

ACCEPTED MANUSCRIPT • OPEN ACCESS

Foundations of Plasma Standards

To cite this article before publication: Luis Alves *et al* 2023 *Plasma Sources Sci. Technol.* in press <https://doi.org/10.1088/1361-6595/acb810>

Manuscript version: Accepted Manuscript

Accepted Manuscript is “the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an ‘Accepted Manuscript’ watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors”

This Accepted Manuscript is © 2023 The Author(s). Published by IOP Publishing Ltd.

As the Version of Record of this article is going to be / has been published on a gold open access basis under a CC BY 3.0 licence, this Accepted Manuscript is available for reuse under a CC BY 3.0 licence immediately.

Everyone is permitted to use all or part of the original content in this article, provided that they adhere to all the terms of the licence <https://creativecommons.org/licenses/by/3.0>

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions may be required. All third party content is fully copyright protected and is not published on a gold open access basis under a CC BY licence, unless that is specifically stated in the figure caption in the Version of Record.

View the [article online](#) for updates and enhancements.

Foundations of Plasma Standards

Luís L. Alves¹, Markus M. Becker², Jan van Dijk³, Timo Gans⁴, David B. Go⁵, Katharina Stapelmann⁶, Jonathan Tennyson⁷, Miles M. Turner⁴ and Mark J. Kushner^{8,9}

1. Instituto de Plasmas e Fusão Nuclear, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal, llalves@tecnico.ulisboa.pt
2. Leibniz Institute for Plasma Science and Technology (INP), Greifswald, Germany markus.becker@inp-greifswald.de
3. Eindhoven University of Technology, Department of Applied Physics, j.v.dijk@tue.nl
4. School of Physical Sciences and National Centre for Plasma Science and Technology, Dublin City University, Dublin, Ireland, timo.gans@dcu.ie, miles.turner@dcu.ie
5. University of Notre Dame, Department of Aerospace and Mechanical Engineering, Notre Dame, IN 46556-5637, USA, dgo@nd.edu
6. North Carolina State University, Department of Nuclear Engineering, Raleigh, NC 27695, USA, kstapel@ncsu.edu
7. Department of Physics and Astronomy, University College London, United Kingdom, j.tennyson@ucl.ac.uk
8. University of Michigan, Electrical Engineering and Computer Science Department, Ann Arbor, MI 49109-2122, USA, mjkush@umich.edu
9. Author to whom correspondence should be addressed.

Abstract

The field of low-temperature plasmas (LTPs) excels by virtue of its broad intellectual diversity, interdisciplinarity and range of applications. This great diversity also challenges researchers in communicating the outcomes of their investigations, as common practices and expectations for reporting vary widely in the many disciplines that either fall under the LTP umbrella or interact closely with LTP topics. These challenges encompass comparing measurements made in different laboratories, exchanging and sharing computer models, enabling reproducibility in experiments and computations using traceable and transparent methods and data, establishing metrics for reliability and in translating fundamental findings to practice. In this paper, we address these challenges from the perspective of LTP standards for measurements, diagnostics, computations, reporting and plasma sources. This discussion on standards, or recommended best practices, and in some cases suggestions for standards or best practices, has as the goal improving communication, reproducibility and transparency within the LTP field and fields allied with LTPs. This discussion also acknowledges that standards and best practices, either recommended or at some point enforced, are ultimately a matter of judgment. These standards and recommended practices should not limit innovation nor prevent research breakthroughs from having real-time impact. Ultimately, the goal of our research community is to advance the entire LTP field and the many applications it touches through a shared set of expectations.

1 2 3 4 5 I. Introduction

6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
59 In the realm of research and development in the field of low-temperature plasmas (LTPs), there is a natural transition or evolution from concept to implementation. An embryonic idea or concept is first investigated with there being little understanding of the underlying processes. Measurements are difficult to perform in this new parameter space, and models are difficult to formulate in the absence of fundamental data and knowing the dominant processes (and how to computationally represent them). Experimental data and model results both have large uncertainty bars. In spite of these outcomes not having high precision, the *first-time* trends revealed by these perhaps not precise results provide remarkable insights to new processes or a better understanding of underlying physics. They stimulate thought and innovation. As a result of these works, it becomes more clear what data are required to make more precise measurements and build more focused models. Through this improvement in understanding, more definitive experiments are designed and conducted, more relevant data is produced and more representative models are constructed. Uncertainty bars are reduced and understanding is improved. The precision of the investigation and of our understanding is improved.

30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
59 This cycle of improved understanding enabling more definitive experiments in turn enables precision to be improved. When there is sufficient fundamental understanding of the basic underlying processes, then attention expands to address higher levels of precision. This higher level of precision produces refined reaction mechanisms and data, and smaller linewidths in spectroscopic measurements. This precision then enables yet higher levels of precision. When understanding and precision reach a threshold level, technology transfer begins wherein, for example, plasma sources are designed and built to provide reproducible and predictable *doses* of plasma activated species. The end result is often a commercial product.

43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
59 Underlying this transition in concept to implementation, and increasing degrees of decision, should be a culture of standards. A dictionary definition of standards is "something set up and established by authority as a rule for the measure of quantity, weight, extent, value, or quality" [1]. Standards are absolutely critical and necessary in applications involving life-safety, and are embodied in building codes and crash-worthiness requirements for automobiles. Another perspective of standards is "an established norm or requirement for a repeatable technical task which is applied to a common and repeated use of rules, conditions [and] guidelines..." [2]. This perspective emphasizes the need for common practices to enable exchange of information and to gauge the goodness of a process.

59 The field of LTPs extends from concept to commercialization, and so has an extremely

1
2
3 diverse need for standards enabling increasing levels of precision, with different levels of ex-
4
5 pectation for compliance. One class of standards is, in principle, independent of where one lies
6
7 in the continuum between concept and precision. These standards address the exchange of
8
9 information and data consistently, unambiguously and reproducibly; verifying a computer
10
11 model, calibrating a diagnostic, or reporting on a result. Another class of standards applies
12
13 when moving beyond conceptualizing, a realm in which higher levels of precision are possible.
14
15 These standards address validation of codes, confirmation of reaction mechanisms, production
16
17 and dissemination of data, and development of standard sources to enable measurements to be
18
19 collated across laboratories. A final class of standards apply to the technology transfer end of
the innovation chain. These standards address reproducibility, safety and reliability.

20
21 The proper and measured use of standards is critical to fostering the entire breadth of
22
23 the innovation chain, from ideation to commercialization. Some standards should be thought
24
25 of as being universal. These standards address transparency in reporting, defining techniques
26
27 and methods, and making data available. As one transitions from concept to precision, the
28
29 enforcement of standards becomes more appropriate. Premature enforcement of standards runs
30
31 the risk of stifling innovation. Neglecting standards can lead to inefficiency and questionable
conclusions.

32
33 In this paper, we discuss the development and use of standards in the LTP community.
34
35 The term standards is used here in the most general way and is synonymous with recommended
36
37 best practices and processes. The intent of the discussion in this paper is to provide guidance
38
39 for how your research results can be communicated and utilized by the LTP community more
40
41 efficiently, more reliably, more reproducibly and less ambiguously. Achieving these goals
42
43 works towards transparency, improving the acceptance of the research and accelerating ad-
vances in the field while not stifling innovation.

44
45 This paper has the following sections addressing the wide range of experimental, com-
46
47 putational, theoretical and technology transfer elements of the LTP field. With the discussion
48
49 in each section being wide ranging, here we also provide summary statements, *action-items* or
50
51 *recommended path* for each section. These items are intended to provide the reader with a high
52
53 level perspective of our recommendations.

54
55 II. *Plasma Sources*: A discussion of how standard plasma sources may benefit exchange
of data and experiences between laboratories.

56
57 While it is recognized that developing and deploying community-driven stand-
58
59 ard plasma sources is a large undertaking, experience with the GEC Reference Cell and
60 the COST Jet has shown how valuable standard sources can be. The community should,

1
2 whenever possible, use the already existing standard sources as “calibration tools” for
3 their own plasma devices to improve fundamental understanding of plasma processes
4 and accelerate translation of scientific findings to applications. Further standards that
5 are easier to realize than a complex standard plasma source, for example a standard
6 substrate to investigate plasma-surface interactions, should be the next focus of com-
7 munity-driven standardization efforts.
8
9

10 **III. *Plasma Diagnostics:*** In developing and reporting on measurements, what information
11 and practices are required to ensure transparency and reproducibility.
12
13

14 In the absence of established diagnostics standards, it is recommended that in
15 their reporting, researchers follow a minimum framework to support transparency. This
16 framework should include: A clear and detailed description of the equipment used for
17 the diagnostics, the precise implementation of the diagnostics, the input data to the anal-
18 ysis and the analysis techniques as well as access to the raw data. This would allow
19 potential reanalyzing of the data using different techniques or when analysis techniques
20 are improved.
21
22

23 **IV. *Data and Reaction Mechanisms:*** Diagnostics and modeling require fundamental data
24 and reaction mechanisms. How should these data and mechanisms be distributed, ar-
25 chived and validated?
26
27

28 Data standards in LTP physics have been developed in a meritocratic way where
29 individuals or organizations propose standards. The standards that succeed are those
30 that bring the most benefits to the community. These successes heavily depend on the
31 availability and usability of tools that can utilize the data. A good example of such a
32 successful combination is the LXCat database and the BOLSIG+ program for solving
33 Boltzmann's equation. This meritocratic process will likely continue for some time.
34 Wider adoption of standards than has happened to date will depend on the willingness
35 of the standards developers to listen better to the needs and wishes of the community,
36 and to look beyond the boundaries of their respective application domains.
37
38

39 **V. *(Quick) Data Generation:*** In the absence of experiments or first principles calculations,
40 are there standard methods to quickly produce needed data for models and diagnostics?
41
42

43 The input data, regardless of the method of generation, need proper statements
44 of uncertainties. Procedures for uncertainty quantification, particularly for theoreti-
45 cal/calculated data, need to be developed and routinely applied. To speed the generation
46 of this needed data, machine learning should be explored as a route to providing data on
47 the many processes for which little or nothing is now known.
48
49

1
2
3 VI. *Verification & Validation (V&V):* How can standards work towards improving the re-
4
5 liability and accuracy of computer models?

6
7 In the short term, authors should ensure that their reporting of calculations fully
8
9 meets the requirements of scientific reproducibility, by making certain that the algo-
10
11 rithms, numerical parameters and data that have been employed are fully specified and
12
13 traceable (in the case of data). Whenever practical, source code and data files should
14
15 be made available. Editors and referees should consider carefully what minimal stand-
16
17 ards are appropriate in their context, taking note of the considerations discussed above.

