Evolutions in Photoelectric Cross Section Calculations and Their Validation

Tullio Basaglia, Maria Grazia Pia, and Paolo Saracco

***Abstract*— This paper updates and complements a previously published evaluation of computational methods for total and partial cross sections, relevant to modeling the photoelectric effect in Monte Carlo particle transport. It examines calculation methods that have become available since the publication of the previous paper, some of which claim improvements over previous calculations; it tests them with statistical methods against the same sample of experimental data collected for the previous evaluation. No statistically significant improvements are observed with respect to the calculation method identified in the previous paper as the state of the art for the intended purpose, encoded in the EPDL97 data library. Some of the more recent computational methods exhibit significantly lower capability to reproduce experimental measurements than the existing alternatives.**

***Index Terms*— Cross sections, photoionization, Monte Carlo simulation.**

I. Introduction

PHOTOIONIZATION has attracted theoretical and experimental interest for many decades, which is reflected in the continuing expansion of theoretical and empirical methods to calculate the cross sections associated with this process.

A previous publication [1] examined several total and partial photoionization cross section calculations with respect to a wide collection of experimental measurements to assess the state of the art for modeling the photoelectric effect in Monte Carlo particle transport. The validation tests documented in [1] identified the tabulations in the Evaluated Photon Data Library (EPDL), 1997 version [2], based on Scofield's 1973 calculations [3], as the calculation method best reproducing experimental cross sections. An overview of EPDL and Scofield's calculations is available in [1]; a recent review can be found in [4].

Other calculation methods have become available since completion of the tests documented in [1], either based on theoretical approaches or on empirical parameterizations: calculations by Sabbatucci and Salvat [5], modified tabulations in Penelope 2014 [6], a modified version of EPDL [7] released in 2018, new parameterizations à la Biggs-Lighthill [8]–[10], developed in the context of GeantV [11] and adopted in Geant4 [12]–[14].

Manuscript received December 17, 2019; accepted January 24, 2020. Date of publication February 3, 2020; date of current version March 13, 2020.

Tullio Basaglia is with CERN Scientific Information Service, CH-1211 Geneva, Switzerland (e-mail: tullio.basaglia@cern.ch).

Maria Grazia Pia and Paolo Saracco are with INFN (Istituto Nazionale di Fisica Nucleare) Sezione di Genova, I-16146 Genova, Italy (e-mail: mariagrazia.pia@ge.infn.it; paolo.saracco@ge.infn.it).

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

Digital Object Identifier 10.1109/TNS.2020.2971173

This paper updates the computational scenario documented in the previous one by evaluating the newly available modeling options with respect to the same collection of experimental data retrieved from the literature. The cross section calculations subject to test are summarized in Table I.

II. Evolutions in Photoelectric Cross Sections

Like [1], this paper focuses on calculations relevant to the simulation of the photoelectric effect in general-purpose Monte Carlo codes, which typically deal with single ionization of neutral atoms. The analysis concerns cross sections that are available as tabulations of precalculated values or can be expressed as simple analytical formulations, suitable for implementation in particle transport codes.

An overview of calculation methods of photoelectric cross sections is summarized in [1]. A few relevant developments that became available at a later stage are briefly discussed below.

It is worthwhile to stress that all the computational methods considered in this paper assume free atoms. This assumption is also adopted by all major general-purpose particle transport codes; it is liable to introduce inaccuracies for molecules and solids, especially near absorption edges, due to aggregation effects on the atomic potential and EXAFS (extended X-ray absorption fine structure). Therefore, some discrepancies observed with respect to experimental measurements are to be ascribed to the common underlying theoretical approach, rather than to specific deficiencies of any of the computational methods tested.

A. Theoretical Calculations

Sabbatucci and Salvat [5] applied first-order perturbation theory and Dirac one-electron wave functions for the Dirac-Hartree-Fock-Slater (DHFS) self-consistent potential of free atoms, which are also the basis of EPDL97 tabulations, to generate a database of atomic cross sections for the elements of the periodic table, including total cross sections and partial cross sections for the K shell and for L, M and N subshells with binding energies greater than approximately 50 eV.

The calculations are implemented in a code named *PHOTACS*. Additional effects are accounted for by a post-processing code, identified as *PHOTACS-PP*: the contribution from excitation to bound levels, the effect of the atomic level widths and a normalization correction. The possible computational configurations are listed in Table II.

TABLE I
CALCULATION METHODS OF PHOTOELECTRIC CROSS SECTIONS CONSIDERED IN THIS STUDY

Method	Identifier	Energy range		Z range		Shell
EPDL, 1997 version [2]	EPDL97	10 eV	100 GeV	1	100	all
EPDL as in EPICS2017 [7]	EPICS2017	10 eV	100 GeV	1	100	all
Biggs-Lighthill parameterization [10]	Biggs	10 eV	100 GeV	1	100	
Geant4 10.5 parameterization	Param6	∼5 keV	100 TeV	1	100	all
Penelope 2011 [23]	Pen2011	50 eV	1 GeV	1	99	K, L, M, N
Penelope 2014 [6]	Pen2014	50 eV	1 GeV	1	99	K, L, M, N
Sabbatucci-Salvat [5]	Sab[*option*]	Ionization threshold	10 GeV	1	99	all

TABLE II
OPTIONS FOR THE CALCULATION OF SABBATUCCI-SALVAT PHOTOELECTRIC CROSS SECTIONS

Identifier	Binding energy	Pratt correction	Level width correction	Excitation
SabCar	Carlson	no	no	no
SabWCar		no	yes	no
SabPCar		yes	no	no
SabWPCar		yes	yes	no
SabWil	Williams	no	no	no
SabWWil		no	yes	no
SabPWil		yes	no	no
SabWPWil		yes	yes	no
SabDef	Default	no	no	no
SabWDef		no	yes	no
SabPDef		yes	no	no
SabWPDef		yes	yes	no
SabECar	Carlson	no	no	yes
SabEWCar		no	yes	yes
SabEPCar		yes	no	yes
SabEWPCar		yes	yes	yes
SabEWil	Williams	no	no	yes
SabEWWil		no	yes	yes
SabEPWil		yes	no	yes
SabEWPWil		yes	yes	yes
SabEDef	Default	no	no	yes
SabEWDef		no	yes	yes
SabEPDef		yes	no	yes
SabEWPDef		yes	yes	yes

The normalization correction is based on a method devised by Pratt [15]–[17]; it aims to address possible inaccuracies of the adopted central potential by applying an energy-independent correction factor, derived from a more elaborate atomic model, to the subshell cross sections calculated from the DHFS potential. This correction exploits calculations according to the multi-configuration Dirac–Fock *MCDF* code by Desclaux [18], [19].

The database derived from [5] is used in the Penelope Monte Carlo code [5], [20].

B. Penelope 2014 Cross Sections

The latest version of Penelope publicly available at the time of writing this paper is Penelope 2014 [6], which was released in 2015.

According to the associated documentation [6], the photoelectric cross sections tabulated in Penelope 2014 derive from calculations performed by a computer code named *PHOTOABS*, which is based on the same Dirac-Hartree-Slater-Fock method as Scofield's 1973 calculations, but "implements more accurate numerical algorithms" and computes them over a denser grid of energies.

