	0.77±0.05	0.79±0.05	0.89±0.04
Coulomb	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.02	< 0.02	< 0.02	< 0.02
CoulombP	10.1	0.47±0.04	0.47±0.04	0.47±0.04	0.46±0.04	0.45±0.05	0.44±0.05	0.44±0.05	0.43±0.05	0.81±0.05	0.81±0.05	0.77±0.05	0.88±0.04
GSBRF	9.3	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.07±0.03	0.07±0.03	0.07±0.03	0.07±0.03
GSBRF	9.4	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.07±0.03	0.07±0.03	0.07±0.03	0.07±0.03
GSBRF	9.6	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.58±0.06	0.58±0.06	0.60±0.06	0.58±0.06
GSBRF	10.0	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.58±0.06	0.56±0.06	0.58±0.06	0.54±0.06
GSBRF	10.1	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.39±0.06	0.40±0.06	0.39±0.06	0.39±0.06
WentzelBRF	9.3	0.18±0.03	0.18±0.03	0.19±0.03	0.20±0.03	0.09±0.03	0.09±0.03	0.09±0.03	0.10±0.03	0.60±0.06	0.60±0.06	0.60±0.06	0.65±0.06
WentzelBRF	9.4	0.02±0.01	0.02±0.01	0.02±0.01	0.04±0.02	0.21±0.04	0.20±0.04	0.20±0.04	0.20±0.04	0.61±0.06	0.60±0.06	0.60±0.06	0.68±0.06
WentzelBRF	9.6	0.46±0.04	0.46±0.04	0.46±0.04	0.46±0.04	0.44±0.05	0.43±0.05	0.43±0.05	0.42±0.05	0.79±0.05	0.77±0.05	0.75±0.06	0.81±0.05
WentzelBRF	10.0	0.49±0.04	0.48±0.04	0.48±0.04	0.49±0.04	0.44±0.05	0.43±0.05	0.43±0.05	0.42±0.05	0.81±0.05	0.79±0.05	0.82±0.05	0.88±0.04
WentzelBRF	10.1	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.42±0.06	0.42±0.06	0.42±0.06	0.46±0.06
WentzelBRFP	9.3	0.02±0.01	0.02±0.01	0.02±0.01	0.03±0.02	0.04±0.02	0.04±0.02	0.04±0.02	0.04±0.02	0.33±0.06	0.33±0.06	0.33±0.06	0.37±0.06
WentzelBRFP	9.4	0.02±0.01	0.02±0.01	0.02±0.01	0.02±0.01	0.03±0.02	0.03±0.02	0.03±0.02	0.03±0.02	0.21±0.05	0.21±0.05	0.21±0.05	0.21±0.05
WentzelBRFP	9.6	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.30±0.05	0.28±0.05	0.30±0.05	0.33±0.05
WentzelBRFP	10.0	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.44±0.06	0.42±0.06	0.44±0.06	0.46±0.06
WentzelBRFP	10.1	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.42±0.06	0.42±0.06	0.42±0.06	0.46±0.06
EmLivermore	9.6	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.05±0.03	0.07±0.04	0.07±0.04
EmLivermore	10.0	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.05±0.03	0.05±0.03	0.05±0.03	0.05±0.03
EmLivermore	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.09±0.04
EmStd	9.6	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.40±0.06	0.40±0.06	0.42±0.06	0.44±0.06
EmStd	10.0	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.09±0.04
EmStd	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.05±0.03	0.05±0.03	0.05±0.03	0.05±0.03
EmOpt1	9.6	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.33±0.06	0.33±0.06	0.33±0.06	0.35±0.06
EmOpt1	10.0	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.39±0.06	0.39±0.06	0.39±0.06	0.40±0.06
EmOpt1	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.14±0.05	0.12±0.05	0.14±0.05	0.18±0.05
EmOpt2	9.6	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.32±0.06	0.32±0.06	0.32±0.06	0.35±0.06
EmOpt2	10.0	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.37±0.06	0.37±0.06	0.37±0.06	0.40±0.06
EmOpt2	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.16±0.05	0.16±0.05	0.16±0.05	0.19±0.05
EmOpt3	9.6	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
EmOpt3	10.0	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
EmOpt3	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
EmOpt4	9.6	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.07±0.04	0.07±0.04	0.07±0.04	0.07±0.04
EmOpt4	10.0	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	0.09±0.04	0.09±0.04	0.09±0.04	0.09±0.04
EmOpt4	10.1	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01	< 0.02	< 0.02	< 0.02	< 0.02
EmWVI	10.1	< 0.01	< 0.01	< 0.01	< 0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.01±0.01	0.40±0.06	0.40±0.06	0.44±0.06	0.46±0.06

very similar; a categorical analysis comparing the results of the Standard and Penelope models with those based on EEDL fails to reject the hypothesis of equivalent behaviour with 0.05 significance. Both Fisher's and Barnard's exact tests lead to the same conclusion.

From this analysis one can infer that the electron backscattering test studied in this paper is insensitive to how electron interactions, apart from multiple scattering, are modeled in Geant4.

The results reported in the following sections are obtained using Geant4 electron interaction models based on EEDL.

The interactions of secondary photons were modeled based on the EPDL [98] evaluated data library in all simulations, based on validation studies [127]–[130]: they play a negligible role in the validation test, as they could only affect the estimated electron backscattering fraction as a contamination originating from secondary electrons they produce. No such contamination was observed in the simulated data sample subject to analysis.

TABLE VIII
EFFICIENCY FOR THE URBAN MULTIPLE SCATTERING CONFIGURATION WITH DIFFERENT ELECTRON PHYSICS MODELS

Energy (keV)	Version	EEDL	Standard	Penelope
1-20	9.1	< 0.01	< 0.01	< 0.01
	9.2	< 0.01	< 0.01	< 0.01
	9.3	< 0.01	< 0.01	< 0.01
	9.4	< 0.01	< 0.01	< 0.01
	9.6	< 0.01	< 0.01	< 0.01
	10.0	< 0.01	< 0.01	< 0.01
	10.1	< 0.01	< 0.01	< 0.01
20-100	9.1	0.10 ± 0.03	0.11 ± 0.03	0.07 ± 0.02
	9.2	0.03 ± 0.02	0.06 ± 0.02	0.03 ± 0.02
	9.3	0.09 ± 0.02	0.04 ± 0.02	0.07 ± 0.02
	9.4	0.10 ± 0.03	0.08 ± 0.03	0.09 ± 0.03
	9.6	0.17 ± 0.04	0.15 ± 0.03	0.12 ± 0.03
	10.0	< 0.01	< 0.01	< 0.01
	10.1	< 0.01	< 0.01	< 0.01
>100	9.1	0.79 ± 0.05	0.75 ± 0.06	0.77 ± 0.05
	9.2	0.79 ± 0.05	0.81 ± 0.05	0.74 ± 0.06
	9.3	0.74 ± 0.06	0.65 ± 0.06	0.67 ± 0.06
	9.4	0.58 ± 0.06	0.61 ± 0.06	0.61 ± 0.06
	9.6	0.68 ± 0.06	0.72 ± 0.06	0.68 ± 0.06
	10.0	0.11 ± 0.04	0.09 ± 0.04	0.09 ± 0.04
	10.1	0.07 ± 0.04	0.07 ± 0.04	0.09 ± 0.04

C. *Geant4 Urban Models*

The fraction of backscattered electrons produced with the Urban, UrbanB and UrbanBRF configurations is illustrated in Figs. 4–17 for Geant4 versions 9.1 to 10.0. Plots concerning Geant4 10.1 are omitted, as they look very similar to those produced with version 10.0. In these figures and the following ones the labels associated with experimental data encode the name of the first author and the publication year of the reference they derive from, thus allowing the tracebility of the experimental data points appearing in the plots. In all figures, experimental errors are plotted whenever they are documented in the associated papers; in some cases they may be smaller than the marker size. The statistical uncertainties of simulated data are not drawn for better clarity of the plots; in most cases they are smaller than the marker size.

As a qualitative appraisal, one can observe that none of the three configuration variants is capable of describing electron backscattering accurately below a few tens of keV. The simulations adopting the Urban and UrbanB multiple scattering configurations exhibit different characteristics of compatibility with experimental data along the evolution of Geant4 from version 9.1 to 10.1, while simulations adopting the UrbanBRF options appear to behave consistently over all the examined versions.

For energies above a few tens of keV the evolution of the compatibility with experiment of simulations using the Urban configuration seems to follow a similar pattern to that observed in the validation of Geant4 simulation of deposited energy in [1] over Geant4 versions 9.1 to 9.6. Visibly different results are obtained with Geant4 version 10.0, which manifestly worsens the capability of the simulation to reproduce experimental data. This behaviour persists in Geant4 10.1. A qualitative illustration of the evolution of simulations using the Urban configuration is summarized in Figs. 18 and 19. These empirical observations are quantitatively supported by the results of goodness-of-fit tests in Table VII. The highest compatibility with experiment is achieved with the Urban multiple scattering configuration in the earlier Geant4 versions among those evaluated in this paper.

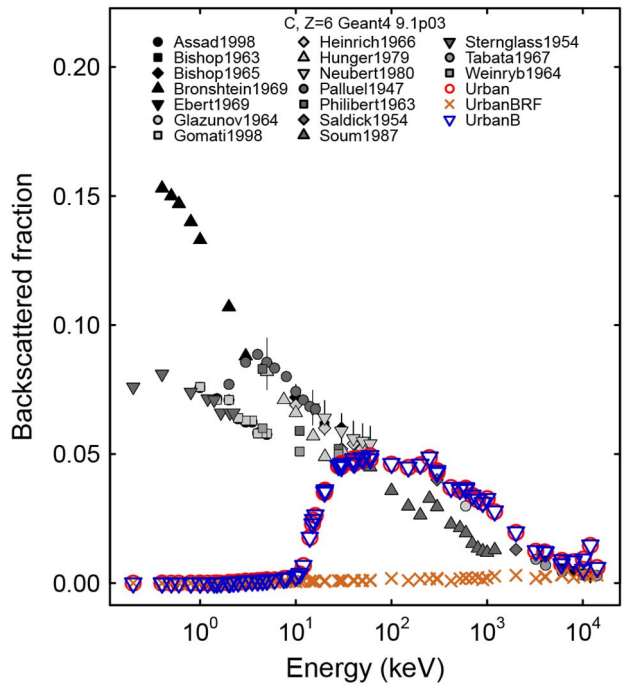

Fig. 4. Fraction of electrons backscattered from a carbon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.1 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

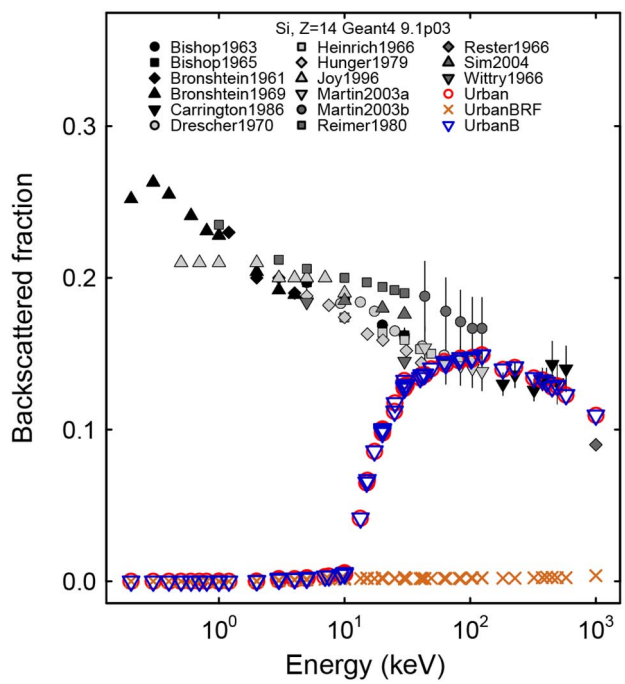

Fig. 5. Fraction of electrons backscattered from a silicon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.1 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

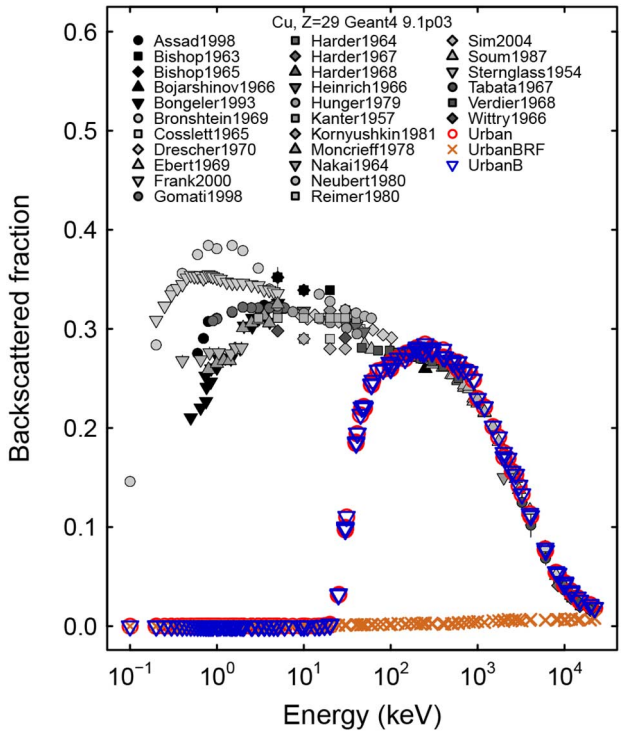

Fig. 6. Fraction of electrons backscattered from a copper target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.1 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

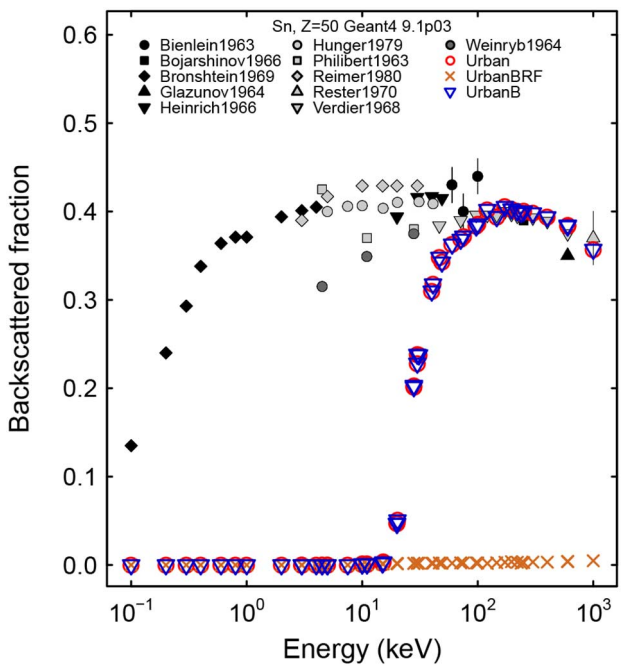

Fig. 7. Fraction of electrons backscattered from a tin target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.1 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

The evolution of the capability of the Urban multiple scattering configuration to produce backscattering simulations compatible with experiment is quantified in Table IX: the second column reports the p-values of McNemar's test

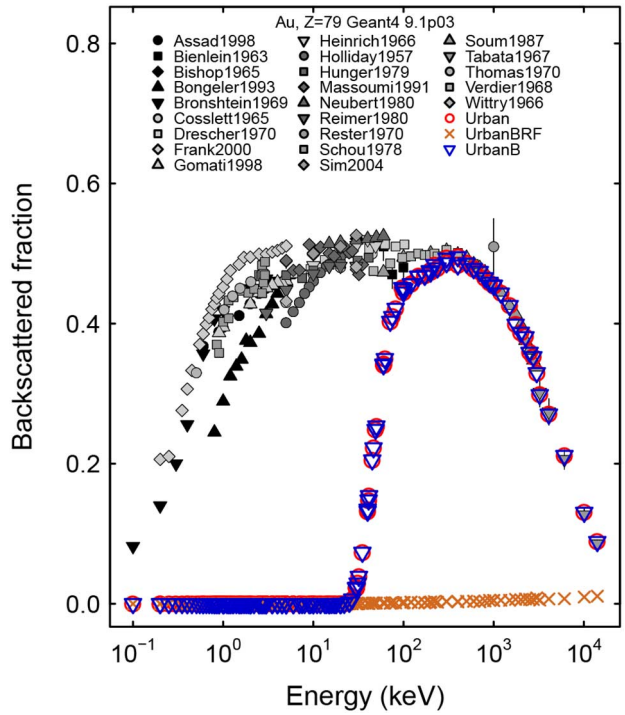

Fig. 8. Fraction of electrons backscattered from a gold target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.1 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

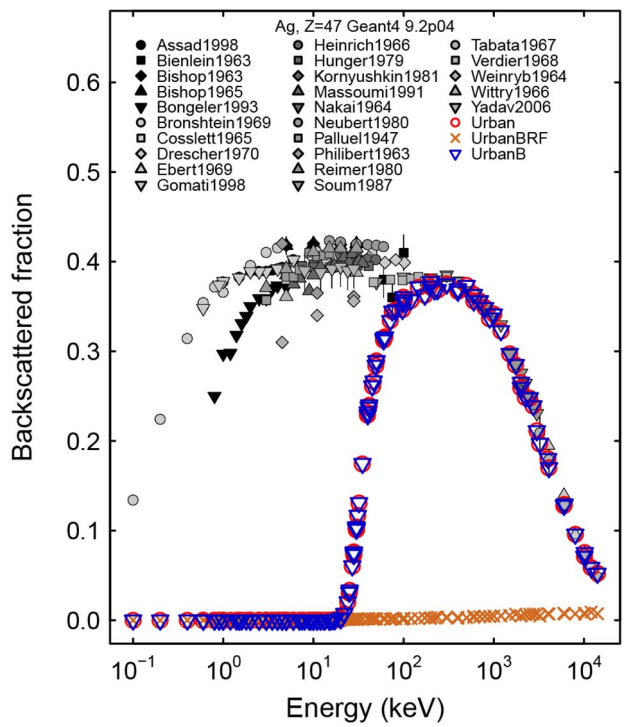

Fig. 9. Fraction of electrons backscattered from a silver target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.2 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

comparing the compatibility with experiment of the Urban configuration in Geant4 9.1, which according to Table VII produces the highest efficiency, with later versions. The hypoth-