18 VII. *Open Source and Publicly Available Codes:* In development and maintenance of com-
19
20 puter models, how can standards improve access and speed innovation?

21
22 The LTP community should use the main conferences in the field as a forum to
23
24 design and carry out collective efforts for the verification and benchmarking of codes.
25
26 These efforts could be directed towards open-source codes as a means to develop spe-
27
28 cific recommendations on revisions to the codes, as well as directed towards less avail-
29
30 able codes as a means to guide the authors of those codes.

31 VIII. *Reporting:* Nearly every technical endeavor produces a report – a journal article or
32
33 archival document. What are the expectations for transparency in reporting that aid in
34
35 the communication and reproducibility of results?

36
37 Reporting standards for the most part simply reflect proper actualization and
38
39 effective communication of the scientific method. That is, proper reporting should nec-
40
41 cessarily be a central element of journal articles. Practically, editors and publishers of
42
43 journals popular among the LTP community, including the one publishing this paper,
44
45 could adopt submission checklists or similar mechanisms, not unlike Nature, that ensure
46
47 that published papers follow expected reporting conventions.

48 IX. *Plasma Dose:* Many fields have standards for quantifying reactivity. Can such a stand-
49
50 ard be developed for LTPs?

51
52 The definition and use of dose (e.g., energy deposited to produce a given prod-
53
54 uct) has been successfully adopted in fields such as radiation physics. Those successes
55
56 result in part from the generation and delivery of the activating energy being separable
57
58 from interaction with the target, and due to the dose itself being weakly sensitive to the
59
60 absolute energy of delivery. The definition of dose for LTPs has been considerably less
successful due to the lack of this separability and due to product generation being sen-
sitive to the distribution of the delivered energy. Advancing the concept of dose for
LTPs, useful for calibrating plasma sources and protocols, would benefit from having

1
2
3 different categories of doses associated with, for example, reactant generation and use
4 of those reactant fluxes. These definitions would best be refined by community led
5 efforts.
6
7

8 X. *Technology Transfer:* One outcome of translational research is technology transfer and
9 commercialization. What standards will enhance this process?
10
11

12 To support technology transfer, data should be collected systematically and un-
13 der well-defined conditions. The definition of reporting standards and the use of in-
14 teroperable data formats will further help to overcome the hurdles in establishing new
15 plasma applications.
16
17

18 XI. *Concluding Remarks*
19
20
21

22 II. Plasma Sources 23

24 Standard plasma sources have the goal of providing *identical* experimental platforms
25 for researchers in different laboratories. These identical platforms would enable a variety of
26 experimental techniques to be employed and modeling performed with there being a minimum
27 in laboratory-to-laboratory variation in reactor configuration and operating conditions. In this
28 way, researchers can exchange and compare experimental data and modeling results with the
29 goal of gaining fundamental understanding of the physical and chemical properties of the
30 plasma and plasma-chemistry. This would be accomplished without needing to correct or ac-
31 count for differences in these data due to known (or unknown) differences in the reactors or
32 operating conditions. Community-driven efforts to use one experimental platform that can be
33 deployed in various laboratories where researchers with different expertise contribute to the
34 development of a “big picture” have been successfully employed with the *GEC Reference Cell*
35 and later with the *COST reference microplasma jet*, both described in more detail below. Data
36 generated from different research groups can help contextualize and cross correlate results, and
37 to accelerate the understanding of fundamental processes.
38
39

40 A. Currently Available Standard Plasma Sources 41

42 A. 1. GEC Reference Cell 43

44 The concept of the GEC Reference Cell (GECRC) grew out of a workshop held at the
45 Gaseous Electronics Conference (GEC) in Minneapolis, Minnesota (USA) in 1988 and was
46 further refined by the GEC community in the following year. (The GEC in 1989 was held in
47 Palo Alto, California (USA) south of San Francisco. The “Workshop on the Reference System
48 for RF Plasma Processing Research” took place on October 17, scheduled for 3:45 pm - 5:30
49 pm. The GECRC was developed to provide a standard plasma source for the GEC community.
50
51

pm. Just as Joseph T. Verdeyen summarized his points at 5:04 pm, the Loma Prieta earthquake occurred, concluding the workshop with Prof. Verdeyen having the final word.) The GECRC was originally designed to advance the understanding of radio frequency excited capacitively coupled plasmas (CCPs) of the type used in microelectronics fabrication [3] and later to address inductively coupled plasmas (ICPs) as those devices began to be used in the microelectronics industry. By having a standard cell that would be implemented in laboratories throughout the world, measurements made in different laboratories and modeling could be compared on a side-by-side basis (see example in Figure 1). Eliminating lab-specific aspects of the plasma source (e.g., wall materials, distance from the powered electrodes to side walls, pumping scheme) enabled there to be more focus on the properties of the plasma. The GECRC was widely adopted with dozens of cells employed throughout the world, enabled in part by a commercial vacuum equipment company having a part-number for ordering an assembled cell at moderate cost. The GECRC is still actively used more than 30 years later, having served as a valuable asset to the research community [4,5] and continues to be a benchmark for validating codes (see Figure 1 as a recent example).

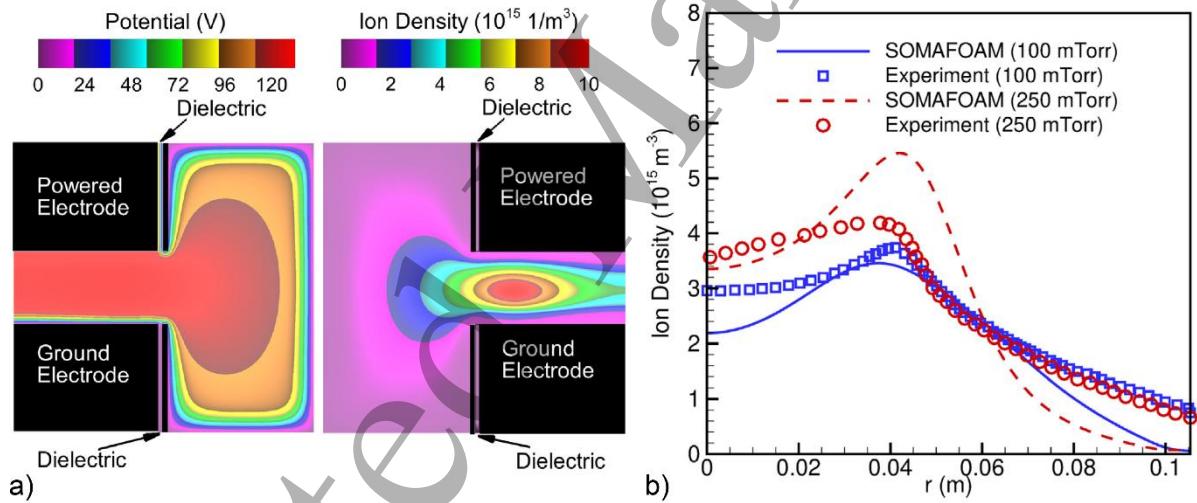


Figure 1: The GECRC continues to be a benchmark for validating codes and investigating fundamental plasma properties. (a) Simulation results using the SOMAFOAM platform for plasma potential and Ar^+ density in a CCP sustained in 100 mTorr Ar powered with 400 V peak-to-peak at 13.56 MHz. (b) Comparison to experimental results. Adapted from Ref. [6]

This initial foray into standard plasma sources was an eye-opening experience for all concerned. This eye-opening was a result of under-estimating the importance to measurements and plasma properties of what were considered minor differences in power-supplies, matching networks and electrical connections. After deployment of the first GECRC, a workshop was held to compare measurements of the most basic properties of a CCP sustained in

1
2
3 argon - current, voltage, phase and electron density. There were hugely unexpected variations
4 in even the current-voltage measurements made in different laboratories. These variations were
5 eventually attributed to subtle and sometimes not so subtle differences between laboratories in
6 the matching networks, how cables were configured between the power supplies, matchboxes
7 and plasma chamber; and stray capacitance and inductance due to what were thought to be
8 inconsequential modifications of the cell. This experience focused the entire community's at-
9 tention on viewing plasma sources as a system, beginning with the power supply, cables and
10 matchboxes and extending through the plasma chamber and method of electrical termination.
11 Guidelines were established for those components to achieve more reproducible results between
12 laboratories. These experiences were translated to practice in the semiconductor plasma pro-
13 cessing industry to help achieve reproducibility between plasma tools and in the same tools
14 before and after maintenance. Although not appreciated at the time, gas impurities can also
15 have a large impact on plasma properties (as discussed in Section III) and future reference
16 plasma sources should include impurity specifications in at least their calibration procedures.
17
18

28 **A.2 COST Reference Microplasma Jet**

29 The COST Reference Microplasma Jet (COST-jet) originated through efforts supported
30 by the European COST (Cooperation in Science and Technology) Action MP1011 on “Bio-
31 medical Applications of Atmospheric Pressure Plasma Technology” [7]. Within this COST
32 Action it was recognized that measurements of plasma properties made in a wide variety of
33 mostly home-made atmospheric pressure plasma jets complicated comparison of results ob-
34 tained in different laboratories, leading to delays in interpretation of those results and in fur-
35 thering understanding of fundamental properties of the plasma. A working group was formed
36 to define a device that could serve as a reference plasma jet for researchers in the sub-field of
37 plasma medicine which relies heavily on plasma jets. The motivation to use a reference plasma
38 source to advance understanding of processes occurring in atmospheric pressure plasma jets is
39 similar to that of the GECRC. The end product was the COST-jet. Much effort has been de-
40 voted to improving the reproducibility of the COST-jet itself and among multiple devices in
41 different laboratories [8]. See Figure 2 as an example of these efforts.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

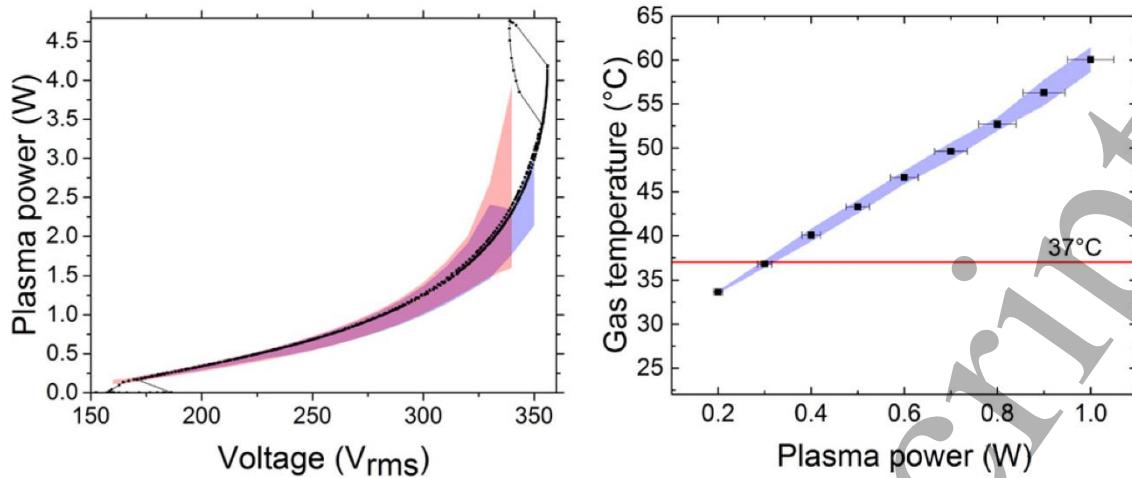


Figure 2: Reproducibility of the COST-jet: Plasma and device properties in four sources were measured independently. The black dots indicate measured data from one jet, the shaded area indicates the deviation between the four sources. (left) Plasma power as a function of voltage and (right) measured effluent gas temperature at 3 mm distance from the nozzle as a function of plasma power. Adapted from Ref. [8].