Penelope 2014 documentation [6] cites an associated reference to be published in 2014 as a source of further details. No pertinent paper appears to be published in 2014; [5], which was published in 2016, states that the database generated by the cross section calculations it describes was adopted in Penelope 2014.

The cross sections associated with Penelope 2014 were calculated for photon energies from the ionization threshold up to 1 GeV using the DHFS approach; they were extrapolated to energies higher than the calculation cut-off by means of Pratt's extrapolation formula [21], shifted in energy to have the absorption edges coinciding with the experimental subshell ionization energies given by Carlson [22] and renormalized using MCDF/DHFS density ratios. The normalisation correction is meant "to correct some inaccuracies of the DHFS wave functions".

The tabulations of photoelectric cross section distributed with previous versions of Penelope (up to Penelope 2011 [23]) were based on EPDL97.

C. EPICS2017 Cross Sections

A new version of EPDL was released in 2018 within ENDF/B-VIII.0 [24] and in a collection of data libraries called EPICS2017 [25]. It accounts for a modification of the binding energies previously encoded in EADL [26], [27]; according to [7], the photoelectric cross sections in EPICS2017 were obtained by modifying the EPDL97 [2] tabulations by smoothly interpolating or extrapolating the values close to absorption edges consistently with the modified binding energies.

In addition, photoelectric cross sections are tabulated in EPICS2017 according to a finer grid of energies than in EPDL97 to support linear interpolation of the data, while logarithmic interpolation was recommended for EPDL97.

Deficiencies in version control were observed [28] in the distribution of EPICS2017: different content has been distributed in various instances under the same identifier of EPICS2017, therefore it is impossible to identify EPICS2017 photoelectric cross sections univocally. The cross sections examined in this paper correspond to those available in ENDF format within ENDF/B-VIII.0 [24]; no further versions of these cross sections are mentioned in the ENDF/B-VIII.0

errata [29]. Different versions of photoelectric cross sections were released in ENDL format [25] in January and February 2018, both identically identified as EPICS2017; the latter appears equivalent to the tabulations included in ENDF/B-VIII.0.

D. Parameterized Cross Sections

A new parameterization of photoelectric cross sections is available in Geant4 version 10.5. It is inspired by Biggs-Lighthill [8]–[10] empirical expression of total photoelectric cross sections

$$\sigma_{ij} = \frac{A_{ij1}}{E} + \frac{A_{ij2}}{E^2} + \frac{A_{ij3}}{E^3} + \frac{A_{ij4}}{E^4} \quad (1)$$

for element i and energy range j. The modified parameterization, identified as "Param6" in Table I, is formulated as

$$\sigma_{ij} = \frac{A_{ij1}}{E} + \frac{A_{ij2}}{E^2} + \frac{A_{ij3}}{E^3} + \frac{A_{ij4}}{E^4} + \frac{A_{ij5}}{E^5} + \frac{A_{ij6}}{E^6} \quad (2)$$

over two energy ranges ("low" and "high"), whose limits depend on the atomic number. According to the Geant4 10.5 documentation [30], the coefficients appearing in (2) are calculated by fitting EPICS2014 [31] tabulations, which are identical to EPDL97 ones [28]. Although the bibliographical reference to the fitted tabulations in [30] refers to EPICS2017 at the time of writing this paper, it specifies that it was accessed on 26 October 2017, while EPICS2017 was released in January 2018; presumably, the online content of the reference was modified after the release of Geant4 10.5 documentation.

The low energy end of applicability of the parameterization is approximately 5 keV; below this limit, which varies according to the atomic number, the cross section is computed by interpolation of EPDL tabulations. Equation (2) can calculate total and shell cross sections.

This parameterized model is used for the simulation of the photoelectric effect by the Geant4 "low energy Livermore" model; it replaces a previous implementation, which calculated total and partial photoelectric cross sections by interpolating EPDL97 tabulations.

III. Validation Method

The validation [32] process applies the methodology described in [1]: comparison with experimental data to determine the compatibility of each calculation model with experiment, and categorical data analysis to identify the state of the art among the calculations subject to test.

The following subsections illustrate the additional computational methods evaluated in this paper and briefly summarize the main features of the validation method. The reader is referred to [1] for further details.

A. Cross Section Calculations

The cross section calculations examined in this paper include some relevant computational methods previously evaluated in [1], along with the recent calculations mentioned in Section II, to facilitate the comparison of the evolution of similar computational approaches.

TABLE III
Outcome of the χ^2 Test Over Total Cross Sections Above 100 eV

Model	Pass	Fail	Efficiency
EPDL97	49	15	0.77 ± 0.05
EPICS2017	48	16	0.75 ± 0.05
Biggs	48	16	0.75 ± 0.05
Pen2011	48	16	0.75 ± 0.05
Pen2014	27	37	0.42 ± 0.06
SabCar	44	20	0.69 ± 0.06
SabWCar	44	20	0.69 ± 0.06
SabPCar	27	37	0.42 ± 0.06
SabWPCar	27	37	0.42 ± 0.06
SabWil	43	21	0.67 ± 0.06
SabWWil	42	22	0.66 ± 0.06
SabPWil	29	35	0.45 ± 0.06
SabWPWil	29	35	0.45 ± 0.06
SabDef	45	19	0.70 ± 0.06
SabWDef	44	20	0.69 ± 0.06
SabPDef	28	36	0.44 ± 0.06
SabWPDef	28	36	0.44 ± 0.06
SabECar	44	20	0.69 ± 0.06
SabEWCar	44	20	0.69 ± 0.06
SabEPCar	27	37	0.42 ± 0.06
SabEWPCar	27	37	0.42 ± 0.06
SabEWil	43	21	0.67 ± 0.06
SabEWWil	43	21	0.67 ± 0.06
SabEPWil	29	35	0.45 ± 0.06
SabEWPWil	28	36	0.44 ± 0.06
SabEDef	45	19	0.70 ± 0.06
SabEWDef	44	20	0.69 ± 0.06
SabEPDef	28	36	0.44 ± 0.06
SabEWPDef	28	36	0.44 ± 0.06

A few configuration options can be applied to the cross sections calculated by Sabbatucci and Salvat [5] through different settings of the *PHOTACS-PP* code; they are listed in Table II with the corresponding identifiers used in this paper. They consist of incorporating the excitation to bound levels in the cross section calculation along with ionization, accounting for finite level width, and applying Pratt's renormalization correction; additionally, the *PHOTACS-PP* code offers the choice of two compilations of atomic binding energies of empirical origin, by Carlson [22] and Williams [33], respectively, as an alternative to the default ones. If atomic binding energies other than the default ones are chosen, the PHOTACS-PP code adjusts the energy corresponding to the cross section calculated with the default value to let the absorption edge coincide with the selected ionization energy. Details about the PHOTACS-PP code are documented in [5].

The parameterized cross section model described in Section II-D was refactored [34] consistently with the software design documented in [1] to facilitate the validation process. Tabulated cross sections were interpolated logarithmically or linearly, according to the recommendations of use of the respective data libraries.

B. Experimental Data

The experimental data samples used for the validation tests reported in this paper are the same as in [1]. Details can be found in Tables II and III of [1].

Test cases were defined exactly as in [1] for the comparison of total cross sections.