TABLE IX
P-VALUES OF MCNEMAR EXACT TEST COMPARING THE COMPATIBILITY WITH EXPERIMENT OF THE URBAN CONFIGURATION IN GEANT4 9.1 WITH LATER GEANT4 VERSIONS AND WITH VARIANTS OF THE URBAN CONFIGURATION

Geant4 Version	Configuration		
	Urban	UrbanB	UrbanBRF
9.1		1.000	<0.001
9.2	1.000	1.000	<0.001
9.3	0.453	<0.001	<0.001
9.4	0.002	<0.001	<0.001
9.6	0.070	<0.001	<0.001
10.0	<0.001	<0.001	<0.001
10.1	<0.001	<0.001	<0.001

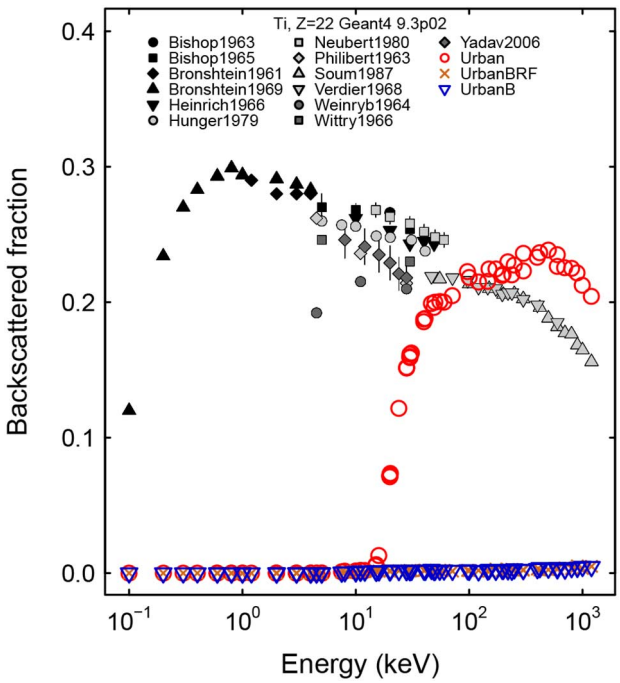

Fig. 10. Fraction of electrons backscattered from a titanium target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.3 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

esis of equivalent compatibility with experiment is rejected with 0.01 significance for the Urban configuration in Geant4 versions 9.4, 10.0 and 10.1. These results concern the energy range above 100 keV; the statistical comparison of lower energy results is less relevant due to the reduced ability of the Urban multiple scattering configuration to reproduce experimental measurements below a few tens of keV.

The backscattering fraction resulting from the UrbanBRF options appears to be very small in all test cases and differs significantly from experimental data. In depth inspection of the results shows that the fraction of detected electrons mainly consists of secondary particles produced by ionization of target atoms. These qualitative observations are confirmed by the results of goodness-of-fit tests listed in Table VII: the number of test cases that fail to reject the hypothesis of compatibility between experimental and simulated backscattering fractions is very small for the UrbanBRF multiple scattering configuration at all energy intervals.

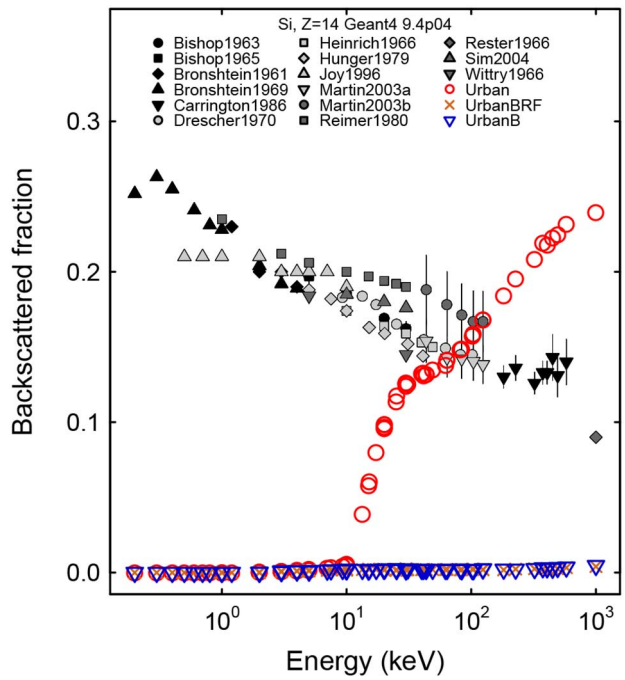

Fig. 11. Fraction of electrons backscattered from a silicon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.4 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

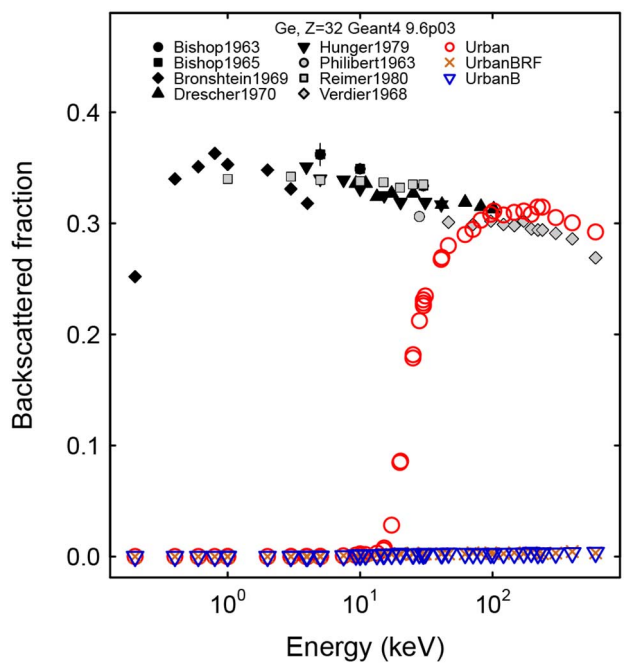

Fig. 12. Fraction of electrons backscattered from a germanium target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.6 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

The UrbanB configuration produces identical results to the Urban one in Geant4 versions 9.1 and 9.2, while it approaches the behaviour of the UrbanBRF configuration in later Geant4 versions. Both configurations apply the *DistanceToBoundary*

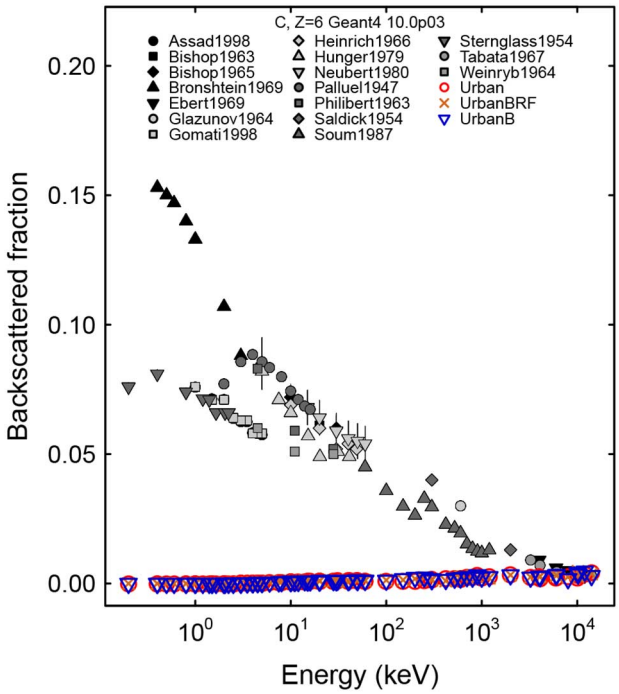

Fig. 13. Fraction of electrons backscattered from a carbon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

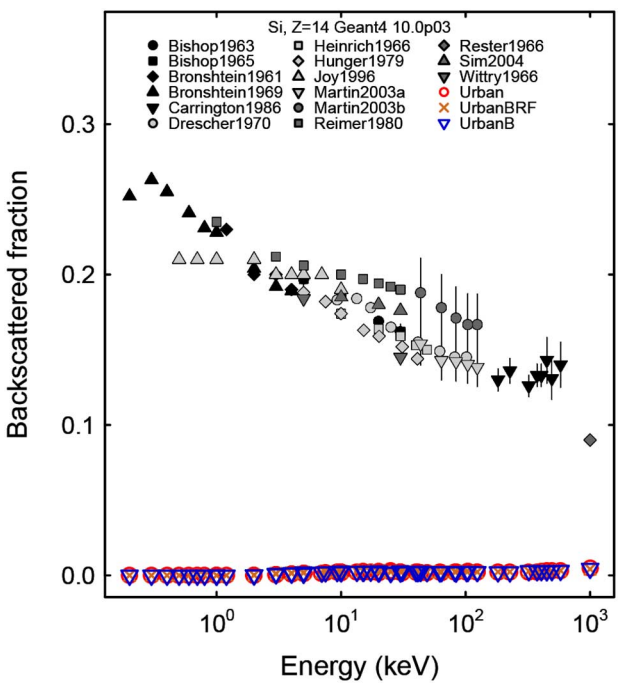

Fig. 14. Fraction of electrons backscattered from a silicon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

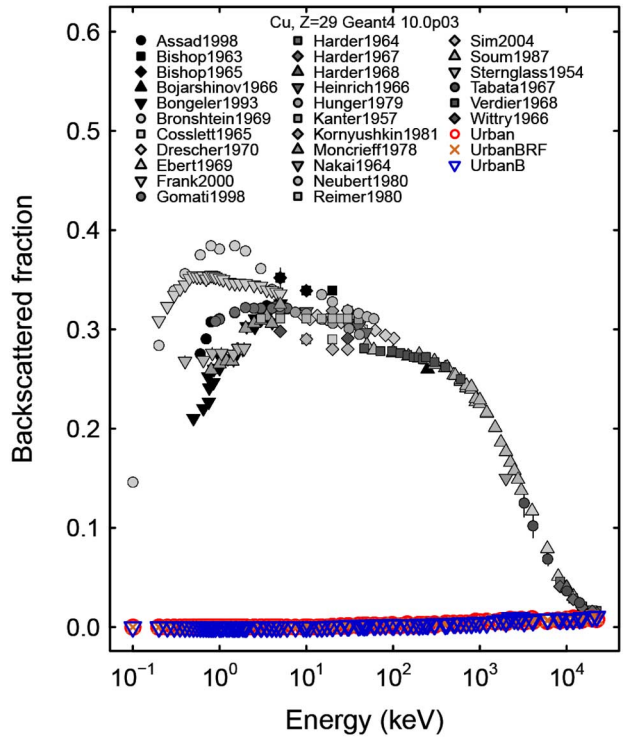

Fig. 15. Fraction of electrons backscattered from a copper target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

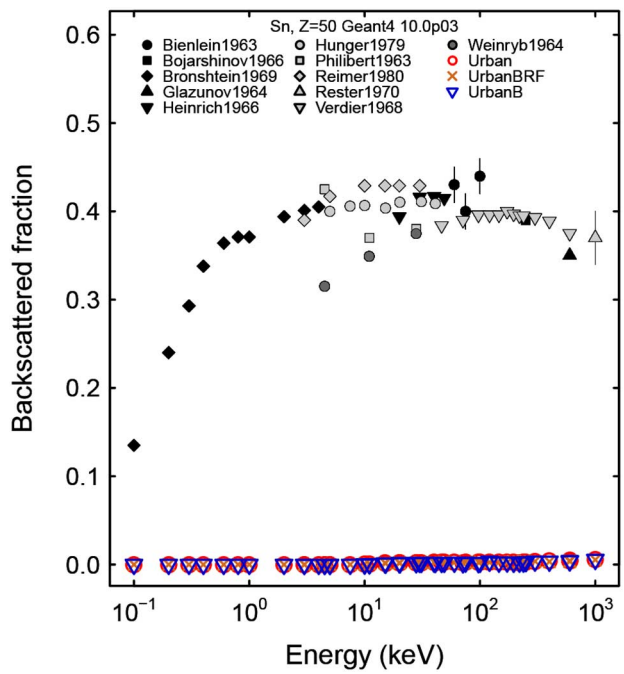

Fig. 16. Fraction of electrons backscattered from a tin target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

step limit type in the multiple scattering algorithm, while they differ in the *RangeFactor* setting, which is not modified with respect to its default value in the UrbanB configuration and is assigned the value of 0.01 in the UrbanBRF configuration. The observed evolution in their compatibility with experiment hints

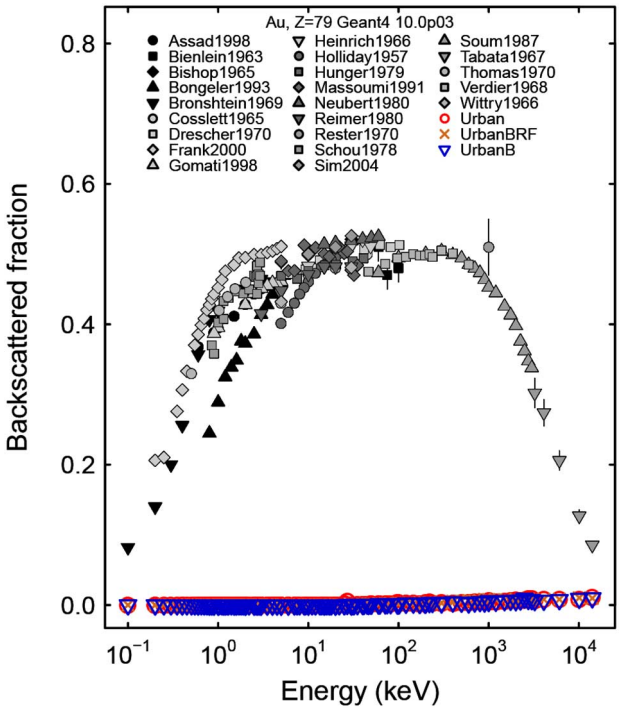

Fig. 17. Fraction of electrons backscattered from a gold target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with Urban (red empty circles), UrbanB (blue empty triangles) and UrbanBRF (brown crosses) multiple scattering configurations.

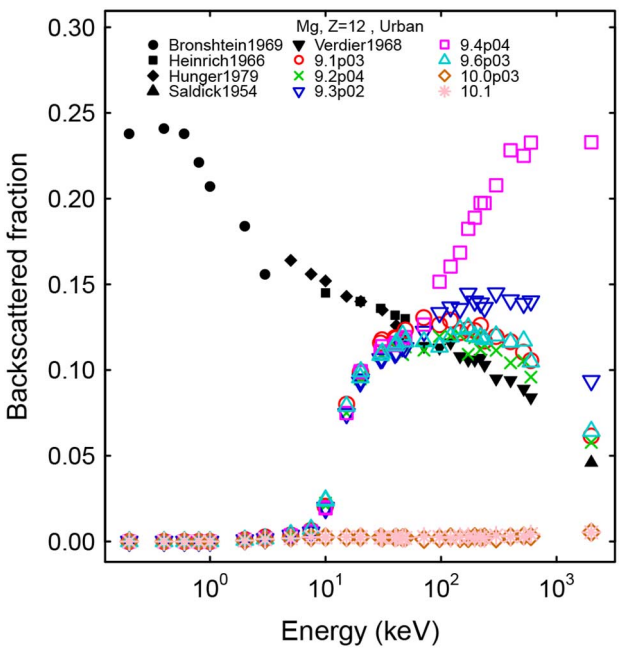

Fig. 18. Fraction of electrons backscattered from a magnesium target as a function of the electron beam energy: experimental data (black and grey filled symbols) and simulation results (empty symbols) with the Urban multiple scattering model, complemented by user defined step limitation, in Geant4 version 9.1 (red circles), 9.2 (green crosses), 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles), 10.0 (brown diamonds) and 10.1 (pink asterisks).

at an effect of the multiple scattering implementation related to this parameter.

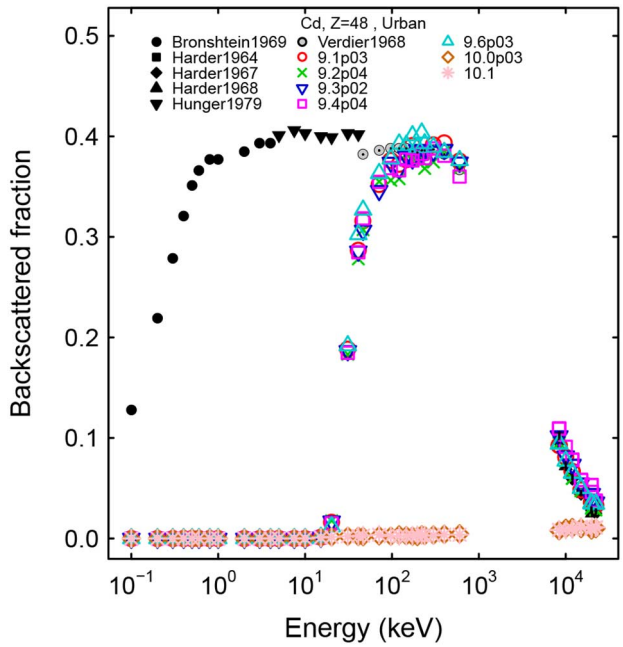

Fig. 19. Fraction of electrons backscattered from a cadmium target as a function of the electron beam energy: experimental data (black and grey filled symbols) and simulation results (empty symbols) with the Urban multiple scattering model, complemented by user defined step limitation, in Geant4 version 9.1 (red circles), 9.2 (green crosses), 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles), 10.0 (brown diamonds) and 10.1 (pink asterisks).