The COST-jet has integrated voltage and current probes. A calibration protocol, instructions on how to measure and calculate dissipated power, and operation protocols for taking reproducible measurements have been published [7]. Several computational models to capture the chemical kinetics in the gas phase have been developed [9,10,11]. Applications of the COST-jet have been recently reviewed by Gorbanov et al. [12] and range from polymer surface modifications [13] and studying anti-cancer effects [14], to investigating plasma-driven biocatalysis [15].

A.3 Other Commercial Sources

In addition to community-derived standard plasma sources such as the GECRC and COST-jet, commercial plasma sources are also available that could serve as standard sources. While these commercial sources were originally not intended to be research tools, they are manufactured to extreme precision and reproducibility. However, most commercial plasma sources are designed for application areas in which the plasma is delivered in a highly constrained manner (e.g., plasma etching reactors, hollow cathode lamps, and fluorescent lamps) with little access for diagnostics. Standard research plasma sources are intended to be reproducible with flexibility in operating conditions (pressure, power, gas mixture) and with access for diagnostics. In new application areas it is difficult to motivate commercial entities to offer standard sources or modify existing products, as the market is admittedly small, with there being an early success in doing so for the GECRC. That said, several commercial plasma sources have become de-facto standard sources due to their widespread use. For example, the commercial kINPen plasma jet has been extensively characterized and modeled [16], while being used

1
2
3 in laboratories throughout the world.
4

5 The responsibility for building and maintaining community-driven standard sources
6 may exceed the capability of the research community, which may explain that although suc-
7 cessful, only two community-driven standard sources have been developed to date. The ideal
8 situation may be a community-driven design that is transferred to a company which then com-
9 mercializes the device, while also being open to making periodic community-driven modifica-
10 tions.
11
12

13 B. Opportunities and Challenges for the LTP Community

14

15 Standard plasma sources have the potential to advance the understanding of a particular
16 type of plasma configuration and to accelerate research, including the translation of findings
17 from basic research to industrial applications. Plasma medicine and plasma agriculture as well
18 as plasma-based water purification have shown promising results over the last decade [17].
19 Many of the plasma devices used in these investigations have been under-diagnosed and oper-
20 ating conditions have been under-specified due to the emphasis being on the outcome of the
21 application. For example, the electrical environment around plasma jets and the substrate are
22 rarely mentioned in publications in spite of their having strong effects on the performance of
23 the plasma apparatus. The end result has been a lack of reproducibility of the application fo-
24 cused results between laboratories and even within laboratories. This lack of reproducibility
25 and lack of knowledge of fundamental properties, such as the flux of plasma generated radicals
26 onto the sample, has resulted in limiting our understanding of the basic processes of the inter-
27 action of plasma with liquid or biological interfaces. Standard plasma sources in all of these
28 areas would help to i) contextualize results, ii) compare and benchmark experiments, and iii)
29 develop and validate models for the plasmas.
30
31

32 The use of standard plasma sources should be a tool towards improving fundamental
33 understanding of plasma processes and speeding translation of scientific findings to applica-
34 tions, while not hindering innovation. It would be detrimental to the field to mandate that to be
35 credible, every researcher must use the same plasma sources, as it would also be detrimental to
36 have no standard plasma sources. Comparing experimental and computational results obtained
37 with standard plasma sources with other well characterized sources can facilitate broadening
38 and extending the operational space. Doing so will similarly advance understanding and help
39 to contextualize results and to identify mechanisms, as it has been done successfully with the
40 COST-jet in comparison with the well-studied kINPen [16,18,19] and a ns-pulsed DBD [20].
41
42

In this mode of operation, standard plasma sources could act as “calibration sources” for diagnostics and models, which may be particularly important in research areas using several reactor designs, developing new sources and scaling up processes.

The major emphasis to date on standard sources has been on plasma generation with there being less emphasis on plasma-surface interactions, in spite of the latter being a major outcome of researchers’ efforts. A community-driven design for a reference plasma-surface treatment system would enable combined fundamental investigations of the plasma source and interface behavior. The outcomes would directly assist in determining limitations and advantages of the process, performance levels in comparison to other plasma devices or existing technology, and to identify the key technical challenges that must be overcome before the technology can be upscaled.

In addition to standard plasma sources, standard protocols for making and calibrating measurements would help the community further understand basic processes in the plasma as well as in interfacial interactions. While standard protocols are common in other research areas, the LTP community lacks standard diagnostics and reporting formats, both further discussed in Sections III and VIII, respectively. The LTP community is extremely intellectually diverse, with a large variety of plasma devices and applications as well as a large variety of educational and training backgrounds. Diagnostics are often unique to a single laboratory and part of the research program itself (e.g., mass spectrometry for atmospheric pressure plasmas [21]) or widely used but without a standard procedure on how to calibrate and report results (e.g., optical emission spectroscopy [22,23]). The end result has been difficulties in assembling results from multiple laboratories to create a more complete picture of plasma phenomena.

The LTP community is not the only research community facing challenges like this and several publishers recently launched protocol journals to work towards reproducibility by sharing a detailed, step-by-step protocol and to take advantage of expert peer review to refine and shape protocols. The LTP community can particularly benefit from these efforts as plasmas can be very sensitive to small variations in parameters or protocols due to the underlying complex and non-linear physics. The LTP community is encouraged to develop and share standard protocols to increase the reproducibility between laboratories and to transfer knowledge. That said, the goal of having protocols is to improve reproducibility and sharing of results, and not to stifle innovation. It is sometimes a first-time measurement using a newly developed diagnostic on a new source that opens up an entirely new field of research. For example, the E-FISH (electric field induced second harmonic generation) technique for measuring electric fields was demonstrated using a non-standard, home-built setup, and has since dramatically

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
beneficially affected plasma diagnostics over the last 5 years since that demonstration [24,25].

III. Plasma Diagnostics

LTPs are weakly ionized complex multi-species systems, including charged species (electrons, positive ions, negative ions), ground state neutral species (gas-phase background species, dissociation products, reaction products), excited species (electronic, ro-vibrational) and photons. The different species are not in thermal equilibrium with each other and can exhibit very different temperatures and energy distributions. Light electrons typically have mean energies in the range of 1-5 eV while heavier species in the plasma bulk typically have temperatures only slightly elevated above room temperature, around hundreds of degrees Kelvin. Ions adjacent to surfaces that are accelerated in the boundary sheath electric field can reach energies of hundreds of eV. The densities of and dynamics of charged species transport are closely coupled with the spatial and temporal structure of the electric field, both local and remote.

The investigation and characterization of LTPs are often challenging and require the combination of several specialized diagnostic techniques for the measurement of species densities, energy distributions (temperatures) and electric fields. Frequently employed techniques include electrical diagnostics, mass spectrometry and optical diagnostics. The different techniques provide complementary information about various plasma parameters and each technique exhibits their own specific advantages and disadvantages depending on the plasma environment and parameters of interest. The role of diagnostics tends to be multifold, going beyond diagnosing and characterizing plasmas. Diagnostics also play key roles in validation of models and simulations as well as being sensors for process control in industrial settings.

Of paramount importance in deploying diagnostics is their reliability or accuracy. The assessment of reliability or accuracy is best made by comparing results of diagnostics to known, calibrated standard sources, or to a theoretical expression or computation if the conditions are amenable to theory. In the absence of standard sources or an applicable theory or computation, the reliability of diagnostics is often determined by whether they reproduce the measurements of other diagnostics. In these comparisons, the average value of a plasma quantity obtained from many measurements made in different laboratories is basically declared to be the “correct” value. Outliers (least consistent with other measurements, e.g. outside a certain confidence band) are then considered to be “less correct”. As unsatisfying as this technique is, in the absence of a known, calibrated standard, applicable theory or computation, there are few other options. The liability of this method is that the outlier may indeed be the correct measurement.