Test cases for the comparison of K shell data were defined for this analysis by grouping the experimental data mainly at fixed energies, i.e. assembling data samples consisting of the measurements performed by an experiment at a given energy for various target elements, to increase the power of the tests by means of larger sample sizes, while in the previous paper the test cases were mainly defined on the basis of the atomic number. However, it is worthwhile to stress that the experimental data set is the same as in the analysis documented in [1] and that the validation results obtained with either set of test cases are statistically consistent.

C. Data Analysis Method

The data analysis method is extensively documented in [1]; it is briefly summarized here to facilitate the comprehension of the results reported in Section IV.

The validation process encompasses two stages: the first tests the compatibility between the cross sections calculated by each simulation model and experimental data, while the second determines whether the models exhibit any significant difference in compatibility with experiment.

Both analyses employ statistical methods: the first stage applies the χ^2 goodness-of-fit test [35] to appraise the compatibility of calculated and experimental cross sections; the second applies a set of tests to contingency tables based on the outcome of the first stage. Exact tests (Fisher [36], Boschloo [37], Z-pooled [38] and Barnard [39] in the CSM approximation [40]) are used in the analysis of contingency tables; Pearson's χ^2 test [41] is also used, when the number of entries in the cells justifies its applicability (i.e., it is greater than 5 [42]).

The level of significance applied in all the tests to reject the null hypothesis is 0.01, unless otherwise specified.

The analysis has been designed to ensure that the tests have adequate power to observe sizeable effects of physical interest, given the available sample of experimental data. The need to retain adequate power hinders more detailed analyses, such as a thorough assessment of the incompatibility of the calculation methods with experiment as a function of the atomic number and of energy.

The robustness of the results reported in the following sections has been investigated with respect to several factors, e.g. with respect to measurements with different precision and to different formulations of the test statistic in both stages of the analysis; a few relevant details are discussed below.

The analysis reported in this paper used the R software system [43], version 3.6.1.

IV. Results

The main results of the validation process are reported in the following subsections for total and partial cross sections.

A. Total Cross Sections

Figs. 1–6 represent qualitative examples of how EPDL97 and some more recent computational methods reproduce

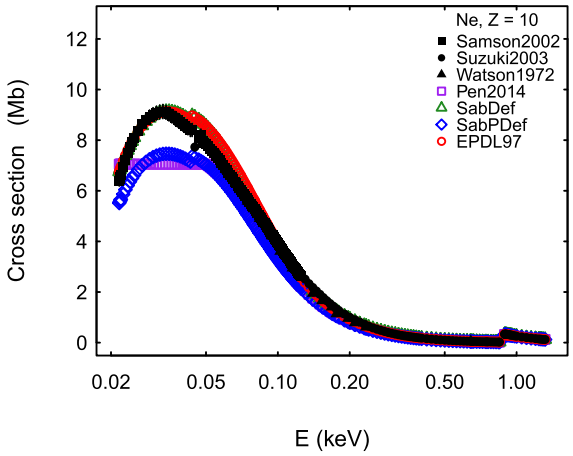

Fig. 1. Total photoelectric cross section for neon as a function of photon energy. The experimental references are listed in the bibliography of [1].

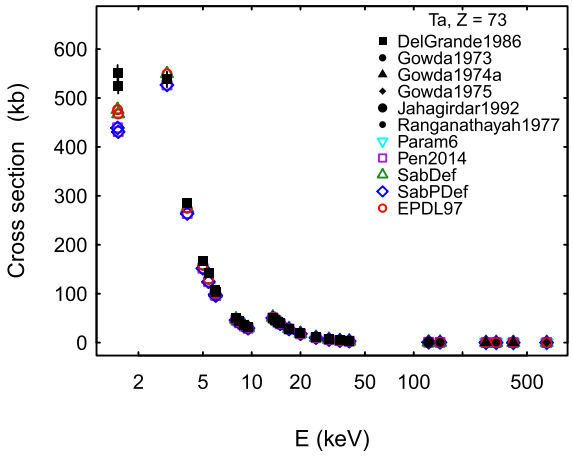

Fig. 2. Total photoelectric cross section for tantalum as a function of photon energy. The experimental references are listed in the bibliography of [1].

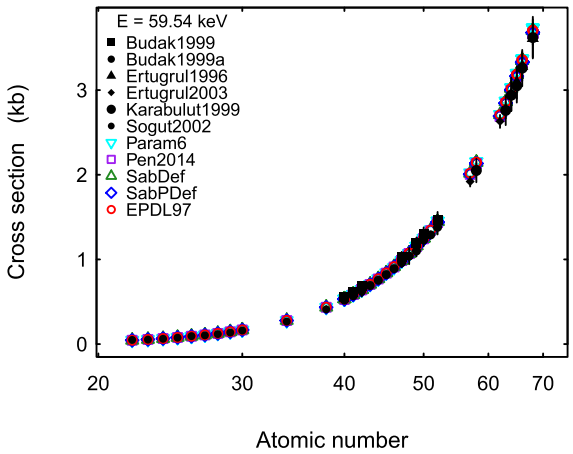

Fig. 3. Total photoelectric cross section at 59.54 keV as a function of the atomic number. The experimental references are in the bibliography of [1].

experimental total photoelectric cross sections. Error bars are omitted in Figs. 4 and 6 to facilitate the appraisal of relevant features in the plots, i.e. the visible discrepancies of data points involving Pratt's correction and the similarities concerning

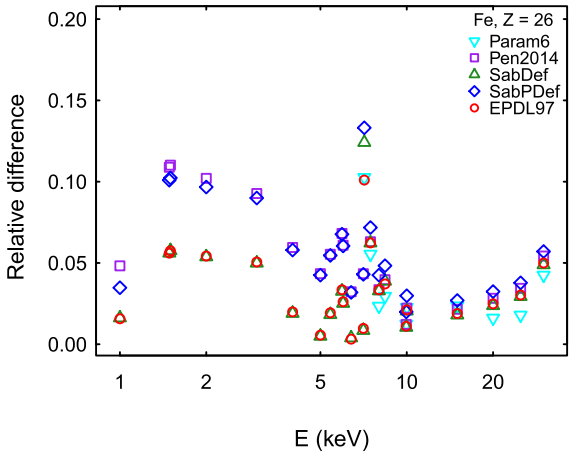

Fig. 4. Absolute relative difference between calculated and experimental total photoelectric cross section for iron as a function of photon energy. The experimental references are documented in Table II of [1]. Error bars are omitted to facilitate the appraisal of relevant features of the data points.

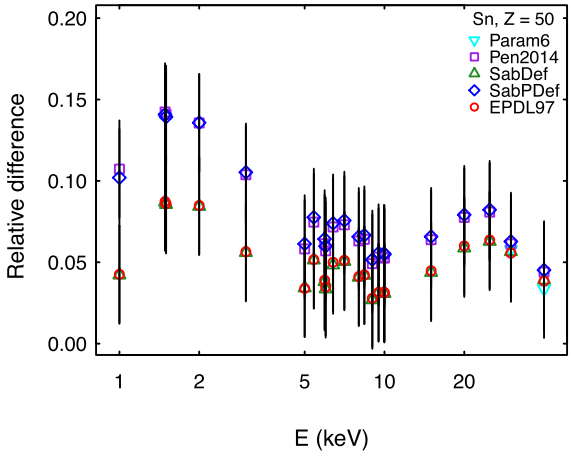

Fig. 5. Absolute relative difference between calculated and experimental total photoelectric cross section for tin as a function of photon energy. The experimental references are documented in Table II of [1].

other calculations. One can qualitatively distinguish some differences in the data plotted in Figs. 1–6: they mainly concern Sabbatucci-Salvat cross sections including the correction à la Pratt, while other data sets appear hardly distinguishable.