Table IX also shows the p-values resulting from the comparison of the compatibility with experiment achieved with the Urban configuration in Geant4 9.1 and the UrbanB and UrbanBRF variants in the same Geant4 version and later. The hypothesis of equivalent compatibility with experiment is rejected in all cases, with the exception of the UrbanB configuration in Geant4 9.1 and 9.2.

From this analysis one concludes that the multiple scattering settings characterizing the UrbanBRF configuration produce significantly different compatibility with experiment with respect to the Urban one, which achieves the best compatibility with backscattering measurements in the earlier Geant4 versions evaluated in this paper. It is worthwhile to note that the multiple scattering settings of the UrbanBRF configuration reproduce those implemented in the recommended G4EmStandardPhysics_option3 PhysicsConstructor, which is claimed in [94] to enable high accuracy simulation.

A similar categorical analysis estimates the statistical significance of the difference in compatibility with experiment between simulations based on the Urban configuration in Geant4 9.1 and on multiple or single scattering models other than the Urban model in the latest Geant4 versions. The results are summarized in Table X regarding energies above 100 keV. In the energy range between 20 and 100 keV the hypothesis of equivalence is rejected in all test cases, while below 20 keV it is rejected only in test cases involving the Coulomb and WentzelBRF configurations. Given the low efficiency of the Urban model below 20 keV observed in Table VII, the test cases where the hypothesis of equivalence is not rejected reflect the

TABLE X
P-VALUES RESULTING FROM THE COMPARISON OF COMPATIBILITY WITH EXPERIMENT BETWEEN SIMULATIONS USING URBAN MULTIPLE SCATTERING IN GEANT4 9.1 AND SIMULATIONS USING OTHER UNRELATED CONFIGURATIONS IN VERSIONS 9.6, 10.0 AND 10.1, CONCERNING ELECTRON ENERGY ABOVE 100 KEV

Geant4 Version	Physics Configuration	Fisher Test	Pearson χ^2 Test	Barnard Test	Boschloo Test
9.6	Coulomb	1	1	1	1
	GSBRF	0.026	0.016	0.019	0.19
	WentzelBRF	1	1	1	1
	WentzelBRFP	<0.0001	<0.0001	<0.0001	<0.0001
	EmLivermore	<0.0001		<0.0001	<0.0001
	EmStd	<0.0001	<0.0001	<0.0001	<0.0001
	EmOpt1	<0.0001	<0.0001	<0.0001	<0.0001
	EmOpt2	<0.0001	<0.0001	<0.0001	<0.0001
	EmOpt3	<0.0001		<0.0001	<0.0001
	EmOpt4	<0.0001		<0.0001	<0.0001
10.0	Coulomb	1	0.815	0.889	1
	GSBRF	0.026	0.016	0.019	0.19
	WentzelBRF	1	0.815	0.889	1
	WentzelBRFP	<0.0001		<0.0001	<0.0001
	EmLivermore	<0.0001		<0.0001	<0.0001
	EmStd	<0.0001	<0.0001	<0.0001	<0.0001
	EmOpt1	<0.0001	<0.0001	<0.0001	<0.0001
	EmOpt2	<0.0001	<0.0001	<0.0001	<0.0001
	EmOpt3	<0.0001		<0.0001	<0.0001
	EmOpt4	<0.0001	<0.0001	<0.0001	<0.0001
10.1	Coulomb	<0.0001		<0.0001	<0.0001
	CoulombP	1	1	1	1
	GSBRF	<0.0001	<0.0001	<0.0001	<0.0001
	WentzelBRF	<0.0001	<0.0001	<0.0001	<0.0001
	WentzelBRFP	0.0001		0.0001	0.0001
	EmLivermore	0.0001		0.0001	0.0001
	EmStd	<0.0001		<0.0001	<0.0001
	EmOpt1	<0.0001	<0.0001	<0.0001	<0.0001
	EmOpt2	<0.0001	<0.0001	<0.0001	<0.0001
	EmOpt3	<0.0001		<0.0001	<0.0001
	EmOpt4	<0.0001		<0.0001	<0.0001
	EmWVI	<0.0001		<0.0001	<0.0001

TABLE XI
P-VALUES OF MCNEMAR EXACT TEST COMPARING THE COMPATIBILITY WITH EXPERIMENT OF THE COULOMB CONFIGURATION IN GEANT4 10.0 AND IN OTHER GEANT4 VERSIONS

Version	1-20 keV	20-100 keV	≥100 keV
9.1	< 0.001	< 0.001	< 0.001
9.2	< 0.001	< 0.001	< 0.001
9.3	< 0.001	< 0.001	< 0.001
9.4	< 0.001	0.008	< 0.001
9.6	0.625	0.754	0.625
10.1	< 0.001	< 0.001	< 0.001

equivalent incapability of Geant4 multiple scattering models to describe backscattering accurately at the lowest energies.

D. Geant4 Coulomb Scattering Model

The simulation application instantiates Geant4 Coulomb scattering process and its associated model for electron scattering through the respective class constructors and uses them in their default configuration.

A sample of the behaviour of this model with respect to experimental data is illustrated in Figs. 20–22. The single Coulomb scattering model appears to have evolved from Geant4 version 9.1 to 10.0 towards better reproducing experimental measurements especially at energies below a few tens of keV. A degradation of the capability of the model in its default configuration

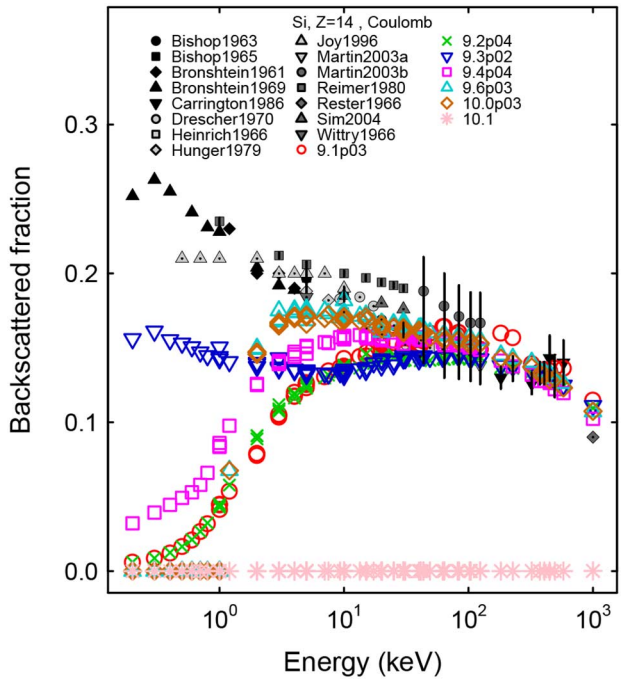

Fig. 20. Fraction of electrons backscattered from a silicon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and simulation results (empty symbols) with the Coulomb scattering model in Geant4 version 9.1 (red circles), 9.2 (green crosses), 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles), 10.0 (brown diamonds) and 10.1 (pink asterisks).

is observed in Geant4 10.1. Physical performance equivalent to that observed in Geant4 10.0 is achieved only by setting the "θ limit" parameter to zero in the singleton G4EmParameters class: this simulation configuration is identified as CoulombP in this paper. Despite our best efforts, we could not retrieve any documentation of the semantic change of the construction of Coulomb scattering objects in Geant4 10.1, which is responsible for the modified behaviour in their default configuration.

These qualitative considerations are confirmed by the results of McNemar's test summarized in Table XI, which compares the compatibility with experiment achieved by the Coulomb scattering confiburation in Geant4 10.0 with the outcome of goodness of fit tests related to earlier versions.

The default Coulomb scattering configuration achieves the highest efficiency over all energy ranges in Geant4 versions 9.6 and 10.0. For convenience, the Coulomb configuration in Geant4 10.0 is defined as the reference Coulomb configuration for further categorical data analysis. Its compatibility with experimental data is compared with the achievements of other multiple scattering configurations in Geant4 versions 9.6, 10.0 and 10.1 in Table XII. The hypothesis of equivalent compatibility with experiment is rejected for all other models implemented in Geant4 10.1 at all energies, as well as for the default Coulomb configuration of that version. Equivalent behaviour is achieved with non-default settings in Geant4 10.1: to the best of our efforts, we could not retrieve mention of the major semantic change of the default instantiation of single Coulomb scattering in Geant4 10.1, nor of the settings required to restore equivalent functionality to previous versions. Regarding earlier Geant4 versions, the hypothesis of

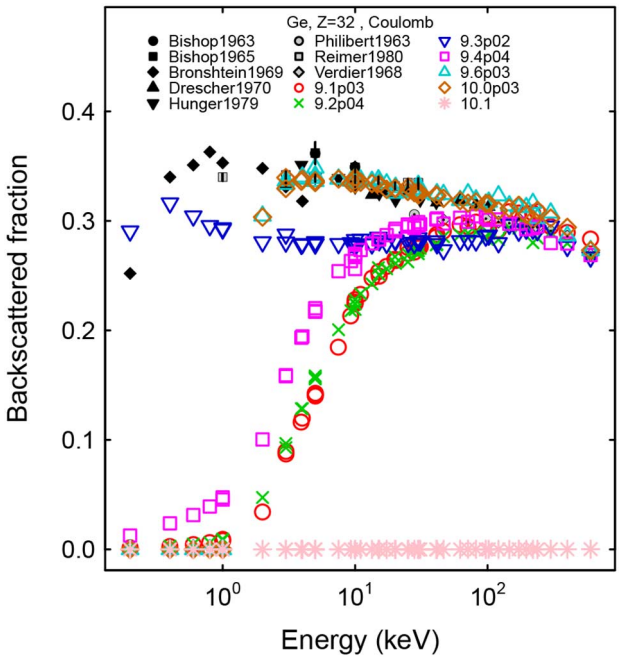

Fig. 21. Fraction of electrons backscattered from a germanium target as a function of the electron beam energy: experimental data (black and grey filled symbols) and simulation results (empty symbols) with the Coulomb scattering model in Geant4 version 9.1 (red circles), 9.2 (green crosses), 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles), 10.0 (brown diamonds) and 10.1 (pink asterisks).

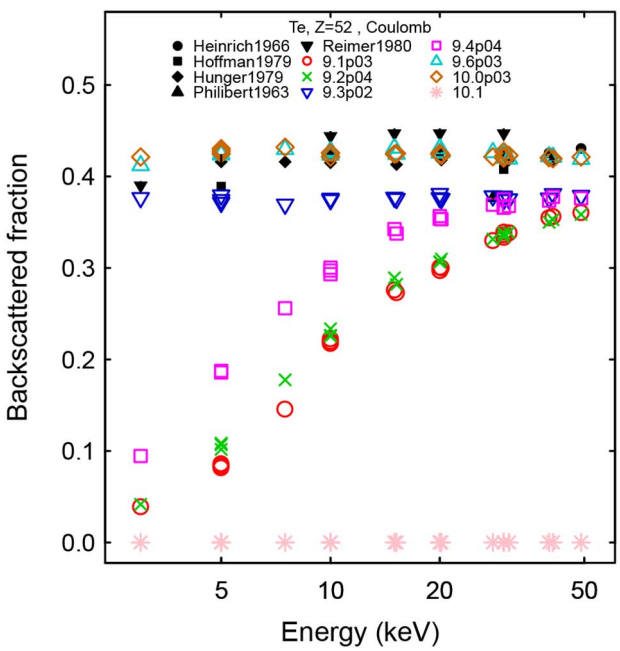

Fig. 22. Fraction of electrons backscattered from a tellurium target as a function of the electron beam energy: experimental data (black and grey filled symbols) and simulation results (empty symbols) with the Coulomb scattering model in Geant4 version 9.1 (red circles), 9.2 (green crosses), 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles), 10.0 (brown diamonds) and 10.1 (pink asterisks).

equivalent compatibility with experiment is not rejected at all energies with respect to the WentzelBRF configuration, which de facto corresponds to enabling single scattering only in the default setting of G4WentzelVIModel. At higher energies the Urban configuration achieves equivalent compatibility with experiment in Geant4 version 9.6, but not in Geant4 10.0. Weak evidence of equivalent compatibility with experiment is reported in Table XII regarding the Goudsmit-Saunderson multiple scattering model in versions 9.6 and 10.0; further details are discussed in Section VI-E.

The CoulombP configuration in Geant4 10.1 exhibits similar behaviour to the reference Coulomb configuration of Geant4 10.0. Its results are not shown in Figs. 20–22 for better clarity of the plots, but they are included in Figs. 32–36. Its associated efficiency is listed in Table VII. McNemar's test confirms

TABLE XII
P-VALUES RESULTING FROM THE COMPARISON OF COMPATIBILITY WITH EXPERIMENT BETWEEN THE SINGLE COULOMB SCATTERING CONFIGURATION IN GEANT4 10.0 AND OTHER UNRELATED CONFIGURATIONS IN VERSIONS 9.6, 10.0 AND 10.1

Energy (kev)	Geant4 Version	Model	Fischer Test	Pearson χ^2 Test	Barnard Test	Boschloo Test
1-20	9.6	Urban	< 0.001		< 0.001	< 0.001
		GSBRF	< 0.001		< 0.001	< 0.001
		WentzelBRF	0.716	0.627	0.683	0.666
	10.0	Urban	< 0.001		< 0.001	< 0.001
		GSBRF	< 0.001		< 0.001	< 0.001
		WentzelBRF	1	0.903	0.951	1
	10.1	Urban	< 0.001		< 0.001	< 0.001
		GSBRF	< 0.001		< 0.001	< 0.001
		WentzelBRF	< 0.001		< 0.001	< 0.001
20-100	9.6	Urban	< 0.001	< 0.001	< 0.001	< 0.001
		GSBRF	< 0.001		< 0.001	< 0.001
		WentzelBRF	0.685	0.588	0.682	0.624
	10.0	Urban	< 0.001		< 0.001	< 0.001
		GSBRF	< 0.001		< 0.001	< 0.001
		WentzelBRF	0.685	0.588	0.682	0.624
	10.1	Urban	< 0.001		< 0.001	< 0.001
		GSBRF	< 0.001		< 0.001	< 0.001
		WentzelBRF	< 0.001		< 0.001	< 0.001
>100	9.6	Urban	0.196	0.132	0.144	0.160
		GSBRF	0.014	0.008	0.009	0.009
		WentzelBRF	1	0.815	0.889	1
		WentzelBRFP	< 0.001	< 0.001	< 0.001	< 0.001
		EmLivermore	< 0.001		< 0.001	< 0.001
		EmStd	< 0.001	< 0.001	< 0.001	< 0.001
		EmOpt1	< 0.001	< 0.001	< 0.001	< 0.001
		EmOpt2	< 0.001	< 0.001	< 0.001	< 0.001
		EmOpt3	< 0.001		< 0.001	< 0.001
		EmOpt4	< 0.001		< 0.001	< 0.001
	10.0	Urban	< 0.001	< 0.001	< 0.001	< 0.001
		GSBRF	0.014	0.008	0.009	0.009
		WentzelBRF	1	1	1	1
		WentzelBRFP	< 0.001	< 0.001	< 0.001	< 0.001
		EmLivermore	< 0.001		< 0.001	< 0.001
		EmStd	< 0.001		< 0.001	< 0.001
		EmOpt1	< 0.001	< 0.001	< 0.001	< 0.001
		EmOpt2	< 0.001	< 0.001	< 0.001	< 0.001
		EmOpt3	< 0.001		< 0.001	< 0.001
		EmOpt4	< 0.001	< 0.001	< 0.001	< 0.001
	10.1	Urban	< 0.001		< 0.001	< 0.001
		GSBRF	< 0.001	< 0.001	< 0.001	< 0.001
		WentzelBRF	< 0.001	< 0.001	< 0.001	< 0.001
		WentzelBRFP	< 0.001	< 0.001	< 0.001	< 0.001
		EmLivermore	< 0.001		< 0.001	< 0.001
		EmStd	< 0.001		< 0.001	< 0.001
		EmOpt1	< 0.001	< 0.001	< 0.001	< 0.001
		EmOpt2	< 0.001	< 0.001	< 0.001	< 0.001
		EmOpt3	< 0.001		< 0.001	< 0.001
		EmOpt4	< 0.001		< 0.001	< 0.001
		EmWVI	< 0.001	< 0.001	< 0.001	< 0.001

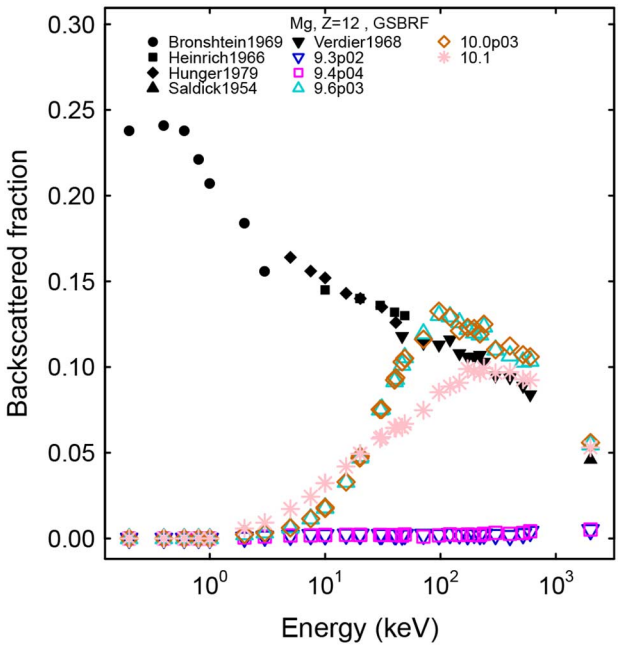

Fig. 23. Fraction of electrons backscattered from a magnesium target as a function of the electron beam energy: experimental data (black and grey filled symbols) and simulation results (empty symbols) with the Goudsmit-Saunderson model in Geant4 version 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles), 10.0 (brown diamonds) and 10.1 (pink asterisks).