1 2 3 A. Brief Overview of Frequently Employed Plasma Diagnostics 4

5 A detailed discussion of various LTP diagnostics can be found in reference [26]. Only
6 a brief overview of frequently employed techniques is outlined here, along with specific ad-
7 vantages and disadvantages.
8

9 Comparatively simple external electrical measurements of voltage and current are
10 mostly non-intrusive and can provide useful information about charged species and fields in the
11 plasma [27]. However, these data are indirect and require interpretation based on model as-
12 sumptions for the plasma. Internal electrical probes [28], such as Langmuir probes, hairpin
13 probes and retarding field analyzers, can provide more direct and spatially resolved information.
14 Nevertheless, the analysis still requires model assumptions. An inherent disadvantage of inter-
15 nal electrical probes is the potentially intrusive nature and limited applicability in the harsh
16 environments of reactive plasmas as well as thermal limitations of probe materials.
17

18 Mass spectrometry is typically an external non-intrusive technique to measure neutral
19 particle and ion densities as well as energy distribution functions [29]. These are particularly
20 valuable as they are key parameters in plasma-surface interactions. Details of the equipment
21 and data interpretation can be complicated. Similar to internal electrical probes, harsh environ-
22 ments can also be a challenge for mass spectrometry due to its direct contact with reactive
23 plasma species. That contact may change the composition of the measured species as well as
24 the dissociation induced by the ionizer requiring the analysis of cracking patterns to identify
25 complex species.
26

27 Optical diagnostics are versatile and can provide non-intrusive information about
28 plasma parameters with high temporal and spatial resolution [30], e.g. through using ICCD
29 imaging [31] or three dimensional computer assisted tomography [32]. Passive optical emis-
30 sion spectroscopy is experimentally comparatively simple and is truly non-obtrusive with the
31 possible exception of modifying the plasma cell to provide optical access. Due to its robustness
32 and comparatively low cost, optical emission is often employed for plasma monitoring. How-
33 ever, optical emission provides indirect information about plasma parameters. The raw optical
34 emission data must be deconvolved to obtain absolute data (such as densities and temperatures)
35 using what are sometimes complex model assumptions. Optical emission also strongly relies
36 on the availability and accuracy of atomic and molecular data for, for example, oscillator
37 strengths and collisional quenching coefficients.
38

39 Active optical diagnostics, often laser-based spectroscopy techniques, can provide di-
40 rect and highly accurate information, typically only requiring moderate model assumptions
41 [33]. Widely used techniques include absorption spectroscopy and laser induced fluorescence
42

(LIF) spectroscopy. Absorption spectroscopy directly measures line-integrated absolute species densities through photon absorption using wavelengths ranging throughout the spectrum of vacuum ultraviolet (VUV), ultraviolet (UV), visible (VIS) and infrared (IR). Light sources vary from synchrotron radiation and a broad variety of laser systems to comparatively simple classical light sources as, for example, commonly employed in Fourier Transform Infrared (FTIR) spectroscopy. Measurement sensitivity is often improved through multipass absorption techniques, either using simple multipass cells or cavity based approaches as in cavity ring down spectroscopy (CRDS). LIF spectroscopy is also based on photon absorption [34]. This technique can encompass either one photon (LIF) or two-photon absorption LIF (TALIF) spectroscopy. In both techniques, LIF and TALIF, the subsequently emitted fluorescence photon is detected and analyzed. These techniques can provide high spatial and temporal resolution for accurate measurements of species densities. The measurements are valuable as reliable reference points for other diagnostic techniques as well as theoretical and computational investigations. However, they are experimentally involved and tend to require costly equipment.

It should be noted that all active optical diagnostics are intrusive (or perturbative) because the diagnostics change the density of the species being observed in a given atomic or molecular level. Even elastic scattering techniques (e.g., Rayleigh or Thomson scattering) can transfer power to the plasma. The degree of intrusiveness ranges from negligible to significant (the latter being the origin of diagnostic techniques such as optogalvanic spectroscopy [35]). It is the researchers' responsibility to determine how perturbing their diagnostic is, and how to account for those perturbations in the analysis of the data.

40 B. Case Study for Diagnostics Benchmark and Simulation Validation

41 In this subsection an example case study is discussed to illustrate benchmarking different
42 diagnostic techniques against each other, and the interplay with computational simulations
43 for validation. The case study focuses on plasma sources similar to the COST-jet discussed in
44 Section II. The sources include slight modifications due to these studies being performed during
45 different stages of development of the reference source, as well as there being modifications
46 made to accommodate specific diagnostic requirements. The plasma sources operated in a helium
47 gas flow through a 1 mm discharge gap with varying humidity admixtures. The power
48 delivery was radio frequency (RF) capacitively coupled at 13.56 MHz. Details of the individual
49 experimental setups can be found in Refs. [36,37,38].

50 The diagnostics measured absolute atomic oxygen densities using different techniques:
51
52
53
54
55
56
57
58
59
60

1
2
3 ns-TALIF [36], synchrotron VUV absorption [37] and ps-TALIF [38]. Synchrotron VUV ab-
4 sorption measurements are the most direct method using a one-photon absorption process hav-
5 ing a well characterized absorption cross section. As such, the method does not require an
6 additional calibration process. On the other hand, VUV absorption is also an experimentally
7 demanding technique and limited to strict vacuum protocols, while providing a column density
8 as opposed to a local density. Details of the VUV absorption process are discussed in Refs.
9 [38,39]. While still complex, ns-TALIF is experimentally the least demanding technique
10 among the techniques that were compared. Nonetheless, ns-TALIF relies on estimates of col-
11 lisional de-excitation rates and requires calibration, producing additional uncertainty. Details
12 are discussed in Refs. [36,38]. Picosecond TALIF (ps-TALIF) reduces the need to estimate
13 collisional de-excitation rates as typically required in ns-TALIF. With its shorter pulse length,
14 direct measurements of effective de-excitation rates can be made. Similar to ns-TALIF, ps-
15 TALIF still requires a calibration procedure, but can also be applied in more application rele-
16 vant laboratory conditions. Nevertheless, ps-TALIF is also experimentally complex and costly
17 to implement. Details are discussed in Ref. [38].

18
19 All three techniques used in this comparison have different advantages and disad-
20 vantages with associated limitations and uncertainties. These differences are important to con-
21 sider in assessing the results of the benchmark study in addition to the slightly different plasma
22 sources and potential differences in the measurements introduced by those differences in
23 sources. Measured absolute atomic oxygen densities for varying humidity admixtures obtained
24 using ns-TALIF [36], synchrotron VUV absorption [37] and ps-TALIF [38] are shown in Fig-
25 ure 3. In general, there is good agreement across the three techniques with there being best
26 agreement between VUV absorption and ps-TALIF. The differences compared to ns-TALIF
27 are systematic and are likely explained by the needed estimate of collisional de-excitation rates
28 based on quenching coefficients [38]. The ns-TALIF and ps-TALIF measurements were cali-
29 brated using the same method using xenon as reference gas [40]. Given this calibration, in
30 principle, the difference between the two TALIF techniques lies with the analysis. Recent stud-
31 ies and direct measurements of the xenon two-photon absorption cross section suggest that there
32 is an additional systematic uncertainty in this calibration technique [41,42]. However, this un-
33 certainty only affects the absolute densities and should not account for the relative differences.
34 That said, there is also a discrepancy between VUV absorption and ps-TALIF at low humidity
35 admixtures. Details of the discrepancy are discussed below.

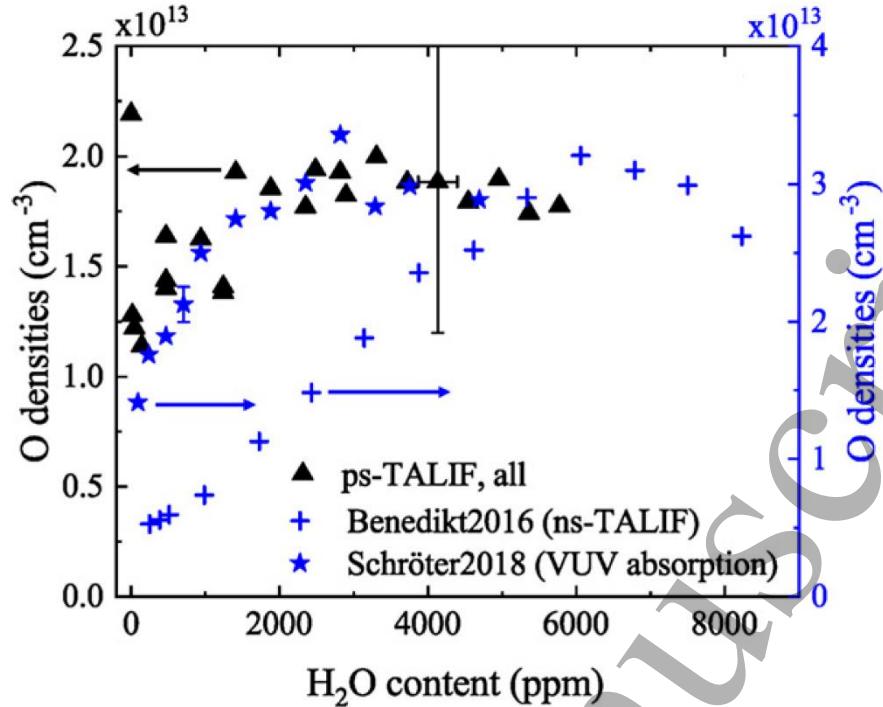


Figure 3: Comparison of measured absolute atomic oxygen densities in similar plasma sources using different diagnostic techniques: ns-TALIF [36], synchrotron VUV absorption [37] and ps-TALIF [38]. The plasma sources operate in a helium carrier gas flow through a 1 mm discharge gap with varying humidity admixtures. The power delivery is a RF-CCP coupled at 13.56 MHz. Reproduced from Ref. [38].

A comparison of the ps-TALIF measurements with computational simulations as a function of humidity admixture is shown in Figure 4. The computational simulations were based on global model assumptions [43] using different levels of O₂ impurities in GlobalKin [44]. Details of the simulations and the reaction mechanism are in Ref. [38]. There is generally good quantitative agreement between simulation and experiment within experimental uncertainties as indicated by error bars in the figure. Without oxygen impurities (0 ppm O₂) the simulation predicts a continuous increase of atomic oxygen as the humidity admixture increases. This behavior agrees with VUV absorption measurements carried out under strict vacuum conditions [37] as well as with ns-TALIF measurements carried out in a controlled helium atmosphere [36]. Particularly interesting is the influence of small O₂ impurities (4 to 12 ppm) at low humidity content. In this regime, the simulation predicts significantly elevated densities of atomic oxygen from dissociation of O₂ impurities. These elevated atomic oxygen densities agree with ps-TALIF experiments operated at ambient laboratory conditions, typically influenced by small impurity levels [38].