The investigation of the compatibility of calculated and measured total cross sections focuses on energies greater than 100 eV and above approximately 5 keV. The former energy range reflects the recommendation of applicability of EPICS2017 photon cross sections [7]; the latter corresponds to the domain of application of the parameterized model described in Section II-D adopted in Geant4 low energy electromagnetic package.

The outcome of the χ^2 test above 100 eV is summarized in Table III, where "fail" and "pass" identify the number of test cases where the hypothesis of compatibility between calculated and experimental data distributions is rejected or not rejected, respectively. For convenience, the table reports a variable named "efficiency", which represents the fraction of test cases where the null hypothesis is not rejected.

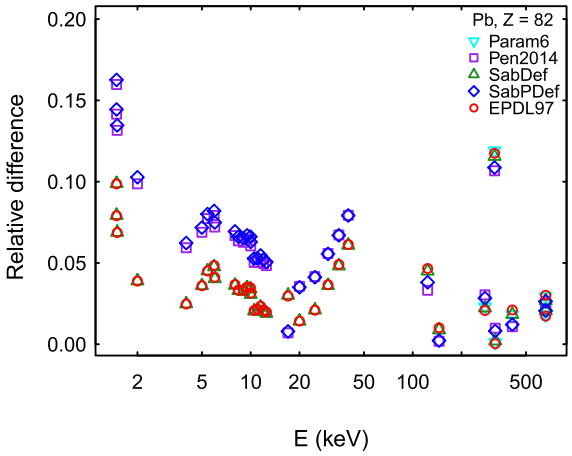

Fig. 6. Absolute relative difference between calculated and experimental total photoelectric cross section for lead as a function of photon energy. The experimental references are documented in Table II of [1]. Error bars are omitted to facilitate the appraisal of relevant features of the data points.

TABLE IV
OUTCOME OF THE χ^2 TEST OVER PARAMETERIZED TOTAL CROSS SECTIONS ABOVE APPROXIMATELY 5 keV

Model	Pass	Fail	Efficiency
EPDL97	28	3	0.90 ± 0.06
Biggs	28	3	0.90 ± 0.06
Param6	19	12	0.61 ± 0.08

Similarly, Table IV summarizes the relevant results for energies where the Geant4 parametrized model is applicable.

The experimental data sample encompasses measurements with different uncertainties; nevertheless, it is worthwhile to note that consistent outcome of the χ^2 test comparing experimental and calculated cross sections is observed over the whole range of data precision: for instance, no statistically significant dissimilarity is present in the "pass" and "fail" results of tests that involve experimental uncertainties in the first quartile of precision (relative uncertainties smaller than 2%) and in the last one (experimental uncertainties larger than 5%). This result suggests that the ability to appraise the capabilities of the computational models subject to evaluation is not significantly affected by the variable precision of the experimental sample, and that the presence of lower precision measurements is not expected to introduce a substantial bias in the categorical data analysis.

Qualitatively, above 100 eV one can observe higher efficiencies for EPDL97, EPICS2017, the original Biggs parameterization, Penelope 2011 and Sabbatucci-Salvat calculations not involving Pratt's correction. Substantial differences are visible in the compatibility with experiment of some categories of computational methods: between the original Biggs parameterization and the Geant4 6-parameter one, between Penelope 2011 and 2014 cross sections, and regarding the role of Pratt's correction in Sabbatucci-Salvat calculations.

Figs. 7–9 highlight some of the observed differences; they focus on the effect of options of cross section calculations based on the DHFS method (Fig. 7), on the performance of

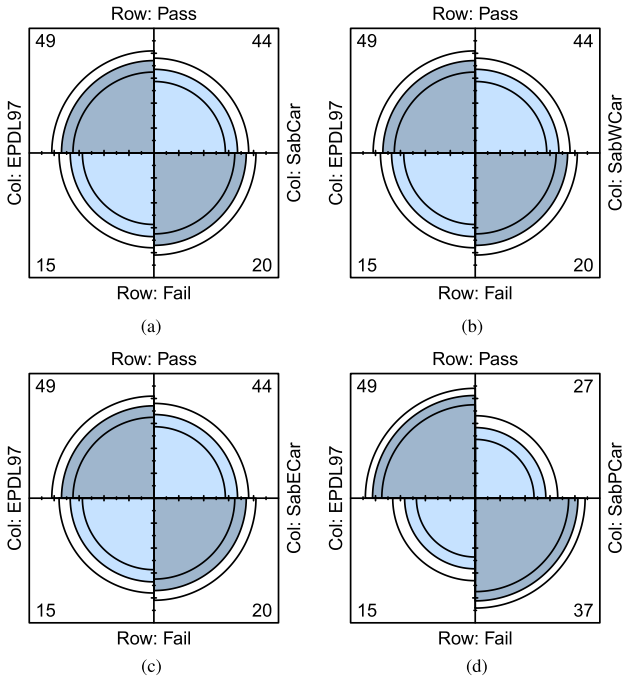

Fig. 7. Visual representation of the 2 by 2 table summarizing the compatibility with experiment of total photoelectric cross sections based on EPDL97 and of different options of Sabbatucci-Salvat calculations, above 100 eV. (a) Carlson binding energies. (b) Accounting for level width. (c) Including excitation to bound levels. (d) Including Pratt's correction.

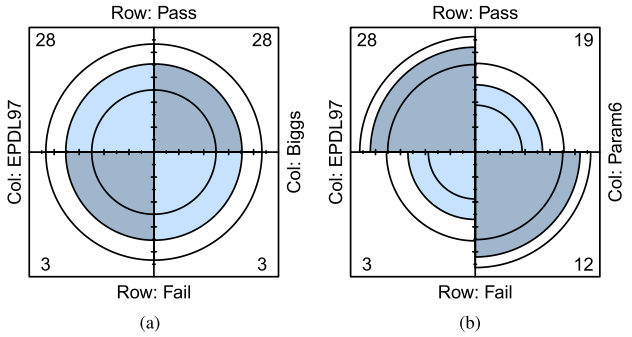

Fig. 8. Visual representation of the 2 by 2 tables summarizing the compatibility with experiment of total photoelectric cross sections calculated by EPDL97 and by parameterized models, above approximately 5 keV. (a) EPDL97—Biggs. (b) EPDL97—Param6.

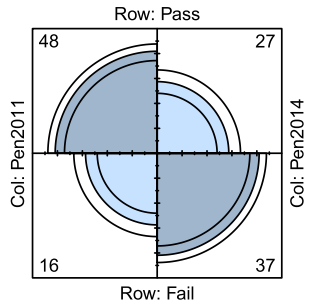

Fig. 9. Visual representation of the 2 by 2 table summarizing the compatibility with experiment of Penelope 2011 and 2014 total photoelectric cross sections.

the original Biggs-Lighthill parameterization and of the new one included in Geant4 10.5 (Fig. 8), and on comparing the compatibility with experiment of Penelope 2011 and 2014 (Fig. 9). The so-called *fourfold* plots provide a graphical representation of the contingency tables that summarize the results of the χ^2 test comparing experimental and calculated cross sections: the data are standardized so that both table margins are equal, while preserving the odds ratio; the area of each quarter circle is proportional to the standardized cell frequency in the corresponding table. An association (odds ratio different from 1) between the row and column variables of the table is indicated by the tendency of diagonally opposite cells in one direction to differ in size from those in the other direction. The innermost and outermost ring show the limits of a 99% confidence interval for the odds ratio; overlapping rings in adjacent quadrants indicate consistency with the null hypothesis of independence, i.e. the hypothesis that the different "pass" and "fail" observed in the two categories could arise from chance only. Details about fourfold plots can be found in textbooks such as [44], [45].