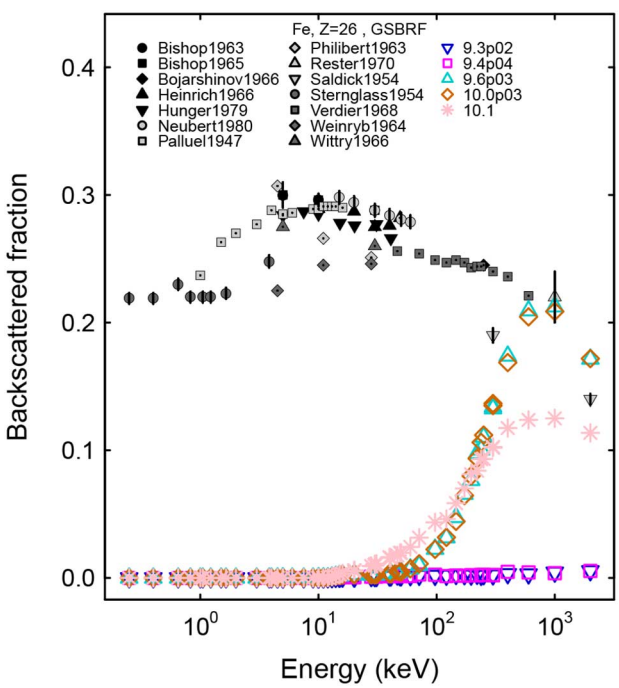

Fig. 24. Fraction of electrons backscattered from an iron target as a function of the electron beam energy: experimental data (black and grey filled symbols) and simulation results (empty symbols) with the Goudsmit-Saunderson model in Geant4 version 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles), 10.0 (brown diamonds) and 10.1 (pink asterisks).

its compatibility with the reference Coulomb configuration of Geant4 10.0 with 0.01 significance.

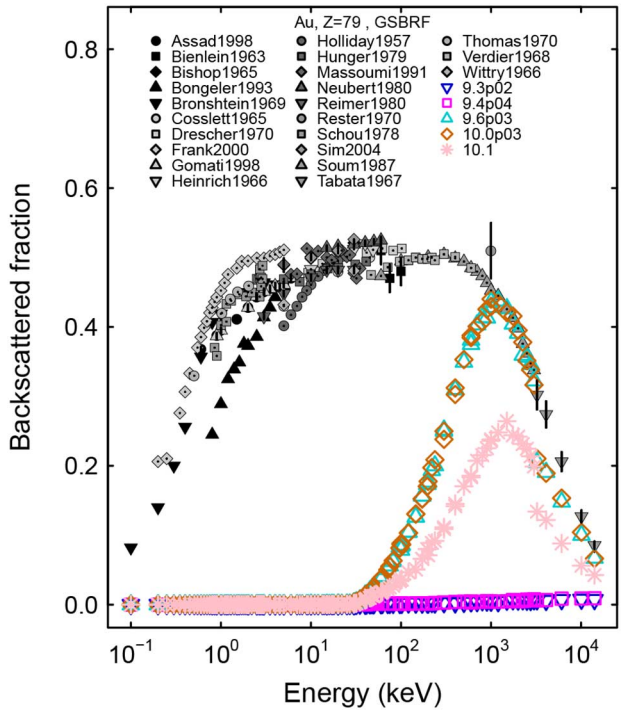

Fig. 25. Fraction of electrons backscattered from a gold target as a function of the electron beam energy: experimental data (black and grey filled symbols) and simulation results (empty symbols) with the Goudsmit-Saunderson model in Geant4 version 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles), 10.0 (brown diamonds) and 10.1 (pink asterisks)..

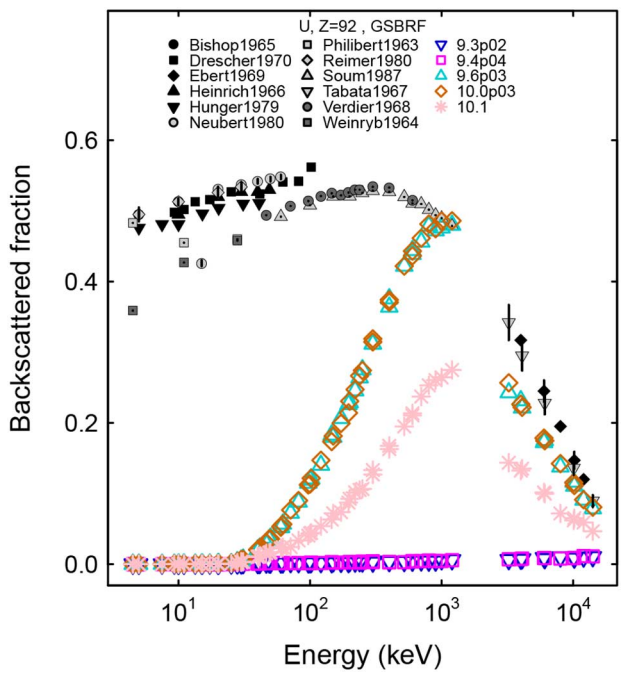

Fig. 26. Fraction of electrons backscattered from an uranium target as a function of the electron beam energy: experimental data (black and grey filled symbols) and simulation results (empty symbols) with the Goudsmit-Saunderson model in Geant4 version 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles), 10.0 (brown diamonds) and 10.1 (pink asterisks)..

The predefined G4EmStandardPhysics_SS PhysicsConstructor, which instantiates single scattering for various particle

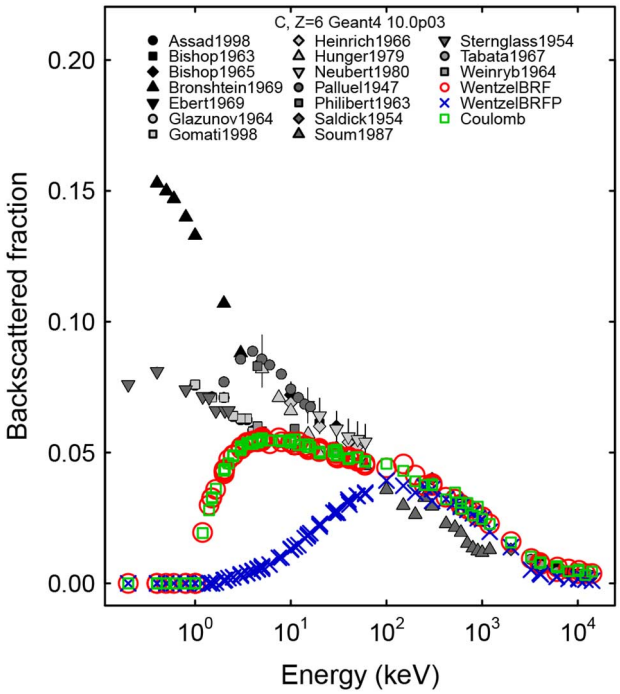

Fig. 27. Fraction of electrons backscattered from a carbon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with WentzelBRF (red empty circles) and WentzelBRFP (blue crosses) multiple scattering settings and Coulomb single scattering model (green empty squares).

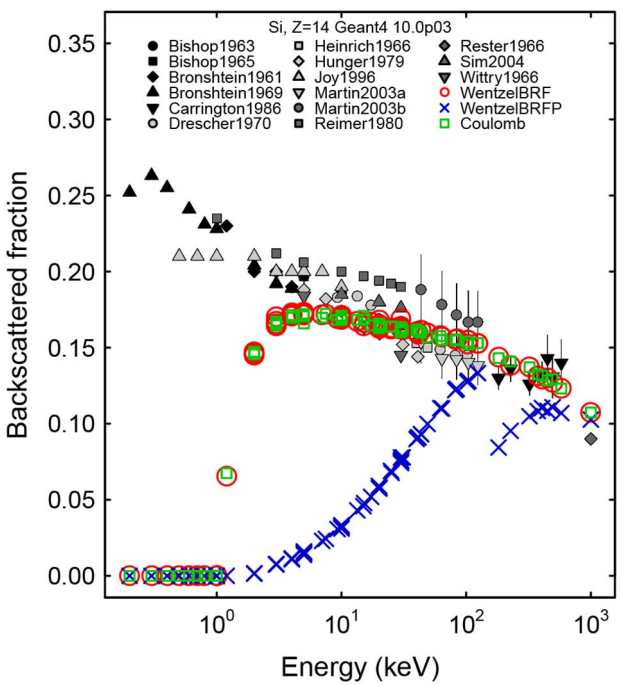

Fig. 28. Fraction of electrons backscattered from a silicon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with WentzelBRF (red empty circles) and WentzelBRFP (blue crosses) multiple scattering settings and Coulomb single scattering model (green empty squares).

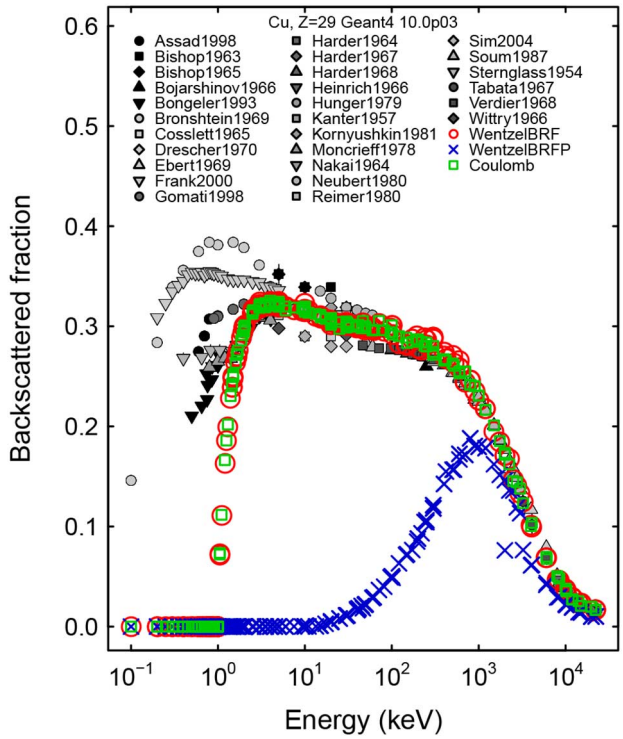

Fig. 29. Fraction of electrons backscattered from a copper target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with WentzelBRF (red empty circles) and WentzelBRFP (blue crosses) multiple scattering settings and Coulomb single scattering model (green empty squares).

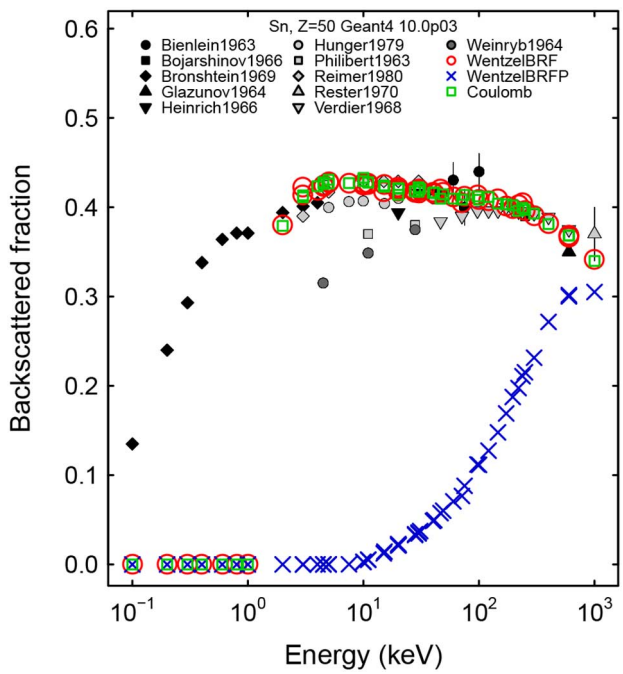

Fig. 30. Fraction of electrons backscattered from a tin target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results withWentzelBRF (red empty circles) and WentzelBRFP (blue crosses) multiple scattering settings and Coulomb single scattering model (green empty squares).

types, produces results statistically consistent to the CoulombP configuration. Since its functionality regarding electron scattering is equivalent to that of the CoulombP configuration, it will not be further discussed in the following sections.

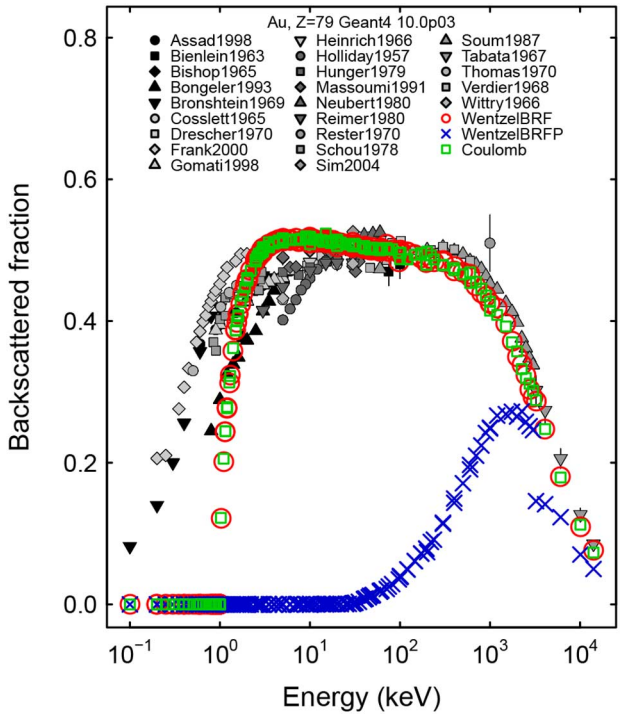

Fig. 31. Fraction of electrons backscattered from a gold target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with WentzelBRF (red empty circles) and WentzelBRFP (blue crosses) multiple scattering settings and Coulomb single scattering model (green empty squares).

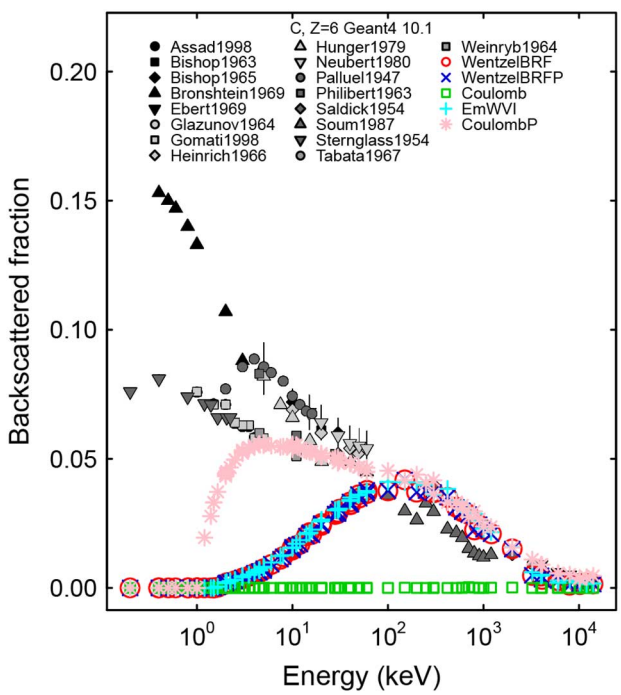

Fig. 32. Fraction of electrons backscattered from a carbon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.1 simulation results with WentzelBRF (red empty circles), WentzelBRFP (blue crosses) multiple scattering settings, Coulomb single scattering model in default configuration (green empty squares) and with modified parameter settings (pink asterisks).

E. Geant4 Goudsmit-Saunderson Model

The implementation of the Goudsmit-Saunderson multiple scattering model appears to have evolved from its first release in Geant4 version 9.3 to version 10.1: qualitatively, one can ob-

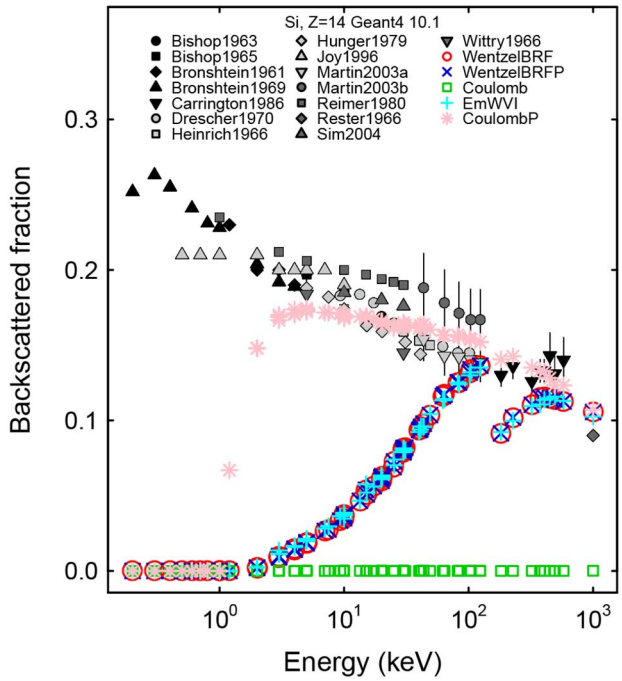

Fig. 33. Fraction of electrons backscattered from a silicon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.1 simulation results with WentzelBRF (red empty circles), WentzelBRFP (blue crosses) multiple scattering settings, Coulomb single scattering model in default configuration (green empty squares) and with modified parameter settings (pink asterisks).