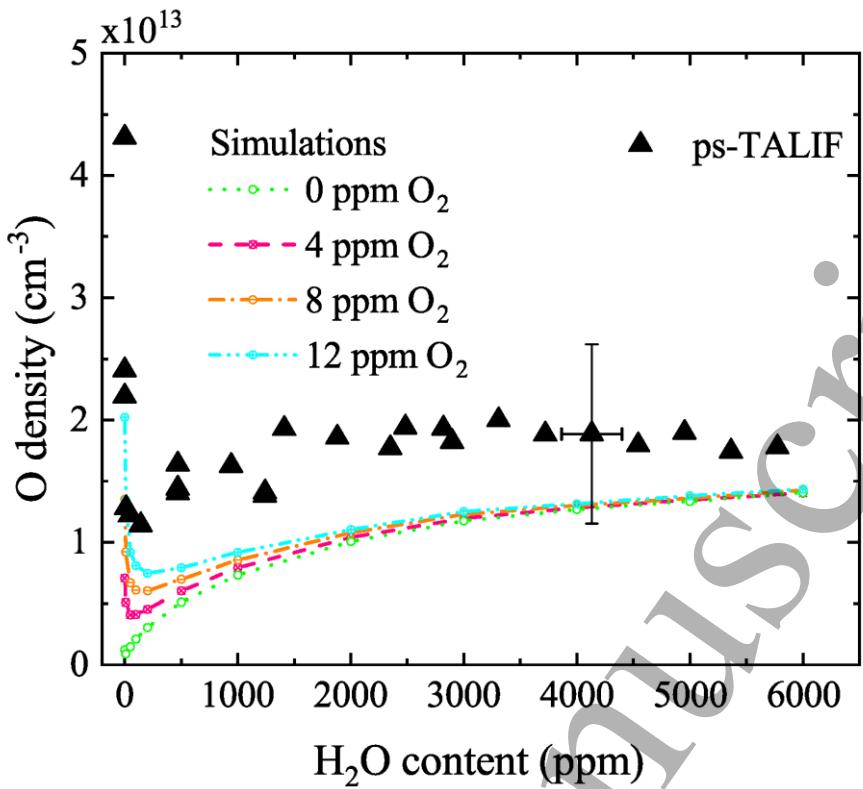


Figure 4: Comparison of experimental measurements of absolute atomic oxygen densities and computational simulations as a function of humidity admixture. Experimental measurements were carried out using ps-TALIF. Computational simulations were based on GlobalKin [44] using different levels of O_2 impurities. Reproduced from Ref. [38].

This case study demonstrates that benchmark experiments for different diagnostic techniques may have an uncertainty envelope. Validation with computational simulations can provide synergy with additional mechanistic insight. This specific case study also demonstrates the potential importance of impurities. The development of standard protocols should include specified impurity levels (or dependence on impurities), as should also be the case in development of reference plasma sources, as mentioned in Section II.

C. Recommendations for Standards in Plasma Diagnostics

In contrast to other disciplines, such as engineering, chemistry and biology, recommended protocols for standard measurement techniques are less developed in plasma science, and for LTPs in particular. This situation is largely due to the challenging and strongly diverse nature of LTPs, often requiring diagnostic techniques specifically adapted to the distinct plasma environment and particular plasma parameters of interest. On the one hand, this makes the development of standards for LTP diagnostics a complex endeavor. On the other hand, it is important to understand the limitations of different diagnostic techniques, associated measurement regimes and uncertainties. The development of standards can provide clear benefits in

1
2
3 transparency, reliability and transferability of experimental measurements. At the same time,
4 standards should minimize additional cost in terms of additional infrastructure and time invest-
5 ment. These standards should not result in unintended constraints and generate potential barri-
6 ers for developing and deploying new diagnostics; or applying current diagnostics to new con-
7 figurations and parameter spaces. To the contrary, the development of standards should be a
8 dynamic ongoing process and support future developments.
9
10

11 Dedicated review or tutorial style articles can be a good starting point for proposing and
12 developing reference standards. This process could begin fairly narrowly in scope, addressing
13 particular diagnostic techniques used to measure selected plasma parameters in a plasma envi-
14 ronment of interest. Valuable articles already exist in the literature in this regard, and should
15 be surveyed as a basis for establishing standards. That said, review articles usually focus on
16 previous research while there is a need for survey articles with a focus on recommendations for
17 establishing standards and protocols for forward looking activities.
18
19

20 Key issues to consider in a proposed framework for diagnostic standards to support
21 transparency, reliability and transferability of experimental measurements are:
22
23

- 24 • Identifying and classifying the diagnostic technique
- 25 • The regime of validity of the technique (that may evolve over time)
- 26 • Analysis technique(s) or requirements for a range of plasma environments
- 27 • The range of equipment and potential calibration procedures
- 28 • Input data required for analysis of the data, including uncertainty assessment.
- 29 • The consequences of impurities.

30 The implementation of such diagnostic standards could be supported by cataloguing
31 community-accepted hardware and software tools and guidelines for their use, as has already
32 been realized with the COST-jet. Care should be taken in developing these catalogues and
33 recommendations for equipment since the situation here is different than with plasma sources.
34 The development of standard plasma sources has been, to date, a bottoms-up community effort
35 producing a unique, and at least initially, non-commercial device. Most sophisticated diagnos-
36 tics, and laser-based diagnostics, employ purchased commercial equipment. Standards should
37 avoid endorsing a particular company's products. Rather, the standards should recommend the
38 equipment's specifications (e.g., wavelength range, bandwidth, energy, pulse length, resolu-
39 tion). Doing so may motivate more companies to develop or offer products having these spec-
40 ifications.
41
42

43 The discussion of diagnostics in this section is intended to cover the entire diagnostic
44 infrastructure. For example, it is not only the capabilities of the laser in a LIF diagnostic that
45
46

1
2
3 requires specifications and description, but also the detectors, spectrometers and critical optics.
4 The latter would be quite important in, for example, Thomson scattering where suppression of
5 the Rayleigh scattered light is important.
6
7

8 In the absence of established standards, it is recommended that in their reporting, re-
9 searchers follow a similar general framework as described here. Doing so supports transpar-
10 ency, reliability and transferability of their results. A clear and detailed description of the equip-
11 ment used for the diagnostic, the precise implementation of the diagnostics (e.g., voltage sweep
12 speed in a probe measurement, repetition rate for a laser), the input data to the analysis and the
13 analysis techniques should be provided. One example of such a framework is the initial draft
14 of the Plasma Metadata Schema [45]. Although this discussion has emphasized processing of
15 diagnostic data, a key element to transparency is to also provide access to the raw data, which
16 would allow potential reanalyzing of data using different techniques or when analysis tech-
17 niques are improved. Enabling access to raw data is already strongly encouraged or required
18 by several international funding agencies and journals.
19
20

21 **D. Links to other sections for plasma standards**

22 There are close links between diagnostics standards with standards discussed in other
23 sections in this article. Standard plasma sources (Section II) require reliable diagnostics for
24 accurate characterization. In turn, well-characterized standard sources provide an ideal plat-
25 form for the development of new diagnostic techniques as well as benchmarking of different
26 diagnostic techniques against each other. Benchmarking is a particularly powerful tool when
27 different techniques rely on different analyses techniques or assumptions and use different input
28 data. Benchmarking can provide an uncertainty envelope and identify weaknesses in model
29 assumptions and input data (Secs. IV and VI). Standard plasma sources and accurate charac-
30 terization using reliable diagnostics are also key in the development of potential concepts for
31 developing a plasma dose (Section IX). Data obtained from measurements should generally be
32 made openly accessible, including raw data (Section VIII). Analysis techniques for diagnostic
33 measurements should be transparent and ideally based on open source analysis codes (Section
34 VII).
35
36

37 **IV. Data and Mechanisms**

38 Over the past decade, the topic of input data for computational models and simulations
39 of LTPs has taken on greater importance and experienced increased activity within the LTP
40 community. The community has realized that reliable and validated data is critical to the use
41
42

of models in both investigating processes in established parameter spaces and extending models into yet-to-be-experimentally investigated parameter spaces. This increased interest has resulted in community-driven projects such as LXCat [46,47], Phys4Entry [48] and VAMDC [49,50], which recently celebrated their tenth anniversaries, and in commercial offerings such as Quantemol's QDB [51]. Another indication of the importance of the topic is the level of activity at dedicated conferences such as the International Conference on Atomic and Molecular Data and Their Applications (ICAMDATA), International Conference on Data Driven Plasma Science (ICDDPS) [52] and the recurring topical sessions on plasma data in more broadly attended meetings such as the Gaseous Electronics Conference (GEC) [53]. The importance of input data has also been acknowledged by publishers. It is becoming more common that scientific papers provide a comprehensive listing of all input data that have been used for a particular investigation. One such example is in the form of digital auxiliary data files hosted on the publishers' websites or in third-party repositories such as Zenodo [54, 55, 56].

Input data have always been crucial ingredients in plasma research. The recent change in that emphasis is that the focus is no longer solely on the data but also on formats for exchange of data, and reliable and convenient web-based dissemination protocols. Examples of projects whose goals are, in part, to provide such access to atomic and molecular data are VAMDC and the XSAMS document format [57] upon which it is based. Another example is INPTDAT [58], which has a broad scope and promotes the FAIR (Findability, Accessibility, Interoperability, and Reuse) [59] data principles for scientific communication. LXCat is another example for data dissemination [47]. LXCat has become a de facto standard for cross section data, in particular for electron-impact processes, and for standardizing input file format.

Is this realization of reliable communication channels the predecessor of a standardized workflow in LTP physics? It definitely is a prerequisite, however standardizing a workflow has other requirements as well. Adoption of a new workflow will occur only if there are demonstrable advantages for the producers and end-users of the data. Since many plasma regimes and types of sources fall under the LTP umbrella, each with particular data needs, it is not clear that a single standard will ever be able to meet the needs of the LTP community at large. Any such effort to establish LTP data standards also calls for an understanding of existing standards in adjacent fields, such as quantum chemistry, transport physics and chemical reactor engineering. This section touches on these issues, but starts with what is perhaps the most important question of all: "Why do we need standardization of data in the first place?"