Fig. 7(a)–(c) represent contingency tables that compare the compatibility with experiment of EPDL97 with that of three variants of Sabbatucci-Salvat calculations using Carlson's binding energies: plain calculations (identified as SabCar), calculations accounting for finite level widths (SabWCar) and calculations accounting for excitation to bound states (SabECar). These plots visually convey the message that these three calculation configurations result in similar compatibility with experiment w.r.t. EPDL97. Fig. 7(d) compares the compatibility with experiment of EPDL97 and Sabbatucci-Salvat calculations using the same binding energies and applying the normalization correction à la Pratt (SabPCar); it qualitatively shows that this correction appears responsible for more pronounced discrepancy in compatibility with experiment w.r.t. EPDL97. A similar effect is observed in Fig. 9, which concerns Penelope 2014, using cross sections derived from Sabbatucci-Salvat calculations, and Penelope 2011, using EPDL97.

These indications are objectively quantified by applying the tests listed in Section III-C to the corresponding contingency tables. Since EPDL97 exhibits the largest efficiency among the computational methods subject to evaluation (as in [1]), the statistical analysis compares its performance with respect to experiment with that of other calculations.

The results of the statistical analysis of contingency tables are summarized in Table V for energies above 100 eV. The hypothesis of equivalent compatibility with experiment with respect to EPDL97 is rejected by all tests for Penelope 2014 cross sections and for Sabbatucci-Salvat calculations involving Pratt's normalization correction; it is not rejected for the other computational methods. Presumably, the difference in compatibility with experiment between Penelope 2011 and Penelope 2014 is related to the application of Pratt's correction in the latter. It is worthwhile to note that the null hypothesis is rejected with 0.001 significance level, which exceeds the significance set a priori in the context of this validation study.

The rejection occurs with 0.001 significance not only over the whole data sample; it occurs even over data samples involving experimental precisions above the median, i.e., encompassing measurements with $> 3\%$ relative uncertainty: that is, the difference in compatibility with experiment with respect to EPDL97 is observed also in the presence of larger experimental uncertainties. This observation confirms that the

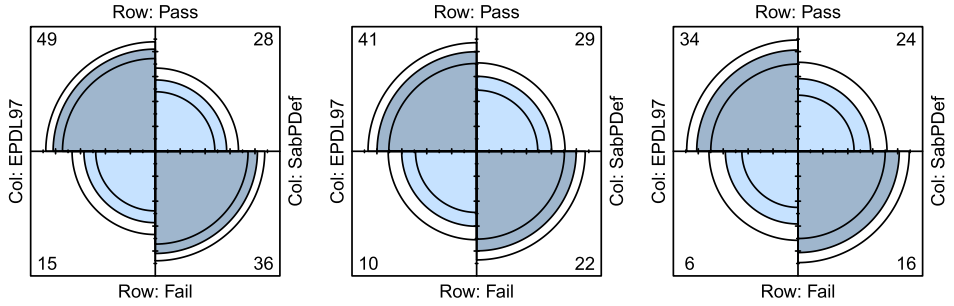

Fig. 10. Visual representation of 2 by 2 tables summarizing the compatibility with experiment of total photoelectric cross sections based on EPDL97 and on Sabbatucci-Salvat calculations including Pratt's correction, concerning different energies. (a) $E > 100$ eV. (b) $E > 250$ eV. (c) $E > 1$ keV.

TABLE V

P-VALUES OF TESTS COMPARING THE COMPATIBILITY WITH EXPERIMENT OF TOTAL CROSS SECTIONS OF EPDL97 AND OF OTHER COMPUTATIONAL MODELS, FOR ENERGIES ABOVE 100 eV

Models	**Fisher**	χ^2	**Boschloo**	**Z-pooled**	**CSM**
EPICS2017	1.000	0.837	1.000	0.905	0.997
Biggs	1.000	0.837	1.000	0.905	0.997
Pen2011	1.000	0.837	1.000	0.905	0.997
Pen2014	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabCar	0.428	0.321	0.338	0.529	0.405
SabWCar	0.428	0.321	0.338	0.529	0.405
SabPCar	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabWPCar	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabWil	0.326	0.238	0.250	0.250	0.308
SabWWil	0.242	0.172	0.185	0.243	0.226
SabPWil	0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabWPWil	0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabDef	0.549	0.423	0.448	0.531	0.601
SabWDef	0.428	0.321	0.338	0.529	0.405
SabPDef	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabWPDef	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabECar	0.428	0.321	0.338	0.529	0.405
SabEWCar	0.428	0.321	0.338	0.529	0.405
SabEPCar	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabEWPCar	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabEWil	0.326	0.238	0.250	0.250	0.308
SabEWWil	0.326	0.238	0.250	0.250	0.308
SabEPWil	0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabEWPWil	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabEDef	0.549	0.423	0.448	0.531	0.601
SabEWDef	0.428	0.321	0.338	0.529	0.405
SabEPDef	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
SabEWPDef	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001

presence of lower precision measurements in the data sample does not hinder the capability to identify statistically significant dissimilarities in the calculated cross sections.

Fig. 10 suggests that the effect of the normalization correction is more pronounced when lower energy data are involved. The plots concern data above 100 eV, 250 eV and 1 keV; they illustrate the compatibility with experiment of total photoelectric cross sections based on EPDL97 and on Sabbatucci-Salvat calculations including Pratt's correction, using the default atomic binding energies of the PHOTACS code. This effect is

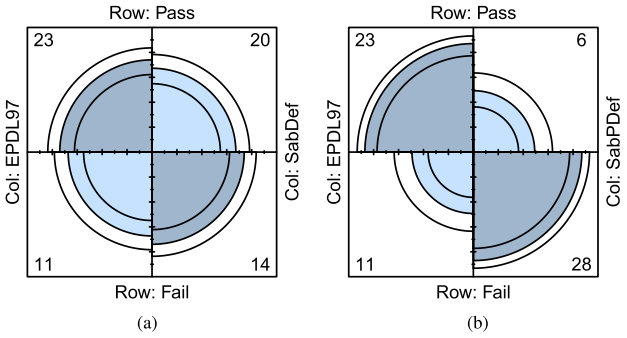

Fig. 11. Visual representation of the 2 by 2 tables summarizing the compatibility with experiment of total photoelectric cross sections calculated by EPDL97 and by Sabbatucci-Salvat, including or not including Pratt's rnormalization correction. The results concern data in the energy range between 100 eV and 1 keV. (a) Without normalization correction. (b) With normalization correction.