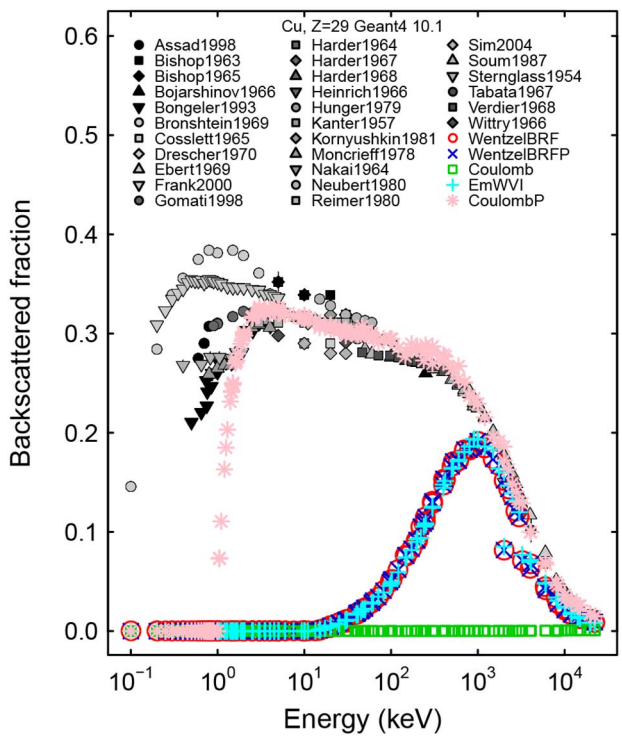

Fig. 34. Fraction of electrons backscattered from a copper target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.1 simulation results with WentzelBRF (red empty circles), WentzelBRFP (blue crosses) multiple scattering settings, Coulomb single scattering model in default configuration (green empty squares) and with modified parameter settings (pink asterisks).

serve in Figs. 23–26 that at higher energies the evolution of this algorithm has contributed to approach experimental backscattering values in simulations based on Geant4 9.6 and 10.0, followed by an apparent deterioration of compatibility in Geant4

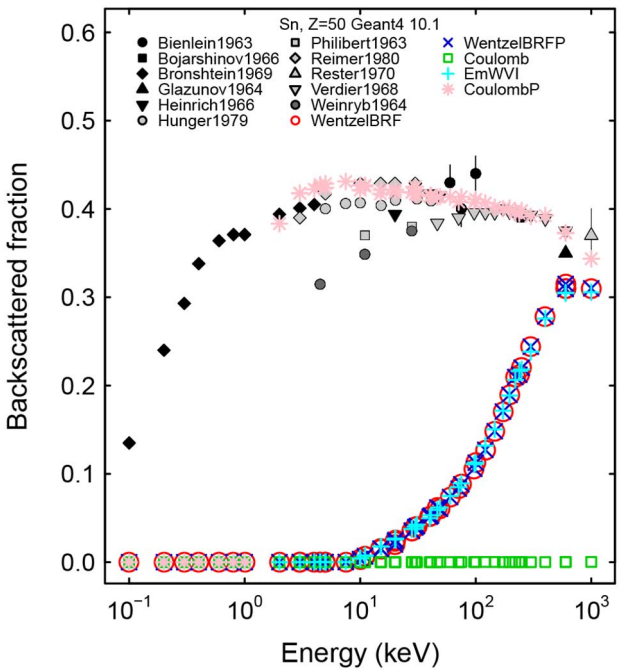

Fig. 35. Fraction of electrons backscattered from a tin target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.1 simulation results WentzelBRF (red empty circles), WentzelBRFP (blue crosses) multiple scattering settings, Coulomb single scattering model in default configuration (green empty squares) and with modified parameter settings (pink asterisks).

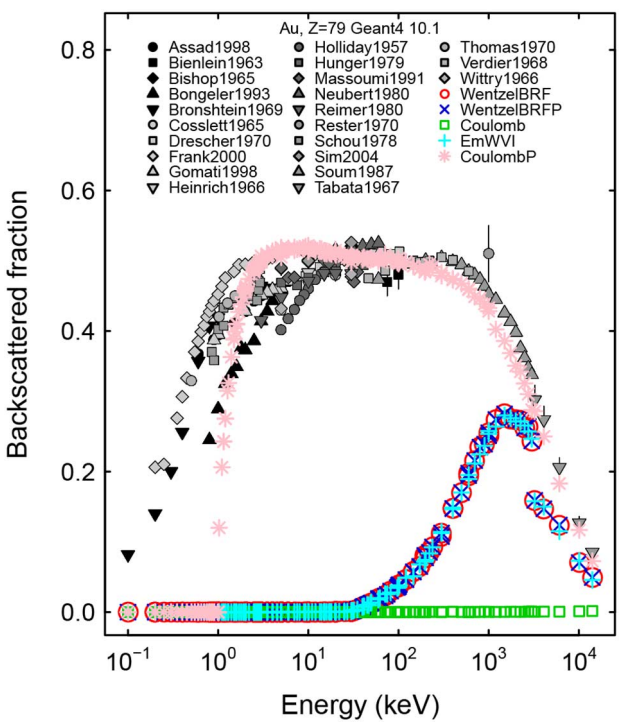

Fig. 36. Fraction of electrons backscattered from a gold target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.1 simulation results WentzelBRF (red empty circles), WentzelBRFP (blue crosses) multiple scattering settings, Coulomb single scattering model in default configuration (green empty squares) and with modified parameter settings (pink asterisks).

10.1. These observations are confirmed by the statistical analysis of compatibility with experiment summarized in Table VII.

The results of the tests reported in Table X show that the compatibility with experiment achieved with the Goudsmit-Saunderson algorithm in Geant4 versions 9.6 and 10.0 is statistically equivalent to that achieved with the Urban configuration in Geant4 version 9.1 in the energy range above 100 keV; nevertheless, equivalent compatibility with experiment with respect to the Coulomb configuration of Geant4 10.0 in the same energy range is assessed only by Fisher test, which is notoriously conservative, while more powerful Z-pooled and Boschloo tests reject the hypothesis of equivalent compatibility with experiment with 0.01 significance. Comparisons at lower energies fail to establish equivalent compatibility with experiment with respect to the 10.0 Coulomb scattering configuration, while they are not physically relevant with respect to the 9.1 Urban configuration due to the very low efficiency exhibited by both multiple scattering models in Table VII. The above mentioned comparisons with the Coulomb scattering configuration of Geant4 10.0 are documented in Table XII.

The evolution of the Goudsmit-Saunderson algorithm in Geant4 10.1 leads to incompatibility with both experimental data and the reference Urban and Coulomb scattering configurations.

F. Geant4 WentzelVI Model

The Geant4 WentzelVI model is used in two configurations evaluated in this paper: WentzelBRF and WentzelBRFP, which use it in its default setting and with a polar angle threshold, respectively. The latter reflects the setting recommended in the electromagnetic PhysicsConstructors that instantiate the WentzelVI model, such as G4EmStandardPhysics and G4EmLivermorePhysics. Their performance in the context of Geant4 10.0 and 10.1 is illustrated in Figs. 27–31 and 32-36, respectively. In addition, the predefined G4EmStandardPhysics_WVI PhysicsConstructor first released in Geant4 10.1 configures multiple scattering with the WentzelVI model for various particle types, including electrons.

The WentzelBRF configuration appears to behave similarly to Geant4 single Coulomb scattering configuration according to the statistical results collected in Tables VIIand XII up to version 10; this similarity also concerns its evolution over the Geant4 versions examined in this paper. It achieves its highest efficiency in Geant4 10.0.

Large differences, especially visible at lower energies, are observed between the results produced by the WentzelBRF and WentzelBRFP configurations in Geant4 versions up to 10.0. The two configurations behave similarly in Geant4 10.1.

The statistical analysis over the earlier Geant4 versions documented in Tables VII, X and XII confirms that, while the WentzelBRF configuration produces statistically equivalent compatibility with experiment with respect to the most efficient configurations (Urban in Geant4 9.1 and Coulomb in Geant4 10.0), the hypothesis of equivalent compatibility with experiment is rejected for the WentzelBRFP one.

From these results one can infer that in the experimental scenarios evaluated in this paper the recommended polar angle setting contributes to worsening the reproduction of measured backscattering with respect to the default settings.

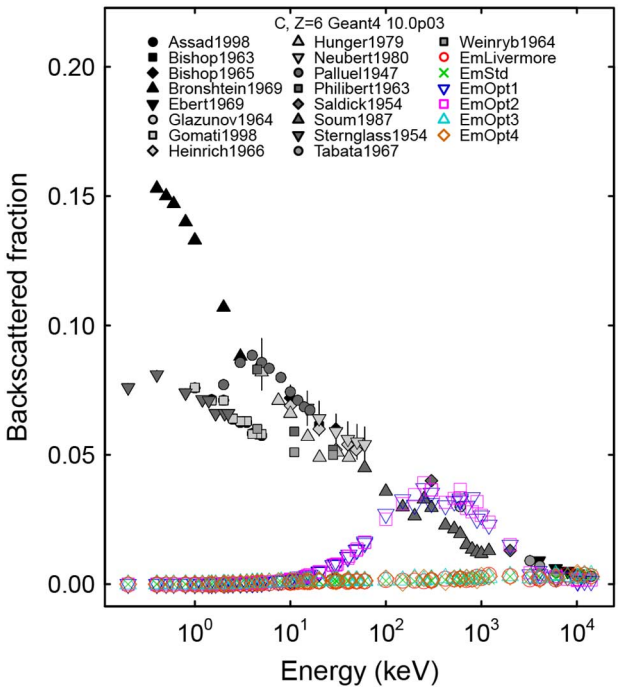

Fig. 37. Fraction of electrons backscattered from a carbon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with G4EmLivermorePhysics (red empty circles), G4EmStandardPhysics (green crosses), G4EmStandardPhysics_option1 (blue empty upside-down triangles), G4EmStandardPhysics_option2 (magenta empty squares), G4EmStandardPhysics_option3 (turquoise empty triangles) and G4EmStandardPhysics_option4 (brown empty diamonds) PhysicsConstructors.

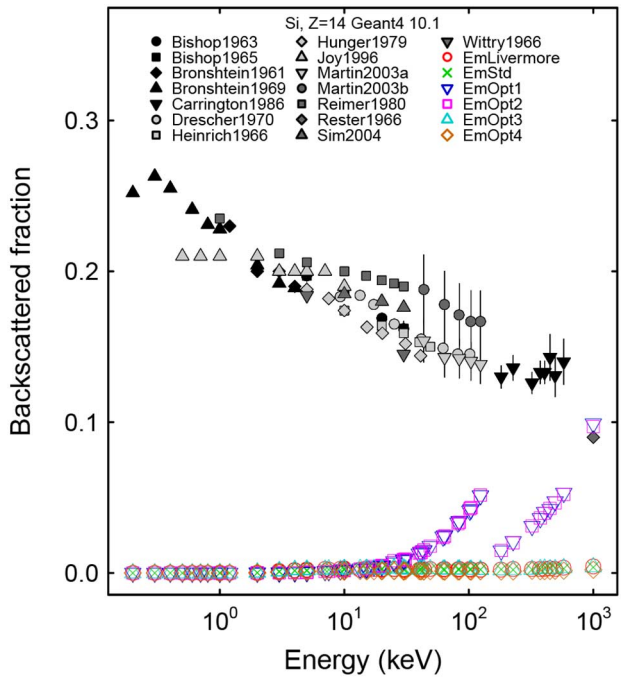

Fig. 38. Fraction of electrons backscattered from a silicon target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.1 simulation results with G4EmLivermorePhysics (red empty circles), G4EmStandardPhysics (green crosses), G4EmStandardPhysics_option1 (blue empty upside-down triangles), G4EmStandardPhysics_option2 (magenta empty squares), G4EmStandardPhysics_option3 (turquoise empty triangles) and G4EmStandardPhysics_option4 (brown empty diamonds) PhysicsConstructors.

The WentzelBRF and WentzelBRFP configurations produce statistically equivalent results in Geant4 10.1: McNemar's test fails to reject the hypothesis of compatibility between the two categories of goodness-of-fit testing results with 0.01 significance.

The behaviour of the predefined G4EmStandardPhysics_WVI PhysicsConstructor is equivalent to that of the WentzelBRF and WentzelBRFP configurations in Geant4 10.1: this conclusion is assessed by McNemar's test with 0.01 significance.

G. Geant4 Recommended PhysicsConstructors

As stated in Section IV-B, the evaluation of Geant4 recommended PhysicsConstructors regarding their simulation of electron backscattering concerns Geant4 versions 9.6, 10.0 and 10.1.

The fraction of backscattered electrons produced by G4EmStandardPhysics, G4EmStandardPhysics_option1, G4EmStandardPhysics_option2, G4EmStandardPhysics_option3 G4EmStandardPhysics_option4 and G4EmLivermorePhysics with Geant4 version 10.0 is illustrated in Figs. 37–41. One can qualitatively observe that simulation results produced by G4EmStandardPhysics, G4EmStandardPhysics_option3, G4EmStandardPhysics_option4 and G4EmLivermorePhysics fail to reproduce the characteristics of experimental distributions over the whole energy range. Simulations using G4EmStandardPhysics_option1 and G4EmStandardPhysics_option2 appear to approach experimental data in the higher energy end above 1 MeV.

These qualitative considerations are reflected in the outcome of the statistical data analysis. At lower energies the recommended PhysicsConstructors appear incapable to reproduce experimental data similarly to other Geant4 multiple scattering configurations. G4EmStandardPhysics_option3 and G4EmStandardPhysics_option4, which according to Geant4 documentation are intended to produce high accuracy simulations, exhibit negligible efficiency at all energies, as well as G4EmLivermorePhysics. At higher energies the efficiency achieved using G4EmStandardPhysics_option1 and G4EmStandardPhysics_option2 is higher than that associated with the other recommended PhysicsConstructors in Geant4 9.6 and 10.0 is approximately a factor two lower than that achieved in the most efficient scenarios (Urban multiple scattering configuration in Geant4 9.1 and Coulomb single scattering configuration in Geant4 10.0). The performance of G4EmStandardPhysics_option1 is consistent with the statement in Geant4 user documentation that this PhysicsConstructor is intended for fast, but less accurate simulation. G4EmStandardPhysics exhibits a similar performance in Geant4 9.6, while its efficiency drops in Geant4 10.0.

For all the recommended PhysicsConstructors the hypothesis of equivalent compatibility with the most efficient configurations (Urban in Geant4 9.1 and Coulomb in Geant4 10.0) is rejected with 0.0001 significance above 100 keV.

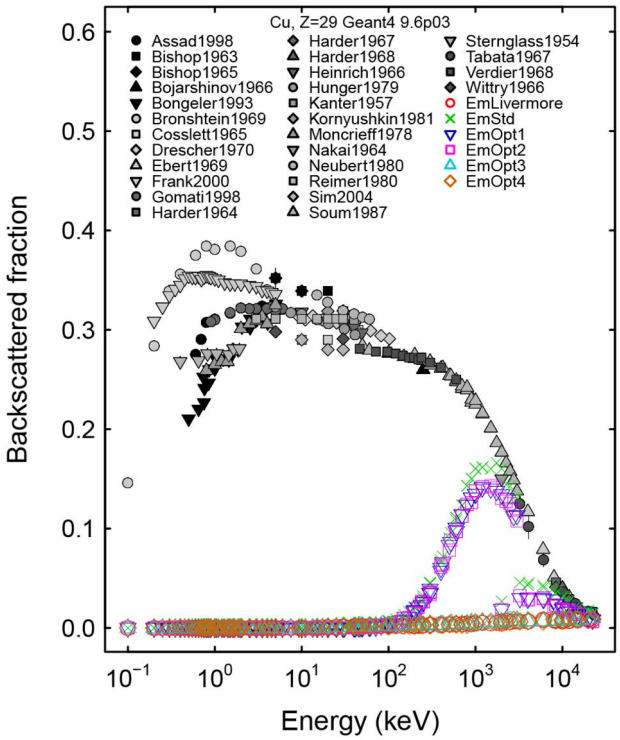

Fig. 39. Fraction of electrons backscattered from a copper target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 9.6 simulation results with G4EmLivermorePhysics (red empty circles), G4EmStandardPhysics (green crosses), G4EmStandardPhysics_option1 (blue empty upside-down triangles), G4EmStandardPhysics_option2 (magenta empty squares), G4EmStandardPhysics_option3 (turquoise empty triangles) and G4EmStandardPhysics_option4 (brown empty diamonds) PhysicsConstructors.