1 2 3 A. The Desirability of Standards Development 4

5 The first issue that needs to be discussed is what can be gained from the development
6 of input data standards in LTP science. There are at least two aspects of data standards that
7 should be considered:
8

- 9 • Standardization of data representation and methods for data handling.
10
- 11 • The establishment of standard (reference) data sets.
12

13 Some immediate advantages of the adoption of standard file formats will first be discussed from
14 the perspective that the real challenge of standardization is in the clarification of terms and
15 labels. The complexity of a simple question such as “what is the meaning of N_2 ?” will be
16 discussed by contrasting the state-to-state approach to plasma modeling with the use of self-
17 consistent electron data sets that is common in gas discharge physics. We will consider the
18 pros and cons of establishing reference data sets.
19

20 When new plasma simulations are developed that use or develop a particular physical
21 model of the plasma, extensive testing should be done to verify the code’s correctness. This is
22 usually done by a combination of testing of individual sub-units of functionality and comparing
23 results of the full model with well-established previous results, obtained either as analytical
24 (asymptotic) solutions, experiments or from numerical calculations. In the latter case, the task
25 requires that the same input data and operational settings are used as in the reference study. In
26 doing so, any differences in the results can be attributed to differences in the algorithms used
27 in the reference work and in the new code. This type of benchmarking is good practice in
28 computational sciences. Examples of benchmarking in other fields include a collection of test
29 problems for matrix solvers (the Matrix Market [60]) and test problems for computational fluid
30 dynamics (CFD) [61]. A developer of a new code will typically test the code against the refer-
31 ence cases. Although this testing is part of the culture of the CFD community, there are no
32 requirements for such testing by the leading journals of the field.
33

34 In plasma science such rigorous benchmarking appears to be less of a standard practice,
35 at least in the published literature, but there are notable examples of such community bench-
36 marks. For example, five independently developed particle-in-cell (PIC) codes have been com-
37 pared in a detailed benchmark study by Turner et al. [62]. The Landmark project [63] provides
38 a platform for comparing low-pressure magnetized plasma codes. In a more recent study [64],
39 six codes for simulating positive streamers in air were compared. The level of detail in the
40 specification of the cases varies in these examples. However, all followed the practice of spec-
41 ifying a relatively small amount of necessary input data, so that emphasis would be on the
42 algorithms.
43

When testing a new code relies on large amounts of input data (e.g., species properties, radiative transitions, reaction mechanism), setting up the test case can become a laborious task, especially when the data have to be reconstructed from tables, graphs and informal descriptions in publications. Interpreting and collating these sources of data inevitably involves individual judgement in issues as simple as the number of data points that are extracted from a line graph, which then impacts the interpolation of that data during execution of the code. As an example, the LoKI-B code for solution of Boltzmann's equation for electron energy distributions [65, 66] (see also Section VII) has been rigorously benchmarked against the BOLSIG+ [67] code. In order to use the same data in both codes, additional information was required on the specific excited states corresponding to the electron impact excitation cross sections used in the original input files of BOLSIG+. Since the codes adopt slightly different models (e.g., to obtain an elastic cross section from effective momentum transfer cross sections) and different levels of detail in the number of excited states, the exercise of benchmarking was limited to simple cases. In another example, at the bi-annual “Non-LTE Code Workshop” (see for example Ref. [68]), harvesting and interpreting the correct input data appears to be part of the challenge that is posed to the participants. Finally, an effort [54] to simulate a CO₂ plasma with dozens of species and thousands of reactions with two different codes and comparing the results has been successful, but revealed the difficulty and intensive labor required to accurately reproduce results of previous works. This state of affairs goes against one of the key principles of science – reproducibility.

In other fields a portion of the challenge of reproducibility has been addressed by establishing (reference) data sets and distributing them in electronic form. An example from a nearby field of science is the GRI-Mech mechanism for the combustion of hydrocarbon gases [69]. The availability of the data set as a computer-parsable file not only addresses the problem of ambiguity, but also mitigates the risk of introducing typing or unit mistakes in the assembly of the test problem. In the absence of such well-established (and named) reference sets, strict testing of codes for plasmas in complex gas mixtures becomes difficult to the point that it is not frequently done, or that differences in output due to unexplained reasons are taken for granted [64].

In emphasizing the importance of reference data sets for benchmarking codes, we have not focused attention on whether the data is physically valid. If the interest is simply benchmarking codes, the best data for comparing results may in-fact be an artificial test suite. However, more interesting reference sets describe an actual physical problem for a particular parameter range, a good example being the GRI-Mech mechanism. Such a mechanism can be used

as input by an end-user who addresses more applied issues, for example a new burner design. Such separation of concerns is becoming more relevant as we experience “democratization” of computational plasma physics. Whereas in the past simulations were usually performed only by plasma-computational experts, today computations are widely performed by others, for example for engineering purposes, using commercial and non-commercial codes. This is a wonderful development – more computations are being performed by more researchers. However, with the computations being done by non-experts, there is also the risk of reduced rigor in the selection of mechanisms and data when the focus is elsewhere and if no reference data sets are a priori defined. On the other hand, the availability of ready-to-use mechanisms may hamper original research and debate about the relevance of species and reactions in the system at hand.

B. Exchange Formats

Standardizing the digitally available files for the representation of data and the methods for data handling is an important aspect for the dissemination of plasma input data. Reading such files should not be a source of confusion and errors itself. Using a standard file format such as XML or JSON relieves the users from the need of programming custom lexers or parsers, with the usual difficulties in getting all of the details correct. Even the apparently simple task of reading a two-column data set from an ASCII file, for example one that represents a cross section as a function of energy, is non-trivial if one cares about the details. Among the difficulties are the handling of local settings (e.g., is a comma a thousands separator, or a decimal sign), being prepared for DOS and Unix line endings, the handling of missing or superfluous data on a line, the identification of different blocks of data and the precision with which the data is written and read (single or double precision). Such a file cannot be automatically processed if the units cannot be reliably inferred from the document. Depending on the task at hand, many additional meta-data may (or should) be required, such as a statement on the accuracy of the data, the method by which that data have been obtained or how extrapolation should be carried out to obtain values outside the range of the table. It is also essential that references to underlying literature can be easily extracted so that proper credit is given to the authors of these publications. As the complexity of the document increases (more fields, data in logical sections, perhaps in recursive format), the complexity of a parser increases. For a typical example we refer to the file Code/Parse.m of the open source LoKI-B code [65,66], which contains the code for parsing an LXCat-style input file. In such cases the advantages of a standard structured file format become even more obvious.

XML has been adopted as the exchange format by the VAMDC consortium [49]. In the

case of LXCat [46], XML output is available as an experimental feature, while a switch to JSON is being considered as part of a reimplementation of the software stack [47]. We are not aware of resistance against the adoption of structured document formats *per se*, but acknowledge common concerns. The format should be vendor-neutral, and software libraries must be available for the most common platforms. Both XML and JSON match these criteria, though they are not the only options. It is also important for acceptance of a new format that backward compatibility is ensured, by providing software that translates new-style documents into prior formats so they can be used with existing software.

C. Standardization of State-to-State Data, XSAMS and VAMDC

The choice of a structured document format should not be confused with the choice of a *particular* format, such as XML or JSON. The importance of that latter choice is sometimes over-emphasized, since lossless conversion between files of such types can be accomplished with standard tools. It is more important that there be a formal, generally accepted specification of the contents of such a file, so the integrity of a document can be checked by a computer and a computer code can rely on particular data being available in the file. In the case of XML and JSON files, so-called "Schema" can be used for this purpose [70,71]. The real work of standardizing an input data document type reduces to the creation of the schema file for such documents.

The XML Schema for Atoms, Molecules and Solids (XSAMS) project [57] is an example of an (XML) schema for atomic and molecular data. The schema was introduced in 2009 and is supported by a consortium that involves organizations such as the International Atomic Energy Agency (IAEA) and the National Institute for Science and Technology (NIST) in the United States. XSAMS has been adopted as the primary output format by the VAMDC project, a community-driven e-science infrastructure that provides access to dozens of databases. The result is that such data can now be obtained via a uniform query interface and that the XML files that are produced have the predictable structure defined by the XSAMS schema. For a recent overview of the databases connected to the VAMDC infrastructure, refer to Table 1 of Ref. [50]. The emphasis of VAMDC is on spectroscopic data which resulted from the involvement of astronomers, for example through the International Virtual Observatory Alliance (IVOA) and the Observatoire de Paris in the project.

Collisional data are also available in VAMDC, and among the many examples are the BASECOL2012 database [72]. The VAMDC project has a formal approach to define the states of atoms and molecules. (See, for example, Ref. [73].) The format is best-suited for processes

1
2
3 involving states for which a complete set of quantum numbers is available (“state-to-state
4 data”), and the BASECOL2012 database is an excellent example of such a data set. Another
5 example of a collection that contains state-to-state data is the Phys4Entry database [48]. While
6 the number of identifiers that is needed to define a state is large, the advantage of this type of
7 data is that there is little room for semantic confusion, which is common when adopting ambig-
8 uous designations such as Ar* when referring to an excited state of argon (see below).
9
10

11 The state-to-state approach has some additional advantages. Among others, the data
12 items (such as radiative transitions and collisions) are elementary, and as a result also context-
13 free. This means that, for example, data on collisions in a hydrogen plasma can be combined
14 with data on a nitrogen plasma when a hydrogen-nitrogen mixture is simulated. Only the
15 “cross-terms” that involve both hydrogen and nitrogen species must then still be added. Data
16 obtained from calculations are usually of this state-to-state type. Depending on the application,
17 vast amounts of such data might be needed, for example for the detailed representation of an
18 emission spectrum or for the calculation of the thermodynamic properties of a plasma in a mo-
19 lecular gas in Local Thermodynamic Equilibrium (LTE), such as partition sums and specific
20 heat. The detailed information on the states and the systematic representation of the data in a
21 standardized form such as that offered by VAMDC then facilitates the implementation of such
22 tasks in a computer code. Examples of codes that can operate with VAMDC are SPECVIEW
23 [74] and CASSIS [75], both spectroscopic tools. For more software using the VAMDC schema,
24 see the list in Ref. [76]. However, the level of detail enabled by the state-to-state description
25 and the subsequent effort in handling large amounts of such data might be excessive when
26 simpler models suffice. Standard input data formats, referring to different reference data sets,
27 should be able to accommodate the different needs of the users.
28
29

43 D. Mechanisms

44 Reality is often not as elegant as the state-to-state approach suggests. In LTP physics,
45 cross sections may have been obtained from drift tube experiments and may describe more than
46 one elementary process. For example, the excited states of a species may be lumped into a
47 single species such as Ar* and an atomic argon plasma is then represented as mixture with
48 species Ar, Ar*, Ar+ and electrons, for example. [Lumped species, or multiple lumped species
49 (Ar*, Ar**, Ar***) may be necessary to reduce the complexity of a simulation to a manageable
50 level.] The energy levels and the excitation cross sections in such a model may be the outcome
51 of fitting experimental data within a theoretical framework such as a two-term spherical har-
52 monic solution of Boltzmann's equation for the electron energy distribution [65, 66, 67]. The
53
54
55
56
57
58
59
60

cross sections are tuned such that experimental values of drift and diffusion coefficients and inelastic rate coefficients are accurately reproduced by the model. Such cross section data are not elementary in the sense that they do not describe individual transitions. As a consequence, such swarm-derived data sets should be combined with other data with caution. That said, momentum transfer cross sections and transport coefficients such as for mobility and diffusivity for many species are only available by unfolding swarm data in this manner. For more details refer to section 5.4 of Ref. [47].