TABLE VI

P-VALUES OF TESTS COMPARING THE COMPATIBILITY WITH EXPERIMENT OF TOTAL CROSS SECTIONS BASED ON EPDL97 AND OF SABBATUCCI-SALVAT CALCULATIONS INCLUDING PRATT'S CORRECTION, OVER DIFFERENT ENERGIES

Energy	**Fisher**	χ^2	**Boschloo**	**Z-pooled**	**CSM**
E > 100 eV	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
E > 250 eV	0.018	0.010	0.013	0.013	0.011
E > 1 keV	0.023	0.012	0.014	0.013	0.015

clearly visible in Fig. 11, which concerns data between 100 eV and 1 keV.

For energies between 100 eV and 1 keV, the hypothesis of equivalent compatibility with experiment with respect to EPDL97 is rejected for all Sabbatucci-Salvat calculation options including the normalization correction, and for Penelope 2014; it is not rejected for the other computational methods. The statistical analysis of the contingency tables associated with Fig. 10, reported in Table VI, shows that the null hypothesis of equivalent compatibility with experiment is rejected with 0.01 significance for the sample involving cross sections down to 100 eV, while it is not rejected for the data samples at higher energies. Nevertheless, the null hypothesis is rejected with 0.05 significance over the higher energy data samples. This result suggests that the inadequacy

TABLE VII

P-VALUES OF TESTS COMPARING THE COMPATIBILITY WITH EXPERIMENT OF TOTAL CROSS SECTIONS OF EPDL97 AND OF PARAMETERIZED MODELS, FOR ENERGIES ABOVE APPROXIMATELY 5 keV

Models	Fisher	Boschloo	Z-pooled	CSM
Biggs	1	1	1	1
Param6	0.0159	0.0080	0.0080	0.0097

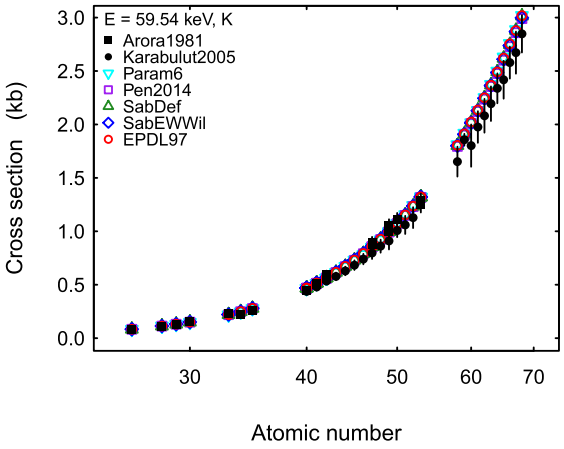

Fig. 12. K shell photoionization cross section at 59.54 keV as a function of the atomic number Z.

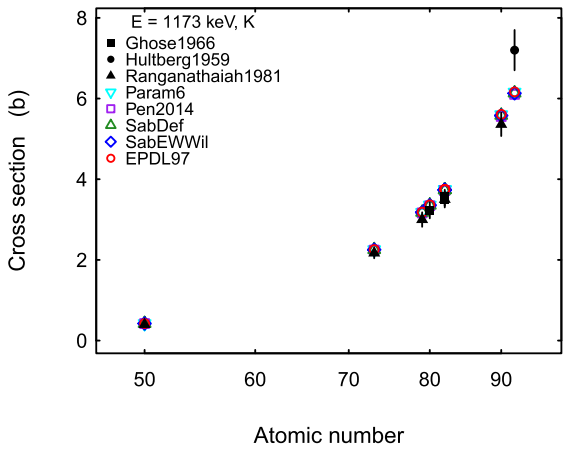

Fig. 13. K shell photoionization cross section at 1.173 MeV as a function of the atomic number Z.

TABLE VIII

OUTCOME OF THE χ^2 TEST OVER K SHELL CROSS SECTIONS ABOVE 100 eV

Model	Pass	Fail	Efficiency
EPDL97	23	3	0.88 ± 0.06
EPICS2017	24	2	0.92 ± 0.06
Pen2011	24	2	0.92 ± 0.06
Pen2014	23	3	0.88 ± 0.06
SabCar	23	3	0.88 ± 0.06
SabWCar	24	2	0.92 ± 0.06
SabPCar	23	3	0.88 ± 0.06
SabWPCar	24	2	0.92 ± 0.06
SabWil	23	3	0.88 ± 0.06
SabWWil	24	2	0.92 ± 0.06
SabPWil	23	3	0.88 ± 0.06
SabWPWil	24	2	0.92 ± 0.06
SabDef	23	3	0.88 ± 0.06
SabWDef	23	3	0.88 ± 0.06
SabPDef	23	3	0.88 ± 0.06
SabWPDef	23	3	0.88 ± 0.06
SabECar	23	3	0.88 ± 0.06
SabEWCar	24	2	0.92 ± 0.06
SabEPCar	23	3	0.88 ± 0.06
SabEWPCar	24	2	0.92 ± 0.06
SabEWil	23	3	0.88 ± 0.06
SabEWWil	25	1	0.96 ± 0.06
SabEPWil	23	3	0.88 ± 0.06
SabEWPWil	25	1	0.96 ± 0.06
SabEDef	23	3	0.88 ± 0.06
SabEWDef	23	3	0.88 ± 0.06
SabEPDef	23	3	0.88 ± 0.06
SabEWPDef	23	3	0.88 ± 0.06

of calculations involving the normalization correction to some extent persists even above 1 keV, which is conventionally considered a conservative lower limit for the reliability for photon transport in general-purpose Monte Carlo codes.

The analysis of cross section parameterizations is summarized in Table VII, which reports the results of the tests comparing the compatibility with experiment of EPDL97 total cross sections and of parameterized models above approximately 5 keV. The hypothesis of equivalent compatibility with experiment of EPDL97 and of Biggs-Lighthill original parameterization is not rejected by any of the tests applied to the corresponding contingency table. Regarding the parameterization included in Geant4 10.5, the null hypothesis is rejected with 0.01 significance by Boschloo [37], Z-pooled [38] and CSM-approximated Barnard [39] test, while it is not rejected by the Fisher test, which is known to be more conservative over 2×2 tables [46]. The rejection occurs even over the data samples involving experimental precisions above the median, i.e. encompassing measurements with $> 3\%$ relative uncertainty: that is, the difference in compatibility with experiment with respect to EPDL97 does not appear to be influenced by lower precision measurements.

B. Partial Cross Sections

Figs. 12 and 13 represent qualitative examples of how EPDL97 and some more recent computational methods reproduce experimental K shell photoelectric cross sections.

The outcome of the χ^2 test comparing calculated and experimental K shell cross sections is summarized in Table VIII for data samples above 100 eV and in Table IX for data above approximately 5 keV. All the models involved in the test qualitatively exhibit similar capabilities to reproduce experimental measurements.