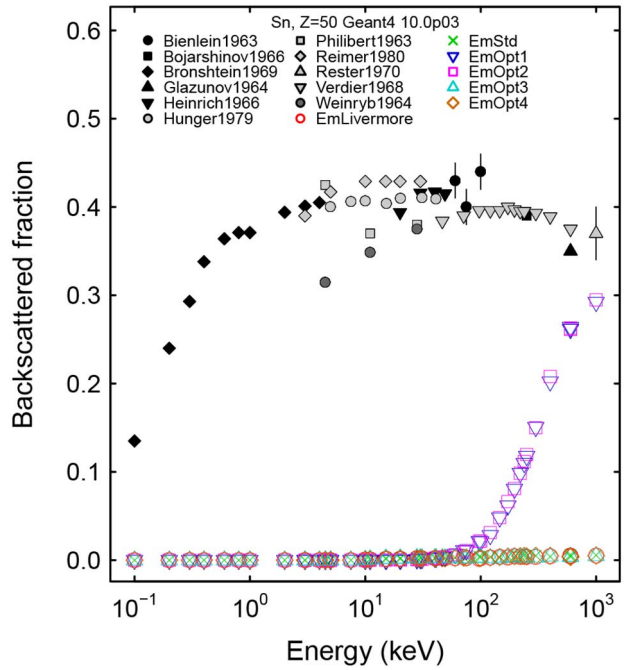

Fig. 40. Fraction of electrons backscattered from a tin target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.0 simulation results with G4EmLivermorePhysics (red empty circles), G4EmStandardPhysics (green crosses), G4EmStandardPhysics_option1 (blue empty upside-down triangles), G4EmStandardPhysics_option2 (magenta empty squares), G4EmStandardPhysics_option3 (turquoise empty triangles) and G4EmStandardPhysics_option4 (brown empty diamonds) PhysicsConstructors.

As a result of this analysis, one can conclude that in the scenario examined in this test significantly better accuracy can be achieved with physics configurations other than those recommended in Geant4 user documentation.

H. Computational Performance

The heterogeneous production environment of the simulations documented in this paper, consisting of different platforms, prevents the absolute comparison of the computational performance of different physics modeling options. Nevertheless, in most test cases the simulations corresponding to a group of multiple scattering configurations (Urban model and its variants, WentzelVI model variants, Coulomb, GSBRF) were executed on the same machine, thus enabling a relative comparison of their computational performance at least at a qualitative level.

A sample of results in Figs. 42 and 43 illustrate the main features regarding the computational performance of backscattering simulations, namely the overhead associated with simulating single scattering rather than multiple scattering, in Geant4 version 9.6. One can observe that the Coulomb and WentzelBRF configurations exhibit similar computational performance, while the Urban configuration is an order of magnitude faster. All recommended PhysicsConstructors are substantially faster than the other configurations shown in the plots, although at the price of significantly degraded compatibility with experiment, as documented in the previous sections.

VII. Correlation Between Backscattering and Energy Deposition

It is physically intuitive that a relation exists between the fraction of electrons that are backscattered from a target volume and the energy deposited in it. A few examples of the energy deposited in the targets used in the simulation of electron backscattering are shown in Figs. 44–49: they illustrate the dependence of the energy deposition on the evolution of the Geant4 Urban model from version 9.1 to 10.0 in Figs. 44–45 and the effects associated with different physics configurations in the simulation in Figs. 46–49. Visible differences are observed both among different Geant4 versions for the same physics configuration, and among different physics configurations in the same Geant4 version.

The capability of Geant4 to reproduce experimental measurements of the energy deposited by low energy electrons is quantitatively analyzed in [1] with respect to two experimental configurations: a longitudinally segmented detector and a detector consisting of a bulk volume. The latter closely resembles the simulation configuration of the backscattering test cases evaluated in this paper. Reference [1] reports the efficiency, representing the compatibility with experimental data of energy deposition simulations based on Geant4 versions 9.1 to 9.6: these values were obtained using the default Urban multiple scattering model corresponding to each Geant4 version and imposing a

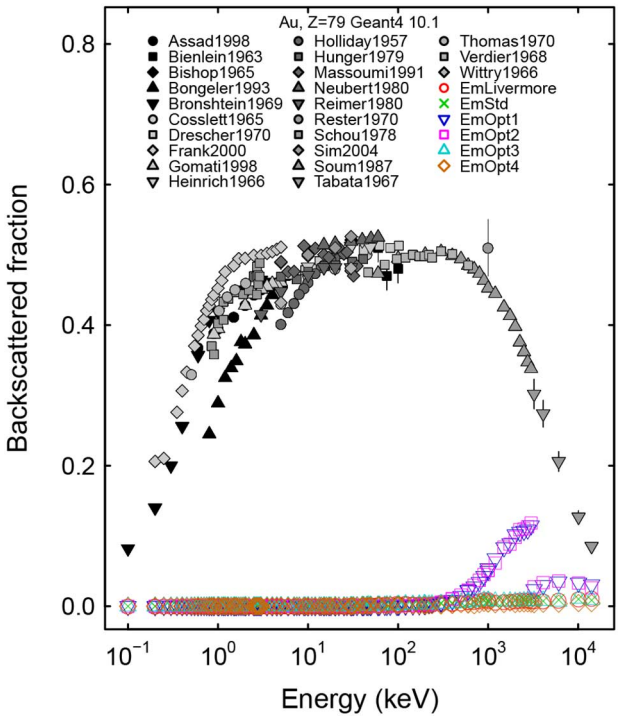

Fig. 41. Fraction of electrons backscattered from a gold target as a function of the electron beam energy: experimental data (black and grey filled symbols) and Geant4 10.1 simulation results with G4EmLivermorePhysics (red empty circles), G4EmStandardPhysics (green crosses), G4EmStandardPhysics_option1 (blue empty upside-down triangles), G4EmStandardPhysics_option2 (magenta empty squares), G4EmStandardPhysics_option3 (turquoise empty triangles) and G4EmStandardPhysics_option4 (brown empty diamonds) PhysicsConstructors.

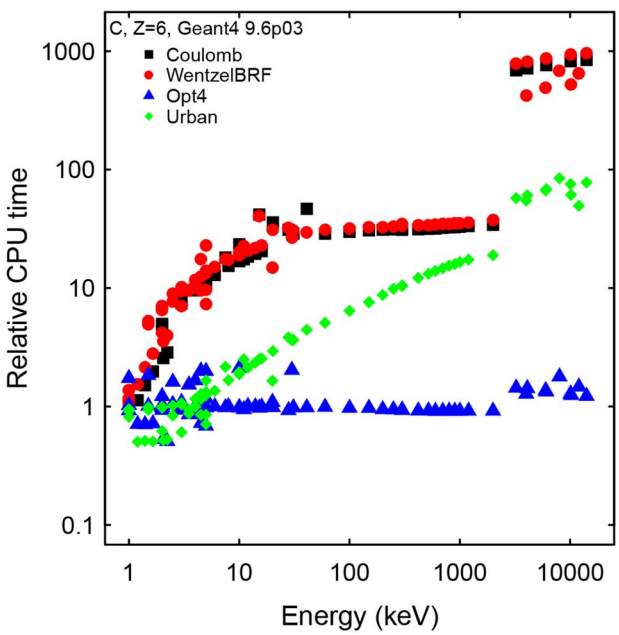

Fig. 42. Relative execution time as a function of beam energy for a set of physics configurations simulating electrons backscattered from a carbon target with Geant4 9.6: single Coulomb scattering (black squares), WentzelBRF (red circles), G4EmStandardPhysics_option4 (blue triangles) and Urban (green diamonds). The results in the plot are scaled with respect to the execution time of a simulation with G4EmStandardPhysics.

user-defined step limitation, similar to the configuration identified as Urban in this paper. The hypothesis was formulated in

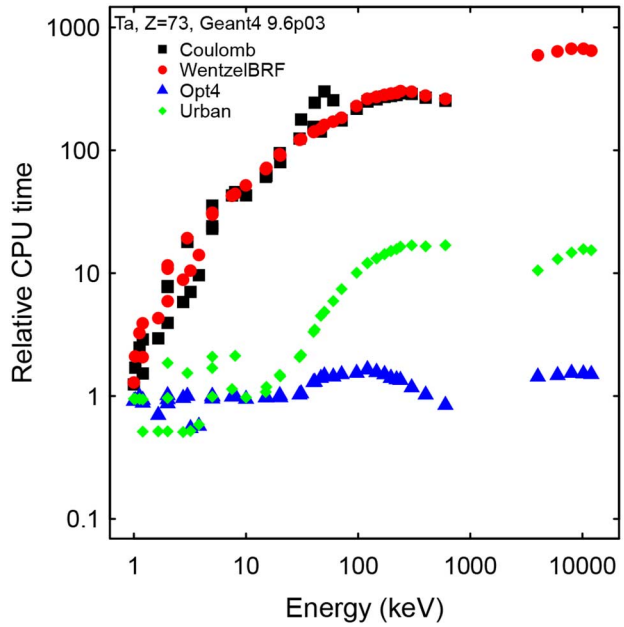

Fig. 43. Relative execution time as a function of beam energy for a set of physics configurations simulating electrons backscattered from a tungsten target with Geant4 9.6: single Coulomb scattering (black squares), WentzelBRF (red circles), G4EmStandardPhysics_option4 (blue triangles) and Urban (green diamonds). The results in the plot are scaled with respect to the execution time of a simulation with G4EmStandardPhysics.

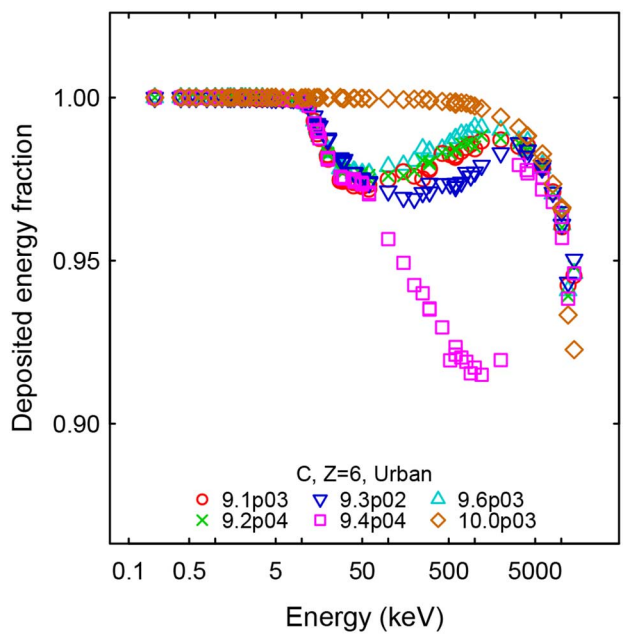

Fig. 44. Energy deposited in a carbon target as a function of the electron beam energy, resulting from simulations with the Urban multiple scattering configuration in different Geant4 versions: 9.1 (red circles), 9.2 (green crosses), 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles) and 10.0 (brown diamonds).

[1] that the discrepancies in compatibility with experiments obtained with different Geant4 versions could be attributed to the evolution of the implementation of the Urban Geant4 multiple scattering algorithm.

The test of electron backscattering simulation, which is a direct effect of multiple scattering algorithms implemented

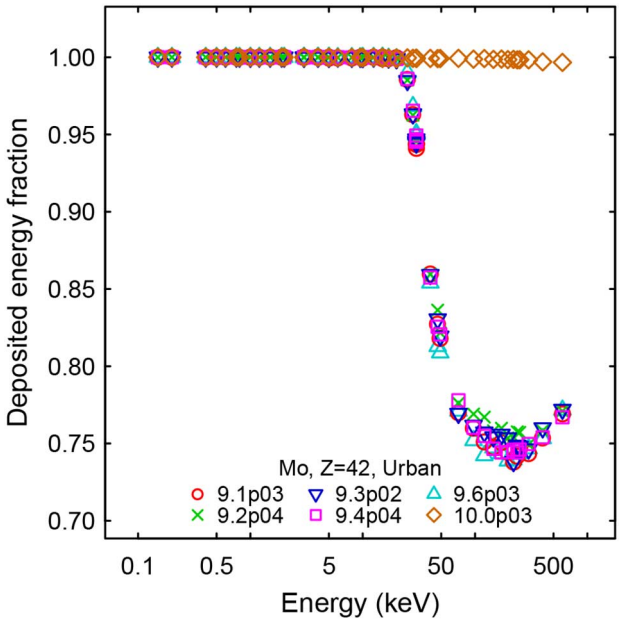

Fig. 45. Energy deposited in a molybdenum target as a function of the electron beam energy, resulting from simulations with the Urban multiple scattering configuration in different Geant4 versions: 9.1 (red circles), 9.2 (green crosses), 9.3 (blue upside down triangles), 9.4 (magenta squares), 9.6 (turquoise triangles) and 10.0 (brown diamonds).

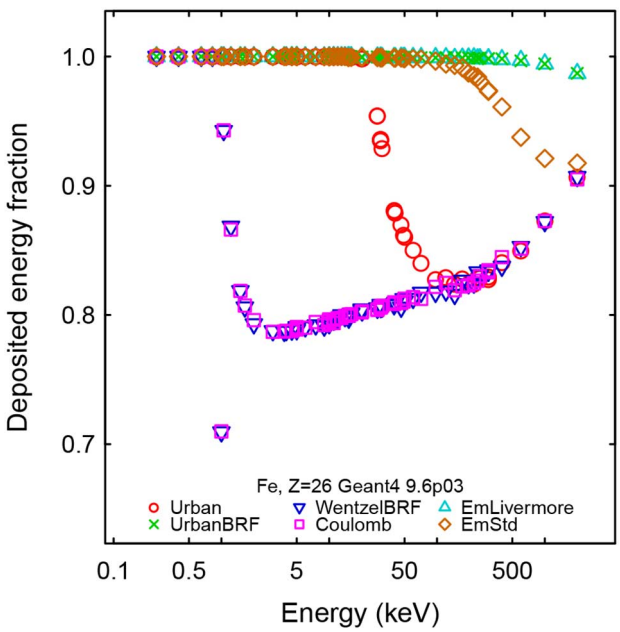

Fig. 46. Energy deposited in an iron target as a function of the electron beam energy, resulting from simulations with different multiple scattering configurations in Geant4 version 9.6: Urban (red circles), UrbanBRF (green crosses) WentzelBRF (blue upside down triangles) and Coulomb (magenta squares) electron scattering configurations, EmLivermore (turquoise triangles) and EmStandard (brown diamonds) pre-packaged PhysicsConstructors).

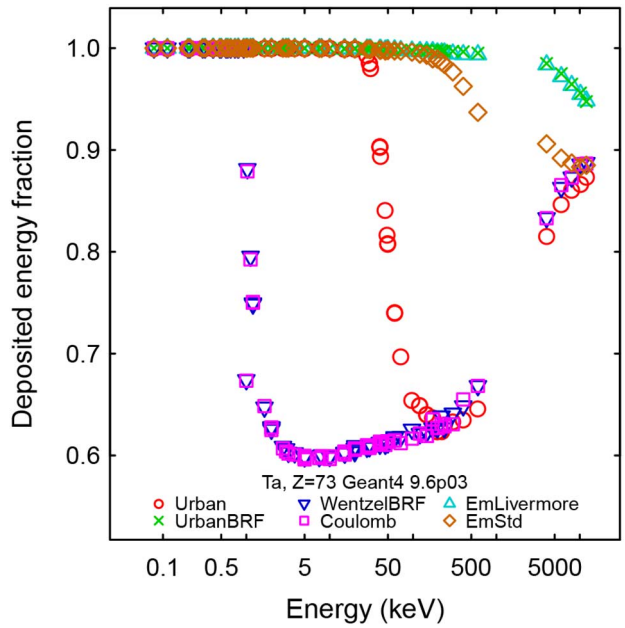

Fig. 47. Energy deposited in a tantalum target as a function of the electron beam energy, resulting from simulations with different multiple scattering configurations in Geant4 version 9.6: Urban (red circles), UrbanBRF (green crosses) WentzelBRF (blue upside down triangles) and Coulomb (magenta squares) electron scattering configurations, EmLivermore (turquoise triangles) and EmStandard (brown diamonds) pre-packaged PhysicsConstructors.

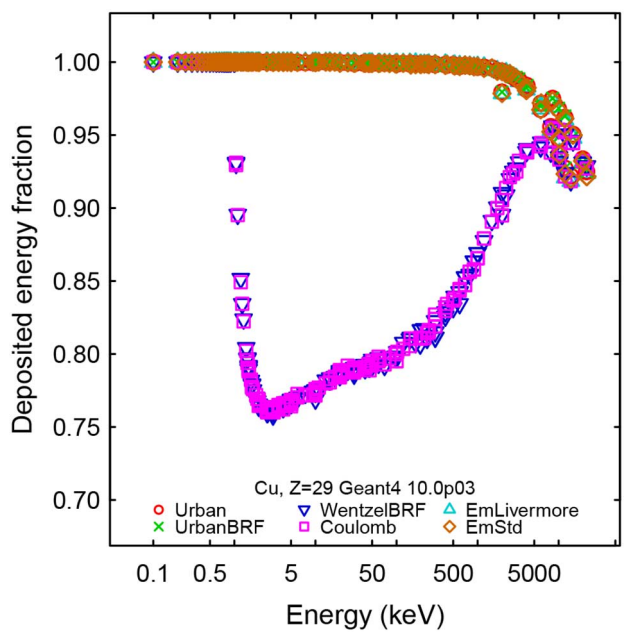

Fig. 48. Energy deposited in a copper target as a function of the electron beam energy, resulting from simulations with different multiple scattering configurations in Geant4 version 10.0: Urban (red circles), UrbanBRF (green crosses) WentzelBRF (blue upside down triangles) and Coulomb (magenta squares) electron scattering configurations, EmLivermore (turquoise triangles) and EmStandard (brown diamonds) pre-packaged PhysicsConstructors.

in Geant4, allows a quantitative test of this hypothesis: the evaluation whether a correlation exists between the two sets of efficiencies, derived from the validation of energy deposition and backscattering simulations, respectively. It is worthwhile to note that measures of correlation are not inferential statistical tests, but are, instead, descriptive statistical quantities, which represent the degree of relationship between two observables. Statistical inference concerning the underlying population is enabled by the analysis of the significance of the measured value through appropriate hypothesis testing methods.