By using lumped-states or effective-states such as Ar^* , the amount of data that is needed by a model in order to get accurate results for a plasma is smaller than a complete state-to-state model. The cost is semantic confusion. For example, what is the meaning of Ar^* or N_2 ? Does a species with the name N_2 include all the states of this molecule or only the rovibrational states of the electronic ground state? The specific intent of a species named Ar^* or N_2 should be made clear in the publication of results. However, the same species Ar^* or N_2 could be defined differently in different papers. Removing such ambiguity could be accomplished by a community accepted definition of Ar^* or N_2 and capturing the definitions in a standard. Ideally, a lumped state such as Ar^* would be defined as the combination of specific, spectroscopically defined states, along with the method of combining those states and their electron impact cross sections. This is particularly important for super-elastic and quenching coefficients. The current practice is to typically add the electron impact cross sections of the individual states for excitation of the lumped state, but use a single cross section for quenching since there is only the single state. This practice has not rigorously been tested nor standardized. In particular, what is the degeneracy for a lumped state, an important consideration in computing super-elastic cross sections? In the LXCat community experiments are going on with a data model that accommodates these types of lumped state data [47]. The VAMDC community appears to acknowledge this issue (see the discussion on “fuzzy matches” in section 3.3.2 of Ref. [50]).

The complete data sets for electron-impact cross sections that are offered by LXCat, called *xs-sets* for brevity here, follow an ontology similar to the reaction mechanisms that appear in other fields, like GRI-Mech in combustion science. Reaction mechanisms are collections of data (reactions, rate coefficients, thermodynamic quantities) that produce good results for a reasonable computational price. Reaction mechanisms are not universally applicable and so the parameter spaces for which they are valid should be specified as part of the mechanism. For example, in combustion this usually means a range of gas temperatures, pressures and initial gas compositions. In the case of the *xs-sets* offered by LXCat, the cross sections alone are valid

for describing the electron kinetics of gases or mixtures of gases where there is negligible interactions between the gases. When there are interactions between gases, the *xs-sets* can be included as part of more complex plasma models with a full reaction mechanism to address applications at higher gas pressures and temperatures, and under non-LTE conditions. Only a few of such reaction schemes are available publicly in electronic form, and then often only in custom file formats [54, 56]. Many mechanisms that are of immediate use are also offered commercially by QDB [51], also in a non-standard format such as qdat, and formats that enable use by a small selection of non-commercial codes, such as HPEM [77].

E. Conclusions

We have discussed the motivation and possible process for developing a standard for plasma input data, and discussed that the data needs of the astronomical, high energy density, LTE and LTP groups of the plasma physics community are quite different. A common standard seems worth the effort, but requires a data model that accommodates state-to-state plasma data as well as mechanisms. Such a standard should not ignore the existing tools and data-formats, particularly in cases where they have wide acceptance. The adoption of a standard or standards will strongly depend on the advantages that are experienced by users. Such advantages can come in the form of applications or postprocessing facilities that operate on the data, such as the spectral software that is available for VAMDC or Boltzmann solvers that can work with LXCat data. Simplified search and re-use of relevant data (e.g., for validation of models, comparison of results, or conducting meta studies) may lead to broader acceptance of the FAIR data principles and sharing of digital data sets with comprehensive and standardized data descriptions, one example being that intended by the data platform INPTDAT. This would open up new possibilities for linking data to associated plasma sources, applied diagnostics and software packages used. In this discussion, it is important to separate the technical aspects such as using standard file formats from the discussion on the content of standard and reference data sets.

V. (Quick) Data Generation

Models of LTPs rely on knowledge of the various chemical processes that can occur in the plasma. Such processes are broadly of two types: the interaction of electrons with atoms and molecules, and heavy particle collisions which give rise to chemical reactions and related processes. Depending on the plasma chamber, processes which occur on surfaces may also be an important driver of the overall composition of the bulk plasma. Indeed, these surface processes are often the reason for studying the plasma in the first place. As constructing a full

1
2
3 chemical model for a plasma can require a significant amount of data for a range of different
4 processes, theoretical methods are increasingly the means for providing such data [78]. Data
5 can be sourced from databases (see Section IV) but often it is necessary to complete datasets
6 by generating extra data using quick but approximate procedures.
7
8

9 These quick methods should also enter into the discussion of standards and best practice.
10 Most complex plasma chemistry models, by necessity, make approximations, analogies and
11 estimates to fill in values for processes for which there are no experimental or computed data
12 for cross sections, mobilities and transition probabilities. In reporting on these models, the
13 rationale for making these approximations should be discussed. That said, a community con-
14 sensus on recommended practices to fill in the missing data would enhance our ability to assess
15 the results of the models, to understand the rationale for making these approximations and to
16 lend systematic consistency between models. This section sketches some of the procedures to
17 quickly produce these data, available starting from the simplest and moving to more sophisti-
18 cated and computationally demanding methods.
19
20

21 A. Scaling laws

22 Perhaps the simplest means of generating new data quickly is the use of scaling laws.
23 Scaling laws are widely used in plasma physics and in the context of constructing chemical
24 networks, scaling laws can be performed based on a number of different properties. For exam-
25 ple, one can use the mass of the species as a scale factor which has a particular use for fusion
26 plasmas. Data are often available for processes involving H and sometimes for D, but are rarely
27 available for T because of the extreme difficulties of performing experiments with this radio-
28 active species. Scaling laws for this situation have recently been proposed by Belli *et al.* [79].
29 Similarly scaling laws for vibrationally-resolved molecular processes have been developed
30 based on vibrational quantum numbers [80]. Other possible parameters that one can scale on
31 are ionization stages for heavily ionized atoms and reaction classes where species come from
32 the same group in the periodic table. Looking to the future, in many cases it should be possible
33 to use the techniques of machine learning to generate new reaction rates using the ideas behind
34 scaling laws but allowing for greater freedom in the parameters choices. This provides the
35 possibility that given an appropriate training set, ML would provide a proper range of parame-
36 ters to generate estimates for reactions whose rates are not known.
37
38

39 B. Electron-driven Processes

40 There are simple or relatively simple formulae that provide rates for a number of phys-
41 ical processes that are important in plasmas. Perhaps the simplest of these are for electron
42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60

1
2
3 impact rotational excitation of molecules which contain permanent dipoles. In this case there
4 are standard analytic formulae for the cases of both neutral [81] and ionized molecules [82].
5 These formulae use the Born approximation which assumes that the interactions are all essen-
6 tially long range. They are found to work well for molecules for which the permanent dipole
7 moment, which dominates long-range interactions, is large with, for example, a value of 2 De-
8 bye or more. A similar formula is available for dipole-driven electron impact vibrational ex-
9 citation of molecules [83]. However, this form is only appropriate for cases where the incoming
10 electron does not get trapped in a long-lived, quasibound resonance state. As such resonances
11 can increase vibrational excitation rates by many orders of magnitude the Born approximation
12 formula for vibrational excitation should be applied with care. In fact, resonances also provide
13 the route to dissociative electron attachment (DEA) which is a key process for the formation of
14 negative ions in molecular plasmas. A procedure to estimate DEA rates is available [84] alt-
15 though this requires estimates of resonance parameters as part of the input.
16
17

18 The Born approximation can be used to estimate electron impact electronic excitation
19 cross sections for so-called optically allowed transitions which are the ones driven by dipoles
20 and are the only excitation cross sections which are important at high impact energies. The so-
21 called BEf method involves use of a scaled version plane-wave Born approximate where the f
22 in the acronym represents the oscillator strength of the corresponding optically allowed transi-
23 tion [85]. The Bef method becomes increasingly accurate for higher energy collisions where
24 other, short-range interactions become increasingly less important and electronic excitation pro-
25 cesses which do not depend on dipoles are not important.
26
27

28 Another higher energy process, electron impact ionization, has been the subject of study
29 by a variety of different approximate and quick to apply methods. The most widely used of
30 these is the Binary Encounter Born (BEB) method of Kim and Rudd [86] which has been found
31 to perform reliably for a whole range of species [87]. Work is now currently focusing on how
32 BEB, or indeed any of the other related methods, can be extended to predict fragmentation
33 patterns for the resulting ionized species [88]. In a similar vein the Drawin approximation [89,
34 90] is used to give electronic excitation (bound-bound) and ionization (bound-free) rates for
35 electron collisions with atoms although this method often gives an overestimate requiring the
36 use of appropriate scaling factors to give useful results.
37
38

39 **C. Heavy Particle Collisions**

40 It is common to represent the rate coefficients for reactive collisions in Arrhenius form.
41 This form is designed to capture the physical behavior of most chemical reactions in that they
42
43

1
2
3 have a barrier to reaction that must be overcome in order for the reaction to proceed. However,
4 there are classes of reactions, particularly ones involving the reactions of ions with neutrals, for
5 which there is no barrier to reaction. The formation of H_3^+ in partially ionized hydrogen gas by
6 the process $H_2 + H_2^+ \rightarrow H_3^+ + H$ is a good example of such a process. These reactions often
7 proceed with every collision which corresponds to the Langevin rate of the process, something
8 that can be easily calculated for a given set of conditions. The Langevin rate therefore provides
9 a good starting approximation for these fast processes. In the case of the reactions with a bar-
10 rier, transition state theory [91] provides a relatively simple method obtaining reaction rates
11 without need to study the dynamics of the atoms involved in the reaction.
12
13
14
15
16
17
18
19

20 VI. Verification & Validation (V&V)

21

22 Computer models of LTPs must meet at least two requirements to faithfully represent
23 an experiment or investigate an experimentally unexplored region. The first is the selection of
24 the proper formulation of the physics, which in LTPs would also include the proper reaction
25 mechanism. For this discussion, we will call the formulation of the physics and the reaction
26 mechanism the *equations*. The equations are transformed into computer code using mathemati-
27 cal techniques such as discretization on numerical meshes and integrated using, for example,
28 linear algebra routines. The process of "validation" is determining whether we have chosen the
29 proper equations (including the reaction mechanism) to represent the physical phenomena. The
30 process of "verification" is determining whether we have solved the equations properly. No
31 method in general use can prove that a computer program is free from error. The challenge,
32 therefore, is to demonstrate that the results of a scientific enquiry are valid in spite of this limi-
33 tation. The formal process of removing errors is a large part of verification.
34
35
36
37
38
39
40
41
42

43 Investigations of source codes for various technical applications have shown a variable
44 level of success in the pursuit of verification, ranging from as many as one fault every ten lines
45 to as few as one in every ten thousand lines [92]. From these observations, we might conclude
46 that careful testing can reduce the frequency of errors by about a factor of a thousand. This
47 reduction in the frequency of errors usually comes at a high price. The difference in the devel-
48 opment time (and hence cost) between these extremes may exceed a factor of ten. So careful
49 verification is expensive.
50
51
52
53
54

55 In recent practice, different fields have approached the verification challenge with dif-
56 ferent emphases. Engineering practice leans to a formal approach, with strong methodological
57 prescriptions often imposed, for example, by editorial policy of journals [93]. Scientific com-
58 munities have favored a less structured approach, which is sometimes characterized as allowing
59
60

a major role for “expert judgement” as a correctness criterion for computer simulations [93]. Suggestions are sometimes made that this approach reflects a lax attitude to verification on the part of some scientific communities. However, the aims of scientific and engineering calculations can be very different, and we should not ignore these differences. Allocating the very large resources required for serious verification may not make sense unless the benefit is clear. So why might scientists find the benefits of verification unclear?