The analysis of contingency tables does not identify any significant differences between the compatibility with experiment of K shell cross sections calculated by EPDL97 and by other computational methods. The results for data above 100 eV are summarized in Table X: the hypothesis of equivalent compatibility with experiment with respect to EPDL97 is not

TABLE IX

OUTCOME OF THE χ^2 TEST OVER RELEVANT K SHELL CROSS SECTIONS ABOVE APPROXIMATELY 5 keV

Model	Pass	Fail	Efficiency
EPDL97	22	3	0.88 ± 0.07
EPICS2017	23	2	0.92 ± 0.06
Param6	21	4	0.84 ± 0.07

TABLE X

P-VALUES OF TESTS COMPARING THE COMPATIBILITY WITH EXPERIMENT OF K SHELL CROSS SECTIONS OF EPDL97 AND OF OTHER COMPUTATIONAL MODELS, FOR ENERGIES ABOVE 100 eV

Models	Fisher	Boschloo	Z-pooled	CSM
EPICS2017	1.000	1.000	0.749	1.000
Pen2011	1.000	1.000	0.749	1.000
Pen2014	1.000	1.000	1.000	1.000
SabCar	1.000	1.000	1.000	1.000
SabWCar	1.000	1.000	0.749	1.000
SabPCar	1.000	1.000	1.000	1.000
SabWPCar	1.000	1.000	0.749	1.000
SabWil	1.000	1.000	1.000	1.000
SabWWil	1.000	1.000	0.749	1.000
SabPWil	1.000	1.000	1.000	1.000
SabWPWil	1.000	1.000	0.749	1.000
SabDef	1.000	1.000	1.000	1.000
SabWDef	1.000	1.000	1.000	1.000
SabPDef	1.000	1.000	1.000	1.000
SabWPDef	1.000	1.000	1.000	1.000
SabECar	1.000	1.000	1.000	1.000
SabEWCar	1.000	1.000	0.749	1.000
SabEPCar	1.000	1.000	1.000	1.000
SabEWPCar	1.000	1.000	0.749	1.000
SabEWil	1.000	1.000	1.000	1.000
SabEWWil	0.610	0.488	0.359	1.000
SabEPWil	1.000	1.000	1.000	1.000
SabEWPWil	0.610	0.488	0.359	1.000
SabEDef	1.000	1.000	1.000	1.000
SabEWDef	1.000	1.000	1.000	1.000
SabEPDef	1.000	1.000	1.000	1.000
SabEWPDef	1.000	1.000	1.000	1.000

rejected for any of the alternative computational methods. It is not rejected either in the tests concerning the domain of applicability of the Geant4 10.5 parameterized model above approximately 5 keV. However, it is worthwhile to note that the smaller size of the K shell data sample reduces the power of the test with respect to the analysis of total cross sections: for instance, the power of the Boschloo test to detect the observed difference between total cross sections calculated by EPDL97 and with Pratt's correction is 0.93, while it drops to 0.48 for observing the same effect in the K shell data sample.

The scarcity of experimental data hinders a meaningful validation analysis for L subshell cross sections and for outer shells. Regarding L subshells, all computational methods exhibit the same compatibility with experiment with the exception of the Geant4 10.5 parameterization à la Biggs; nevertheless, the observed differences are not statistically significant. It was already remarked in [1] that the experimental data sample for outer shells is inadequate to perform validation tests.

V. CONCLUSION

Statistical tests against a large sample of photoelectric cross section measurements allow a rigorous and objective characterization of the available computational methods and identification of the state of the art in this domain.

The validation tests documented in this paper demonstrate that EPDL97 cross section tabulations, based on Scofield's 1973 calculations, still represent the state of the art for total and K shell photoelectric cross sections: there is no evidence that more recent computational methods surpass EPDL97 compatibility with experimental data. The new version of EPDL, released in ENDF/B-VIII.0, exhibits statistically equivalent behaviour with respect to total and K shell photoelectric cross sections. EPDL97 is used by several Monte Carlo particle transport codes; the results of this paper show that there is no need to move to more recent data libraries for the simulation of the photoelectric effect.

The detailed formulation of the theory documented in [5] allows the evaluation of various physics options. The tests fail to reject the hypothesis of equivalent compatibility with experiment with respect to EPDL97 when accounting for finite level widths and excitation to bound states. No significantly different validation results are observed by adjusting the theoretically calculated cross sections according to either Carlson's or Williams' compilations of atomic binding energies.

The normalization applied to calculations adopting the DHFS approach according to [5], based on the multiconfiguration code by Desclaux, appears to deteriorate the accuracy of total cross sections: goodness of fit tests show that the hypothesis of compatibility with experimental data is rejected in a larger number of test cases, resulting in statistically significant differences in reproducing experimental measurements with respect to EPDL97. The effect of deterioration is not visible in the tests concerning the K shell; nevertheless, the power of these tests is lower due to the smaller experimental data sample involved. Caution should be exercised in using this correction, properly identifying the limits of its applicability.

The results of the validation analysis suggest that the application of the normalization correction could be responsible for the apparently degraded compatibility with experiment of Penelope 2014 photoelectric cross sections with respect to the previous 2011 version.

The replacement of EPDL97 interpolation in Geant4 10.5 *low energy* electromagnetic package with an empirical parameterization à la Biggs-Ligthhill above approximately 5 keV significantly degrades the compatibility of total cross sections with experiment. The original Biggs-Lighthill parameterization, used in Geant4 *standard* electromagnetic package, is not affected by this drawback in the same energy range.

The conceptual framework of hypothesis testing [47] does not allow discerning whether the failure to reject the null hypothesis is due to the hypothesis being "true" or to insufficient evidence from the data to reject it. Therefore, more extensive experimental measurements are needed to quantify the capabilities of partial cross section calculations and of other features, such as accounting for excitation to bound levels and finite level widths, which could not be discriminated

with respect to EPDL97 on the basis of currently available experimental data. High-precision measurements in the proximity of absorption edges, with reliable estimates of their uncertainties, would also be helpful to better characterize the capabilities of the various computational methods.

The results of the validation tests documented in this paper provide guidance to the developers and users of Monte Carlo particle transport codes to optimize the choice of photoelectric cross sections in the simulation of experimental scenarios.

ACKNOWLEDGMENT

The authors express their gratitude to Francesc Salvat, who kindly provided tabulations resulting from [5] and the *PHOTACS-PP* code implementing further computational options on top of them. The authors thank Anita Hollier for proofreading the manuscript and Min Cheol Han for his interest in the early stage of the validation project. The CERN Library has provided helpful assistance and essential reference material for the validation tests. The author Tullio Basaglia wishes to specify that his contribution to this paper concerns the cross section data.

REFERENCES

[1] M. C. Han *et al.*, "Validation of cross sections for Monte Carlo simulation of the photoelectric effect," *IEEE Trans. Nucl. Sci.*, vol. 63, no. 2, pp. 1117–1146, Apr. 2016.

[2] D. E. Cullen, J. H. Hubbell, and L. Kissel, "EPDL97: The Evaluated Photon Data Library'97 version, rev. 5," Lawrence Livermore Nat. Lab., Livermore, CA, USA, Tech. Rep. UCRL-LR-50400, 1997, vol. 6.

[3] J. H. Scofield, "Theoretical photoionization cross sections from 1 to 1500 keV," Lawrence Livermore Nat. Lab., Livermore, CA, USA, Tech. Rep. UCRL-51326, 1973.

[4] P. Andreo, D. T. Burns, A. E. Nahum, J. Seuntjens, and F. H. Attix, *Fundamentals of Ionizing Radiation Dosimetry*. Weinheim, Germany: Wiley, 2017, ch. 3.3.

[5] L. Sabbatucci and F. Salvat, "Theory and calculation of the atomic photoeffect," *Radiat. Phys. Chem.*, vol. 121, pp. 122–140, Apr. 2016.