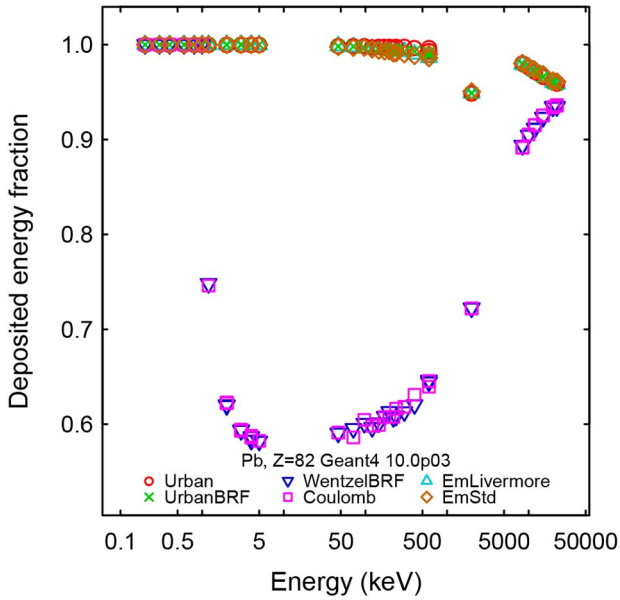

Fig. 49. Energy deposited in a lead target as a function of the electron beam energy, resulting from simulations with different multiple scattering configurations in Geant4 version 10.0: Urban (red circles), UrbanBRF (green crosses) WentzelBRF (blue upside down triangles) and Coulomb (magenta squares) electron scattering configurations, EmLivermore (turquoise triangles) and EmStandard (brown diamonds) pre-packaged PhysicsConstructors.

TABLE XIII
CORRELATION BETWEEN EFFICIENCY AT SIMULATING BACKSCATTERING COEFFICIENT AND ENERGY DEPOSITION COMPATIBLE WITH EXPERIMENT OVER GEANT4 VERSIONS 9.1 TO 9.6

Correlation	Measure	p-value
Kendall τ	1	0.008
Spearman ρ	1	0.008
Pearson correlation coefficient	0.992	0.0005

For the purpose of measuring the correlation between the two sets of efficiencies, the data sample over which the backscattering efficiency is calculated has been restricted to the energy range of the energy deposition test described in [1], and the correlation analysis is limited to Geant4 versions common to both validation tests.

Measures of correlation related to these observables are calculated and complemented by inferential statistical tests [131], [132]. The null hypothesis is formulated as the absence of any correlation between the two sets of efficiencies, expressed by a measure of zero. The alternative hypothesis concerns the existence of a positive correlation between the two sets of efficiencies associated with Geant4 versions; it corresponds to the execution of one-tailed tests.

Two nonparametric correlation measures, Spearman ρ correlation [133] and Kendall τ correlation [134], are reported in Table XIII along with the associated p-values. Table XIII also lists Pearson product-moment correlation coefficient [135] and the associated p-value: although more widely known in the experimental physics environment, Pearson correlation coefficient has a more limited scope, as it describes linear correlation only.

The null hypothesis of no correlation is rejected with 0.01 significance. Consistent with the alternative hypothesis, one can infer that a positive correlation exists. From this analysis one can infer that the accuracy of simulation of electron backscattering, which is a direct effect of multiple scattering modeling, and of simulation of the energy deposited by low energy electrons are correlated.

VIII. CONCLUSION

This paper has analyzed the capabilities of several Geant4 multiple scattering models with respect to a large experimental data sample, and its evolution over Geant4 versions from 9.1 to 10.1. In addition, a single scattering algorithm has been evaluated. The fraction of electrons backscattered from a semi-infinite target has proven to be a sensitive probe of multiple scattering algorithms available in Geant4. A significant correlation between the accuracy of this observable and the accuracy of the simulation of the energy deposition originating from low energy electrons has been established on rigorous statistical grounds.

Large variability is observed in the performance of all models over the range of Geant4 versions considered in this study. The evolution does not always go in the direction of improving the compatibility with experiment: statistically significant regressions are observed for some Geant4 models with respect to their previous behaviour. Although an analysis of the software development process in the electromagnetic physics domain of Geant4 is outside the scope of this paper, the results of this validation test highlight the opportunity to strengthen the discipline of change management and the traceability of changes, including their side effects. Established software process frameworks, such as the Unified Process [136], CMMI (Capability Maturity Model Integration) [137] and the ISO/IEC 15504 Standard [138], provide support for these and other related disciplines.

Due to the variability of physics performance affecting all models, it is not possible to identify a single Geant4 version providing an optimal simulation environment over the whole energy range considered in the test. All the evaluated multiple scattering models encounter difficulties at reproducing low energy backscattering measurements: up to a few tens of keV only the single Coulomb scattering model demonstrates some capability to describe experimental data realistically, limited to Geant4 9.6 and 10.0 versions in its default configuration, although at the price of substantially slower computational performance. Above a few tens of keV, the Urban model in Geant4 9.1, complemented by user defined step limitation, demonstrates the best capability to reproduce experimental data in a condensed transport scheme. Recommended settings advertised for high accuracy, such as *UseDistanceToBoundary*, worsen the compatibility with experiment of this configuration.

In the investigated scenario the predefined electromagnetic PhysicsConstructors do not fulfill the expectations of accuracy deriving from their advertisement in Geant4 documentation and the statement of their intensive validation in conference papers. The recommendations of their use should be revised and based on objective grounds, documenting the scope to which they are pertinent and quantifying the capabilities of the embedded physics configurations with respect to experimental data.

The quantitative assessment documented in this paper, along with the results of the validation test documented in [1], allow

Geant4 users to optimize the physics configuration of simulation applications involving electrons of energy up to a few tens of MeV, either as primary particles or as secondary interaction products. Experimental scenarios for which accurate simulation of electron backscattering is relevant would benefit from selecting multiple or single scattering configurations in their simulation applications, which best reproduce experimental measurements in the energy range of interest.

ACKNOWLEDGMENT

The support of the CERN Library has been essential to this study. The authors thank Mihaly Novak for valuable discussions regarding Geant4 multiple scattering models, Sergio Bertolucci for support, Anita Hollier for proofreading the manuscript, Fabio Pratolongo and the Computing Service at INFN Genova for help with the computing hardware used for simulation productions, and Tatsuo Tabata for sharing his data. S. H. Kim thanks CERN PH/SFT Group for the hospitality at CERN during summer 2014.

REFERENCES

[1] M. Batic, G. Hoff, M. G. Pia, P. Saracco, and G. Weidenspointner, "Validation of Geant4 simulation of electron energy deposition," *IEEE Trans. Nucl. Sci.*, vol. 60, no. 4, pp. 2934–2957, Aug. 2013.

[2] A. Lechner, M. G. Pia, and M. Sudhakar, "Validation of Geant4 low energy electromagnetic processes against precision measurements of electron energy deposit," *IEEE Trans. Nucl. Sci.*, vol. 56, no. 2, pp. 398–416, Apr. 2009.

[3] S. Agostinelli *et al.*, "Geant4 - a simulation toolkit," *Nucl. Instrum. Methods Phys. Res. A*, vol. 506, no. 3, pp. 250–303, 2003.

[4] J. Allison *et al.*, "Geant4 developments and applications," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 1, pp. 270–278, Feb. 2006.

[5] G. Lockwood *et al.*, Calorimetric Measurement of Electron Energy Deposition in Extended Media - Theory vs. Experiment Sandia National Laboratories, Albuquerque, NM, USA, SAND79-0414 Rep., 1980.

[6] G. J. Lockwood *et al.*, Electron Energy and Charge Albedos - Calorimetric Measurement vs. Monte Carlo Theory, Sandia National Laboratories, Albuquerque, NM, USA, SAND80-1968 Rep., 1981.

[7] P. Tolias, "On electron backscattering from dust grains in fusion plasmas," *Plasma Phys. Control. Fusion*, vol. 56, p. 045003, 2014.

[8] M. Dapor, "Transport of energetic electrons in solids," in *Springer Tracts in Modern Physics*. Berlin, Germany: Springer, 2014, vol. 257, pp. 65–79.

[9] M. Hannachi, Z. Rouabah, C. Champion, and N. Bouarissa, "Electron backscattering from solid targets: Elastic scattering calculations," *J. Electron Spectroscopy Related Phenomena*, vol. 195, pp. 155–159, 2014.

[10] J. W. Martin *et al.*, "Measurement of electron backscattering in the energy range of neutron decay," *Phys. Rev. C*, vol. 68, no. 5, p. 055503, 2003.

[11] M. M. El Gomati, C. G. H. Walker, A. M. D. Assad, and M. Zadrazil, "Theory experiment comparison of the electron backscattering factor from solids at low electron energy (250-5,000 eV)," *Scanning*, vol. 30, pp. 2–15, 2008.

[12] A. M. D. Assad and M. M. El-Gomati, "Backscattering coefficients for low energy electrons," *Scanning Microscopy*, vol. 12, no. 1, pp. 185–192, 1998.

[13] J. K. Bienlein and G. Schlosser, "Backscattering of electrons in energy region from 60 to 100 keV," *Z. Phys.*, vol. 174, no. 1, pp. 91–101, 1963.

[14] H. E. Bishop, Ph.D. dissertation, Univ. of Cambridge, Cambridge, U.K., 1963.

[15] H. E. Bishop, "Some electron backscattering measurements for solid targets," in *Proc. 4th Int. Conf. X-ray Optics and X-ray Microanalysis*, 1965, pp. 153–158.

[16] H. E. Bishop, "Some electron backscattering measurements for solid targets," *Optique des rayons X et Microanalyse*, p. 153, 1966.

[17] L. M. Bojarshinov, *At. Energy SSSR*, vol. 21, no. 42, 1966.

[18] R. Bongeler *et al.*, "Electron-specimen interactions in low-voltage scanning electron microscopy," *Scanning*, vol. 15, no. 1, pp. 1–18, 1993.

[19] I. M. Bronshtein and B. S. Fraiman, "Inelastic scattering of electrons and secondary electron emission from certain metals and semiconductors," *Phys. Solid State*, vol. 3, no. 6, pp. 1188–1196, 1961.

[20] I. M. Bronshtein and B. S. Fraiman, "Secondary electron emission," *Nauka, Moscow*, p. 340, 1969.

[21] A. Carrington and H. W. Nicholson, "Electron backscattering from a silicon detector from 0.182 to 0.579 MeV," *Nucl. Instrum. Methods Phys. Res. A*, vol. 248, no. 2, pp. 425–428, 1986.

[22] V. E. Cosslett and R. N. Thomas, "Multiple scattering of 5-30 keV electrons in evaporated metal films lll: Backscattering and absorption," *Brit. J. Appl. Phys.*, vol. 16, no. 6, pp. 779–796, 1965.

[23] H. Drescher, L. Reimer, and H. Seidel, "Backscattering and secondary electron emission of 10-100 keV electrons in scanning electron microscopy," *Z. Angew. Phys.*, vol. 29, no. 6, pp. 331–336, 1970.

[24] P. J. Ebert, A. F. Lauzon, and E. M. Lent, "Transmission and backscattering of 4.0-to 12.0-MeV electrons," *Phys. Rev.*, vol. 183, no. 2, pp. 422–430, 1969.

[25] L. Frank, R. Stekly, M. Zadrazil, M. M. El-Gomati, and I. Mullerova, "Electron backscattering from real and in-situ treated surfaces," *Microchim. Acta.*, vol. 132, no. 2 4, pp. 179–188, 2000.

[26] P. Y. Glazunov and V. G. Guglya, "Reflection of monoenergetic electrons ranging within 600 to 1200 keV from certain metals + from graphite," *Dokl. Akad. Nauk. SSSR+*, vol. 159, no. 3, p. 632, 1964.

[27] M. M. El-Gomati and A. M. D. Assa'd, "On the measurement of the backscattering coefficient for low energy electrons," *Michochim. Acta*, vol. 15, pp. 325–331, 1998.

[28] D. Harder and H. Ferbert, "Fast electron backscattering in the energy region 8 to 22 MeV," *Phys. Lett.*, vol. 9, no. 3, pp. 233–234, 1964.

[29] D. Harder and G. Poschet, "Transmission and range of fast electrons in the energy region 4 to 30 MeV," *Phys. Lett.*, vol. 24B, no. 10, pp. 519–521, 1967.

[30] D. Harder and L. Metzger, "Check of electron backscattering coefficients at 10 and 20 MeV," *Z. Naturforsch*, vol. 23a, p. 1675, 1968.

[31] K. F. J. Heinrich, "Electron probe micoanalysis by specimen current measurement," *X-ray Opt. Microanal.*, pp. 159–167, 1966.

[32] D. H. H. Hoffman *et al., Z. Phys.*, vol. A293, p. 187, 1979.

[33] J. E. Holliday and E. J. Sternglass, "Backscattering of 5-20 keV electrons from insulators and metals," *J. Appl. Phys.*, vol. 28, no. 10, pp. 1189–1193, 1957.

[34] H. J. Hunger and L. Kuchler, "Measurements of the electron backscattering coefficient for quantitative EPMA in the energy range of 4 to 40 keV," *Phys. Status Solidi A*, vol. 56, no. 1, pp. K45–K48, 1979.

[35] J. Jakschik and K. P. Jungst, "Electron back-scattering at 0.25 MeV and 0.5 MeV on aluminum," *Nucl. Instrum. Methods*, vol. 79, no. 2, pp. 240–244, 1970.

[36] V. H. Kanter, "Backscatter of electrons in the 10-100 keV energy region," *Ann. Phys.-Berlin*, vol. 6, pp. 144–166, 1957.

[37] H. Kanter, "Contribution of backscattered electrons to secondary electron formation," *Phys. Rev.*, vol. 121, no. 3, pp. 681–684, 1961.

[38] Y. D. Kornyushkin, N. A. Muzil, V. V. Trunev, and I. G. Stoyanova, "Backscattering of electrons with energy 4-30 keV from the Al-Cu-Al system," *Sov. Phys. Tech. Phys.*, vol. 26, p. 624, 1981.

[39] H. Kulenkampff and K. Ruttiger, "Energy and angular distribution of backscattered electrons," *Z. Phys.*, vol. 137, pp. 426–434, 1954.

[40] A. Liuzzi and J. G. Martin, "Electron backscatter fractions for aluminum and copper," *Trans. Amer. Nucl. Soc.*, vol. 32, pp. 253–254, 1979.

[41] G. R. Massoumi, W. N. Lennard, H. H. Jorch, and P. J. Schultz, "Backscattering of 35 keV electrons from thick targets," in *Proc. 4th Int. Workshop Slow- Positron Beam Techniques for Solids and Surfaces*, 1991, vol. 218, no. 1, pp. 39–43.

[42] B. L. Miller, "A multiple-cavity linear electron accelerator," *Rev. Sci. Instrum.*, vol. 23, no. 8, pp. 401–408, 1952.

[43] D. A. Moncrieff and P. R. Barker, "Secondary electron emission in the scanning electron microscope," *Scanning*, vol. 1, no. 3, pp. 195–197, 1978.

[44] Y. Nakai, K. Matsuda, T. Takagaki, and K. Kimura, "Energy dissipation of electron beams in matter. IV. Measurement of electron energy dissipation in polystyrene by using ionization chamber," *Ann. Rep. Japn. Ass. Rad. Res. Polym.*, vol. 6, no. 7, 1964.

[45] G. Neubert and S. Rogaschewski, "Backscattering coefficient measurements of 15 to 60 keV electrons for solids at various angles of incidence," *Phys. Status Solidi A*, vol. 59, no. 1, pp. 35–41, 1980.

[46] P. Palluel, "Backscattered components of electron secondary emission from metals," *C. R. Acad. Sci.*, vol. 224, pp. 1492–1494, 1947.

[47] J. Philibert and E. Weinryb, "The use of specimen current in electron-probe microanalysis," *X-ray Opt. X-ray Microan.*, pp. 451–476, 1963.