[6] F. Salvat, J. M. Fernandez-Varea, and J. Sempau, "Penelope-2014—A code system for Monte Carlo simulation of electron and photon transport," Tech. Rep. NEA/NSC/DOC(2015)3, 2015.

[7] D. E. Cullen, *A Survey of Photon Cross Section Data for Use in EPICS2017, Rev. 1*, document IAEA-NDS-0225, Vienna, Austria, 2017.

[8] F. Biggs and R. Lighthill, "Analytical approximations for X-ray cross-sections," Sandia Nat. Lab., Albuquerque, NM, USA, Tech. Rep. SC-RR-66-452, 1967.

[9] F. Biggs and R. Lighthill, "Analytical approximations for X-ray cross-sections II," Sandia Nat. Lab., Albuquerque, NM, USA, Tech. Rep. SC-RR-710507, 1971.

[10] F. Biggs and R. Lighthill, "Analytical approximations for X-ray cross sections III," Sandia Nat. Lab., Albuquerque, NM, USA, Tech. Rep. SAND87-0070, 1988.

[11] G. Amadio *et al.*, "GeantV alpha release," *J. Phys., Conf. Ser.*, vol. 1085, Sep. 2018, Art. no. 032037.

[12] S. Agostinelli *et al.*, "Geant4—A simulation toolkit," *Nucl. Instrum. Meth. A*, vol. 506, no. 3, pp. 250–303, 2003.

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

[14] J. Allison *et al.*, "Recent developments in Geant4," *Nucl. Instrum. Meth. A*, vol. 835, pp. 186–225, 2016.

[15] R. H. Pratt, "Photoeffect from the L shell," *Phys. Rev.*, vol. 119, no. 5, pp. 1619–1626, 1960.

[16] R. H. Pratt, A. Ron, and H. K. Tseng, "Atomic photoelectric effect above 10 keV," *Rev. Mod. Phys.*, vol. 45, no. 2, pp. 273–325, 1973.

[17] R. H. Pratt, A. Ron, and H. K. Tseng, "Erratum: Atomic photoelectric effect above 10 keV," *Rev. Mod. Phys.*, vol. 45, no. 4, pp. 663–664, 1973.

[18] J. Desclaux, "A multiconfiguration relativistic Dirac-Fock program," *Comput. Phys. Commun.*, vol. 9, no. 1, pp. 31–45, Jan. 1975.

[19] J. P. Desclaux, "Erratum notice," *Comput. Phys. Commun.*, vol. 13, no. 1, p. 71, 1977.

[20] F. Salvat and X. Llovet, "Monte Carlo simulation and fundamental quantities," *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 304, Jan. 2018, Art. no. 012014.

[21] R. H. Pratt, "Atomic photoelectric effect at high energies," *Phys. Rev.*, vol. 117, no. 4, pp. 1017–1028, 1960.

[22] T. A. Carlson, *Photoelectron and Auger Spectroscopy*. New York, NY, USA: Plenum, 1975.

[23] F. Salvat, J. M. Fernandez-Varea, and J. Sempau, "PENELOPE-2011: A code system for Monte Carlo simulation of electron and photon transport," NEA, Boulogne-Billancourt, France, Tech. Rep. NEA/NSC/DOC(2011)5, 2011.

[24] D. Brown *et al.*, "ENDF/B-VIII.0: The 8^{th} major release of the nuclear reaction data library with CIELO-project cross sections, new standards and thermal scattering data," *Nucl. Data Sheets*, vol. 148, pp. 1–142, Mar. 2018.

[25] D. E. Cullen. *EPICS2017*. [Online]. Available: https://www-nds.iaea.org/epics/

[26] S. T. Perkins, D. E. Cullen, M. H. Chen, J. Rathkopf, J. Scofield, and J. H. Hubbell, "Tables and graphs of atomic subshell and relaxation data derived from the LLNL Evaluated Atomic Data Library (EADL), Z = 1–100," Lawrence Livermore Nat. Lab., Livermore, CA, USA, Tech. Rep. UCRL-50400-Vol.30, 1991.

[27] D. E. Cullen, "A survey of atomic binding energies for use in EPICS2017," IAEA, Vienna, Austria, Tech. Rep. IAEA-NDS-0224, 2017.

[28] M. C. Han, M. G. Pia, P. Saracco, and T. Basaglia, "First assessment of ENDF/B-VIII and EPICS atomic data libraries," *IEEE Trans. Nucl. Sci.*, vol. 65, no. 8, pp. 2268–2278, Aug. 2018.

[29] *ENDF/B-VIII.0 Errata*. [Online]. Available: https://www.nndc.bnl.gov/endf/b8.0/errata.html

[30] Geant4 Collaboration. *Geant4 Physics Reference Manual, Geant4 Version 10.5*. [Online]. Available: http://cern.ch/geant4-userdoc/-UsersGuides/PhysicsReferenceManual/fo/PhysicsReferenceManual.pdf

[31] D. Cullen *et al.*, "EPICS2014: Electron photon interaction cross section (version 2014), rev.1," Tech. Rep. IAEA-NDS-218, 2015.

[32] *IEEE Standard for System, Software, and Hardware Verification and Validation*, IEEE Standard 1012-2016, 2017, pp. 1–260.

[33] G. P. Williams, "Electron binding energies of the elements," in *CRC Handbook of Chemistry and Physics*, W. M. Haynes and D. R. Lide, Eds., 91st ed. Boca Raton, FL, USA: CRC, 2011, pp. 221–226.

[34] M. Fowler, *Refactoring: Improving the Design of Existing Code*. Boston, MA, USA: Addison-Wesley, 1999.

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

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

[37] R. D. Boschloo, "Raised conditional level of significance for the 2×2-table when testing the equality of two probabilities," *Stat. Neerland*, vol. 24, no. 1, pp. 1–9, Mar. 1970.

[38] S. Suissa and J. J. Shuster, "Exact unconditional sample sizes for the 2×2 binomial trial," *J. Roy. Stat. Soc. A*, vol. 148, no. 4, pp. 317–327, 1985.

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

[40] E. B. Wilson, "Barnard's CSM test of significance," *Proc. Nat. Acad. Sci. USA*, vol. 38, no. 10, pp. 899–905, Oct. 1952.

[41] K. Pearson, "On the χ^2 test of goodness of fit," *Biometrika*, vol. 14, nos. 1–2, pp. 186–191, 1922.

[42] L. Lyons, *Statistics for Nuclear and Particle Physicists*. Cambridge, U.K.: Cambridge Univ. Press, 1989.

[43] R Core Team, R Foundation for Statistical Computing, Vienna, Austria. (2019). *R: A Language and Environment for Statistical Computing*. [Online]. Available: http://www.R-project.org/

[44] M. Friendly, *Visualizing Categorical Data*. Cary, NC, USA: SAS Institute, 2000, Sec. 3.4.

[45] M. Friendly and D. Meyer, *Discrete Data Analysis With R: Visualization and Modeling Techniques for Categorical and Count Data*. Boca Raton, FL, USA: CRC, 2015, Sec. 4.4.

[46] A. Agresti, "A survey of exact inference for contingency tables," *Stat. Sci.*, vol. 7, no. 1, pp. 131–153, Feb. 1992.

[47] W. J. Conover, *Practical Nonparametric Statistics*, 3^{rd} ed. Hoboken, NJ, USA: Wiley, 1999, ch. 2.