[48] L. Reimer and C. Tollkamp, "Measuring the backscattering coefficient and secondary electron yield inside a scanning electron microscope," *Scanning*, vol. 3, no. 1, pp. 35–39, 1980.
[49] D. H. Rester and J. H. Derrickson, "Electron backscatter measuerments for perpendicular and non-perpendicular incidence at 1.0 MeV bombarding energy," *Nucl. Instrum. Methods*, vol. 86, no. 2, pp. 261–267, 1970.
[50] J. Saldick and A. O. Allen, "Backscattering from targets of low atomic number bombarded with 1-2 MeV electrons," *J. Chem. Phys.*, vol. 22, no. 10, p. 1777, 1954.
[51] J. Schou and H. Sorensen, "The penetration depth of 0.5-3-keV electrons in solid hydrogen and deuterium," *J. Appl. Phys.*, vol. 49, no. 2, pp. 816–821, 1978.
[52] K. S. Sim and N. Kamel, "A new setup and measurement technique for SNR," *Phys. Lett. A*, 2003.
[53] K. S. Sim and J. D. White, "New technique for in-situ measurement of backscattered and secondary electron yields for the calculation of signal-to-noise ratio in a SEM," *J. Microsc.*, vol. 217, no. 3, pp. 235–240, 2005.
[54] P. Sommerkamp, "Electron backscattering from tantalum and nickel as a factor in the power efficiency of electron- beam melting," *Z. Angew. Phys.*, vol. 28, no. 4, pp. 220–232, 1970.
[55] G. Soum, A. Mousselli, F. Arnal, and P. Verdier, "Etude de la transmission et de la rétrodiffusion d'électrons d'energie 0,05 a 3 MeV dans le domaine de la diffusion multiple," *Rev. Phys. Appl.*, vol. 22, no. 10, pp. 1189–1209, 1987.
[56] E. J. Sternglass, "Backscattering of kilovolt electrons from solids," *Phys. Rev.*, vol. 95, no. 2, pp. 345–358, 1954.
[57] T. Tabata, "Backscattering of electrons from 3.2 to 14 MeV," *Phys. Rev.*, vol. 162, no. 2, pp. 336–347, 1967.
[58] S. Thomas and E. B. Pattinson, "Automatic measurement of secondary electron emission characteristics of TaC, TiC and ZrC," *J. Phys. D*, vol. 2, no. 11, p. 1539, 1969.
[59] S. Thomas and E. B. Pattinson, "Range of electrons and contribution of back-scattered electrons in secondary production in aluminium'," *J. Phys. D*, vol. 3, no. 3, pp. 349–357, 1970.
[60] P. Verdier and F. Arnal, "Phénomène de rétrodiffusion d'électrons monocinétiques," *C. R. Acad. Sci. Ser. B*, vol. 267, pp. 1443–1446, 1968.
[61] M. Waldschmidt and S. Wittig, "Backscattering and bremsstrahlung of electrons in a silicon detector," *Nucl. Instrum. Methods*, vol. 64, no. 2, pp. 189–191, 1968.
[62] E. Weinryb and J. Philibert, "Measurement of the coefficient of backward scattering for electrons of 5 to 30 keV," *C. R. Acad. Sci.*, vol. 258, pp. 4535–4538, 1964.
[63] D. B. Wittry, "Secondary electron emission in the electron probe," in *Proc. 4th Int. Conf. on X-ray Optics and X-ray Microanalysis*, 1966, pp. 168–180.
[64] K. A. Wrigh and J. G. Trump, "Back-scattering of magavolt electrons from thick targets," *J. Appl. Phys.*, vol. 33, no. 2, pp. 687–690, 1962.
[65] R. K. Yadav and R. Shanker, "Backscattering of 8-28 keV electrons from a thick Al, Ti, Ag and Pt targets," *J. Electron Spectrosc.*, vol. 151, no. 1, pp. 71–77, 2006.
[66] R. K. Yadav and R. Shanker, "Contribution of backscattered electrons to the total electron yield produced in collisions of 8-28 keV electrons with tungsten," *Pramana*, vol. 68, no. 3, pp. 507–515, 2007.
[67] H. Hirayama, Y. Namito, A. F. Bielajew, S. J. Wilderman, and W. R. Nelson, The EGS5 code system, Stanford, CA, USA, SLAC-R-730 Rep. and KEK Rep. 2005-8, 2007.
[68] I. Kawrakow, E. Mainegra-Hing, D. W. O. Rogers, F. Tessier, and B. R. B. Walters, "The EGSnrc Code System: Monte Carlo simulation of electron and photon transport," NRCC PIRS-701, 6th printing, 2011.
[69] A. Ferrari, P. R. Sala, A. Fassó, and J. Ranft, FLUKA: A multi-particle transport code Geneva, CERN-2005-10 2005 INFN/TC_05/11, SLAC-R-773 Rep., 2005.
[70] B. C. Franke, R. P. Kensek, and T. W. Laub, ITS Version 6: The Integrated TIGER Series of Coupled Electron/Photon Monte Carlo Transport Codes, Sandia National Laboratories, Albuquerque, NM, USA, SAND2008-3331 Rep., 2008.
[71] T. Goorley *et al.*, Features of MCNP6, Los Alamos, NM, USA, LA-UR-13-28114 Rep., 2013.
[72] J. Baro, J. Sempau, J. M. Fernandez-Varea, and F. Salvat, "PENELOPE, an algorithm for Monte Carlo simulation of the penetration and energy loss of electrons and positrons in matter," *Nucl. Instrum. Methods Phys. Res. B*, vol. 100, no. 1, pp. 31–46, 1995.
[73] M. J. Berger and R. Wang, "Multiple-scattering angular deflections and energy-loss straggling," *Monte Carlo Transport of Electrons and Photons*, vol. 38, pp. 21–56, 1988, Ettore Majorana International Science Series.
[74] S. Goudsmit and J. L. Saunderson, "Multiple scattering of electrons," *Phys. Rev.*, vol. 58, pp. 24–29, 1940.
[75] S. Goudsmit and J. L. Saunderson, "Multiple scattering of electrons. II," *Phys. Rev.*, vol. 58, pp. 36–42, 1940.
[76] G. Molière, "Theorie der streuung schneller geladener Teilchen I: Einzelstreuung am abgeschirmten coulomb-feld," *Z. Naturforsch. A*, vol. 2, pp. 133–145, 1947.
[77] H. W. Lewis, "Multiple scattering in an infinite medium," *Phys. Rev.*, vol. 78, pp. 526–529, 1950.
[78] G. Wentzel, "Zwei bemerkungen uber die zerstreuung korpuskularer strahlen als beugungserscheinung," *Z. Physik*, vol. 40, no. 8, pp. 590–593, 1926.
[79] J. M. Fernandez-Varea, R. Mayol, J. Baró, and F. Salvat, "On the theory and simulation of multiple elastic scattering of electrons," *Nucl. Instrum. Methods Phys. Res. B*, vol. 73, p. 447473, 1993.
[80] A. Ferrari, P. R. Sala, R. Guaraldi, and F. Padoani, "An improved multiple scattering model for charged particle transport," *Nucl. Instrum. Methods Phys. Res. B*, vol. 71, no. 4, pp. 412–426, 1992.
[81] Y. Kirihara, Y. Namito, H. Iwase, and H. Hirayama, "Monte Carlo simulation of tabatas electron backscattering experiments," *Nucl. Instrum. Methods Phys. Res. B*, vol. 268, pp. 2384–2390, 2010.
[82] E. S. M. Ali and D. W. O. Rogers, "Benchmarking EGSnrc in the kilovoltage energy range against experimental measurements of charged particle backscatter coefficients," *Phys. Med. Biol.*, vol. 53, p. 15271543, 2008.
[83] R. Jeraj, P. J. Keall, and P. M. Ostwald, "Comparisons between MCNP, EGS4 and experiment for clinical electron beams," *Phys. Med. Biol.*, vol. 44, p. 705717, 1999.
[84] J. F. Briesmeister, MCNP4B MCNP- A General Monte Carlo N-particle Transport Code, Version 4B, Los Alamos, NM, USA, LA-12625-M Rep., 1997.
[85] W. R. Nelson, H. Hirayama, and D. W. O. Rogers, The EGS4 Code System, Stanford, CA, USA, SLAC-265 Rep., 1985.
[86] A. Fassó, A. Ferrari, and P. R. Sala, "Electron photon transport in FLUKA: Status," in *Proc. Monte Carlo Conf.*, Lisbon, Portugal, 2000, pp. 159–164.
[87] L. Urban, "Multiple scattering model in Geant4," CERN-OPEN-2002-070, Geneva, Switzerland, 2002.
[88] L. Urban, "A model for multiple scattering in Geant4," in *Proc. Monte Carlo Method: Versatility Unbounded in a Dynamic Computing World*, La Grange Park, IL, USA, 2005, American Nuclear Society, CD-ROM.
[89] L. Urban, "A model of multiple scattering in Geant4," CERN-OPEN-2006-077 Geneva, Switzerland, 2006.
[90] S. Elles, V. N. Ivanchenko, M. Maire, and L. Urban, "Geant4 and fano cavity: Where are we," *J. Phys., Conf. Series*, vol. 102, pp. 1–8, 2009.
[91] J. Apostolakis *et al.*, "The performance of the Geant4 standard EM package for LHC and other applications," *J. Phys., Conf. Ser.*, vol. 119, p. 032004, 2008.
[92] V. N. Ivanchenko *et al.*, "Geant4 models for simulation of multiple scattering," *J. Phys., Conf. Ser.*, vol. 219, p. 032045, 2010.
[93] O. Kadri, V. Ivanchenkob, F. Gharbi, and A. Trabelsi, "Incorporation of the goudsmit-saunderson electron transport theory in the Geant4 Monte Carlo code," *Nucl. Instrum. Methods Phys. Res. B*, vol. 267, no. 23–24, pp. 3624–3632, 2009.
[94] Geant4 10.1, "User's guide: For application developers," [Online]. Available: http://geant4.web.cern.ch/geant4/UserDocumentation/UsersGuides/ForApplicationDeveloper/html/index.html
[95] V. N. Ivanchenko *et al.*, "Recent improvements in Geant4 electromagnetic physics models and interfaces," *Progr. Nucl. Sci. Technol.*, vol. 2, pp. 898–903, 2011.
[96] V. N. Ivanchenko *et al.*, "Geant4 electromagnetic physics: Improving simulation performance and accuracy," in *Proc. Joint Int. Conf. Supercomputing in Nucl. Appl. and Monte Carlo 2013* [Online]. Available: http://sna-and-mc-2013-proceedings.edpsciences.org/articles/snamc/abs/2014/01/snamc2013_03101/snamc2013_03101.html
[97] S. T. Perkins, D. E. Cullen, and S. M. Seltzer, Tables and Graphs of Electron-Interaction Cross Sections from 10 eV to 100 GeV Derived from the LLNL Evaluated Electron Data Library (EEDL), UCRL-50400 vol. 31, 1997.
[98] D. Cullen *et al.*, EPDL97, the Evaluated Photon Data Library, Lawrence Livermore National Laboratory, Rep. UCRL-50400, 1997, vol. 6, Rev. 5.

[99] M. J. Berger, J. S. Coursey, M. A. Zucker, and J. Chang, Stopping-Power and Range Tables for Electrons, Protons, and Helium Ions [Online]. Available: http://www.nist.gov/pml/data/star/

[100] G. A. P. Cirrone *et al.*, "A goodness-of-fit statistical toolkit," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 5, pp. 2056–2063, Oct. 2004.

[101] B. Mascialino, A. Pfeiffer, M. G. Pia, A. Ribon, and P. Viarengo, "New developments of the goodness-of-fit statistical toolkit," *IEEE Trans. Nucl. Sci.*, vol. 53, no. 6, pp. 3834–3841, Dec. 2006.

[102] R Core Team, "R: A language and environment for statistical computing," R Foundation for Statistical Computing ISBN 3-900051-07-0, Vienna, Austria, 2012 [Online]. Available: http://www.R-project.org/

[103] R. K. Bock and W. Krischer*, The Data Analysis BriefBook*. Berlin, Germany: Springer, 1998.

[104] T. W. Anderson and D. A. Darling, "Asymptotic theory of certain goodness of fit criteria based on stochastic processes," *Ann. Math. Stat.*, vol. 23, pp. 193–212, 1952.

[105] T. W. Anderson and D. A. Darling, "A test of goodness of fit," *J. Amer. Stat. Ass.*, vol. 49, pp. 765–769, 1954.

[106] H. Cramér, "On the composition of elementary errors. Second paper: Statistical applications," *Skand. Aktuarietidskr*, vol. 11, pp. 13–141–74–180, 1928.

[107] R. von Mises and F. Duticke*,* in *Wahrscheinlichkeitsrechnung und ihre Anwendung in der Statistik und theoretischen Physik*, Leipzig, Germany, 1931.

[108] A. N. Kolmogorov, "Sulla determinazione empirica di una legge di distribuzione," *Gior. Ist. Ital. Attuari*, vol. 4, pp. 83–91, 1933.

[109] N. V. Smirnov, "On the estimation of the discrepancy between empirical curves of distributions for two independent samples," *Bull. Math. Univ. Moscow*, 1939.

[110] G. S. Watson, "Goodness-of-fit tests on a circle," *Biometrika*, vol. 48, pp. 109–114, 1961.

[111] K. Pearson, "On the χ^2 test of goodness of fit," *Biometrika*, vol. 14, no. 1–-2, pp. 186–191, 1922.

[112] R. A. Fisher, "On the interpretation of χ^2 from contingency tables, and the calculation of P," *J. Royal Stat. Soc.*, vol. 85, no. 1, pp. 87–94, 1922.

[113] A. Agresti, "A survey of exact inference for contingency tables," *Stat. Sci.*, vol. 7, pp. 131–153, 1992.

[114] G. A. Barnard, "Significance tests for 2 2 tables," *Biometrika*, vol. 34, pp. 123–138, 1947.

[115] R. D. Boschloo, "Raised conditional level of significance for the 2×2-table when testing the equality of two probabilities," *Stat. Neerlandica*, vol. 24, pp. 1–35, 1970.

[116] S. Suissa and J. J. Shuster, "Exact unconditional sample sizes for the 2×2 Binomial Trial," *J. Royal Stati. Soc., Ser. A*, vol. 148, pp. 317–327, 1985.

[117] A. Martin Andres and A. Silva Mato, "Choosing the optimal unconditioned test for comparing two independent proportions," *Comp. Stat. Data Anal.*, vol. 17, pp. 555–574, 1994.

[118] A. Martin Andres, A. Silva Mato, J. M. Tapia Garcia, and M. J. Sanchez Quevedo, "Comparing the asymptotic power of exact tests in 2 2 tables," *Comp. Stat. Data Anal.*, vol. 47, pp. 745–756, 2004.

[119] Q. McNemar, "Note on the sampling error of the difference between correlated proportions or percentages," *Psychometrika*, vol. 12, pp. 153–157, 1947.

[120] B. M. Bennett and R. E. Underwood, "Note: On mcnemar's test for the 2×2 table and its power function," *Biometrics*, vol. 26, no. 2, pp. 339–343, 1970.

[121] F. Yates, "Contingency table involving small numbers and the χ^2 test," *J. Royal Stat. Soc. Suppl.*, vol. 1, pp. 217–235, 1934.

[122] K. J. Lui, "Notes on testing equality in dichotomous data with matched pairs," *Biometr. J*, vol. 43, no. 3, p. 313321, 2001.

[123] J. Apostolakis, S. Giani, M. Maire, P. Nieminen, M. G. Pia, and L. Urban, "Geant4 low energy electromagnetic models for electrons and photons," Frascati, Italy, 1999.

[124] S. Chauvie, G. Depaola, V. Ivanchenko, F. Longo, P. Nieminen, and M. G. Pia, "Geant4 low energy electromagnetic physics," in *Proc. Computing in High Energy and Nuclear Physics*, Beijing, China, 2001, pp. 337–340.

[125] S. Chauvie *et al.*, "Geant4 low energy electromagnetic physics," in *Proc. IEEE Nuclear Sciece Symp. Conf. Rec.*, 2004, pp. 1881–1885.

[126] V. N. Ivanchenko, M. Maire, and L. Urban, "Geant4 standard electromagnetic package for HEP applications," in *Proc. IEEE Nuclear Science Symp. Conf. Rec.*, 2004, pp. N33–179.

[127] M. Batic, G. Hoff, M. G. Pia, and P. Saracco, "Photon elastic scattering simulation: Validation and improvements to Geant4," *IEEE Trans. Nucl. Sci.*, vol. 59, no. 4, pp. 1636–1664, Aug. 2012.

[128] T. Basaglia *et al.*, "Physics methods for the simulation of photoionisation," in *Proc. IEEE Nuclear Science Symp. Conf. Rec.*, 2013, pp. 1–8.

[129] G. Weidenspointner *et al.*, "Validation of compton scattering Monte Carlo simulation models," in *Proc. IEEE Nuclear Science Symp. Conf. Rec.*, 2013, pp. 1–5.

[130] M. Batic *et al.*, "Photons revisited," in *Proc. SNA + MC 2013 - Joint Int. Conf. Supercomputing in Nuclear Applications + Monte Carlo* [Online]. Available: http://dx.doi.org/10.1051/snamc/201402104

[131] D. J. Best and D. E. Roberts, "Algorithm AS 89: The upper tail probabilities of spearman's rho," *Appl. Stat.*, vol. 24, p. 377379, 1975.

[132] M. Hollander and D. A. Wolfe*, Nonparametric Statistical Methods*. New York, NY, USA: Wiley, 1973, p. 185194.

[133] C. Spearman, "The proof and measurement of association between two things," *Amer. J. Psychol.*, vol. 15, pp. 72–101, 1904.

[134] M. Kendall, "A new measure of rank correlation," *Biometrika*, vol. 30, no. 1–2, pp. 81–89, 1938.

[135] K. Pearson, "Notes on regression and inheritance in the case of two parents," in *Proc. Royal Soc. London*, 1895, vol. 58, pp. 240–242.

[136] G. Booch, I. Jacobson, and J. Rumbaugh*, The Unified Software Development Process*. Reading, MA, USA: Addison-Wesley, 1999.

[137] CMMI Product Team, CMMI for Development, Version 1.3, Software Engineering Institute Tech. Rep. CMU/SEI-2010-TR-033, Carnegie Mellon Univ., Pittsburgh, PA, USA, 2010.

[138] International Standards Organization (ISO) and International Electro technical Commission (IEC), ISO/IEC 15504 series, Geneva, Switzerland, 2004.