In scientific work, a computer simulation is often a tool for exploring the qualitative behavior of a physical system. The aim of the work is not the development of the computer simulation, but rather the articulation of a higher-level understanding, often expressed as an analytical theory. A classic expression of this mode of work is found in the Fermi-Pasta-Ulam-Tsinghou problem, which began as a computer investigation of a model nonlinear system, but motivated a vastly fruitful field of enquiry into the general behavior of nonlinear systems, including the evolution of unprecedented mathematical tools [94]. In this and similar scenarios, the eventual outcomes are hardly dependent on the correctness in detail of the original computer simulations, which needed only qualitative validity. Further investment in verification would have added little value. A counter argument to this view is made by the requirements of scientific reproducibility. That is, there is the expectation that the procedure of a scientific study is sufficiently documented that other researchers can reproduce the results. There is a risk that an insufficiently verified code does not actually do what the authors claim that it does. If this is the case, the results are irreproducible with there also being considerable confusion (and an associated waste of resources). This objection is much mitigated if the source code and other relevant data are published, but this is not often the case.

Very different situations arise in engineering [93]. Frequently, the aim of an engineering calculation is to calculate a number (the drag coefficient of a body, for example). Expert judgement cannot determine whether the number is correct or not, but crucial (and costly) decisions will be dependent on the number. In this context, formal verification and validation is the only defense available against incorrect decisions based on faulty numbers. The LTP physics community encompasses rather mature sub-fields, where the engineering approach to verification is likely appropriate, and exploratory areas, where a less formal approach may often be more efficient, so general prescriptions are difficult to find or define.

These extreme examples show that heavy investment in verification and validation is not always the best use of scarce resources. However, they also highlight the differences between these scenarios. A code developed for scientific exploration with little attention paid to verification is unlikely to be useful for engineering prediction, and should be used with caution

1
2
3 in that context. The traditional practices of the scientific community for computations to pro-
4 vide qualitative predictions may not be adequate when the goal is to provide a consequential
5 value or design. In this case, to some degree, the formal machinery of verification, validation,
6 and uncertainty quantification becomes essential.
7
8

9
10 The responsibility of scientific investigators is to understand the role of computer sim-
11 ulations in their chain of argument. In this regard, an important question to ask is by how much
12 could the results of the computer simulations be different without invalidating the conclusions?
13 How much assurance is there that the results of the simulations meet this criterion? These
14 considerations will often lead to the conclusion that some attention to questions of verification
15 and validation is warranted. In that case, what techniques are available?
16
17

18 Verification and validation are distinct concepts [93]. Verification is concerned with
19 the correctness of a computer code, that is, with demonstrating that the code works as intended.
20 Validation is concerned with the correctness of physical models. Demonstrating validity re-
21 quires some form of comparison with experiments, but verification does not. Clearly, then,
22 verification must come first.
23
24

25 The canonical method of verification is to show that the computer code can reproduce
26 an exact solution of the underlying mathematical model [93]. This comparison is not straight-
27 forward when both the “exact” solution and the computed solution are calculated with finite
28 precision. One method of quantifying verification is to measure the change in the distance
29 between the computed solution and the exact solution when some numerical parameter is var-
30 ied, such as mesh spacing or order of integration. This approach assumes that we can express
31 the numerical error in the computed solution as a polynomial in the parameter, so we have an
32 expectation as to how this distance should vary. This expectation will be realized only if the
33 computed solution converges towards the exact solution at the expected rate. An instance of
34 this technique is shown in Figure 5. An objection to this procedure is that in any case where
35 we know an exact solution, we might not be interested in verifying a numerical method. How-
36 ever, the Method of Manufactured Solutions shows how an essentially unphysical solution can
37 be constructed for verification purposes [93]. The premise of the Method of Manufactured
38 Solutions is that verification test solutions need not have direct physical relevance.
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

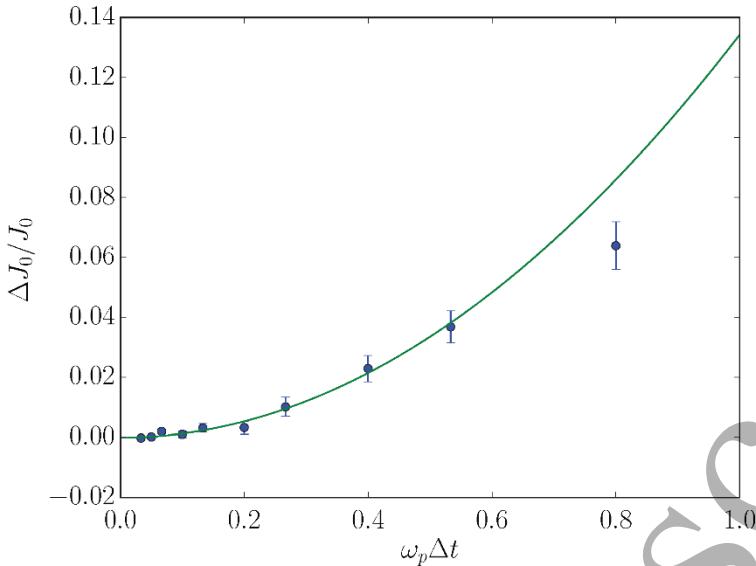


Figure 5: Verification by convergence towards an exact solution as a numerical parameter is changed. In this example, the physical parameter is the conduction current density flowing in a thermionic diode, and the numerical parameter is the time step in a particle-in-cell simulation. The points with error bars are simulation data, and the curve indicates the expected rate of convergence, which is $O(\Delta t^2)$ in this case. Reproduced from Ref. [95].

A corollary of this insight is that useful exact solutions may be found in unexpected places, for example in literature apparently remote from LTP science [95]. For some communities, such as CFD, this combination of techniques appears completely satisfactory. For LTP science, however, challenges remain, in that there are important classes of simulation where these techniques are difficult to apply. For instance, Monte Carlo methods (such as particle-in-cell simulations) often mix statistical effects with other numerical phenomena in a complex fashion, and are for other technical reasons resistant to the Method of Manufactured Solutions [93]. Hybrid models typically use a non-standard mathematical structure, and may contain modules that use different algorithms [96]. Consequently, writing down a single coherent mathematical model for the purposes of verification of the entire code is difficult. For these and other reasons, the canonical methods of verification developed by the CFD community do not solve all the problems faced by LTP scientists. For example, establishing clear mathematical foundations for hybrid models does not appear at all easy, yet it is a prerequisite for formal verification. This does not mean that no progress can be made. For example, one can apply verification tools to individual modules of a hybrid model. However, a general verification framework for the central computational tools of LTP science is not yet available, and appears to present highly non-trivial challenges [95, 97].

An alternative procedure when formal verification methods cannot be applied is benchmarking, which is comparison of different computer codes applied to the same problem [62,

64, 93, 98, 99]. The difficulty with this procedure is that when the results do not agree (which is practically always the situation) there is no way of determining which solution is “correct” if indeed any of them is correct [93]. This problem is again especially acute in the case of hybrid models, where there is not an agreed upon mathematical structure, and there is considerable room for variation in, for example, the choice of boundary conditions and the method of computing transport coefficients and injecting them into the moment equations. Consequently, differences between hybrid codes applied to a benchmark problem will usually arise from a mixture of error and the effects of legitimate alternative choices. A benchmark comparison cannot by itself disentangle these effects. The characteristics of the benchmark problem may also be important. For example, when the benchmark conditions exhibit physical instability, the computed solutions may depend appreciably on factors that are hard to control, such as implementation details like random number generators and aspects of parallel execution environments [99]. One might argue that a benchmark problem with such features is unsuitable, but a natural attraction towards conditions with obvious practical relevance weakens this objection.

Verification provides evidence that a computer simulation is correct, in the sense that the computer program delivers valid solutions to the underlying mathematical model, with a degree of uncertainty that can be characterized by investigating the influence of relevant numerical parameters. This is important, but does not introduce any evidence that the mathematical model appropriately represents any particular experimental situation. This can be demonstrated only by comparison of model calculations with experiments [93]. This process is known as “validation”. Excellent examples of this procedure are found in that part of the LTP science community concerned with the measurement and calculation of transport coefficients [100]. In careful comparisons, error bars are associated with both the experimental and computed data, where the errors in the computations arise both from numerical effects and from the presence of parameters with uncertain values. Uncertain parameters are treated by methods of “uncertainty quantification.” The simplest procedure is to employ a Monte Carlo method to vary the uncertain parameters according to a suitable distribution [101]. This is a simple and direct approach, but certainly not the only one conceivable. But uncertainty quantification is a well-understood procedure. Clearly, the desired outcome of a validation exercise is that the measurements and calculations agree within the combined error bars. When such agreement is not found, identifying the cause of the failure may not be straightforward because there are so many possibilities. These causes are not limited to erroneous calculations: Error (or misinterpretation) of the experiments needs to be considered as well.

1
2
3 These considerations show that a “one size fits all” approach to verification and validation
4 is undesirable and probably not possible. At one extreme, namely exploratory work intended
5 only to motivate future investigations, detailed verification and validation are likely an
6 inappropriate use of scarce resources. At the other, engineering prediction often absolutely
7 requires close attention to detailed questions of verification and validation. Of course, there is
8 a continuum of intermediate situations. Always, the overriding principle is fitness for purpose.
9 Investigators (and referees) should be clear about the aims of calculations, and careful to ensure
10 that those aims are delivered by suitable attention to questions about “V&V.”
11
12

13 An important nuance here is the status of computer codes that are shared, by open sourcing
14 or otherwise. In this case, the code authors do not always know how the code will be used.
15 There have been instances where codes were employed (and improper results published) by
16 code-users for parameter spaces that the developers of the code never intended. This not-good
17 outcome may result from the algorithms not extending into that parameter space (verification)
18 or the reaction mechanism not being proper in that parameter space (validation). The providers
19 of codes should make it very clear the parameter spaces in which their codes and reaction mech-
20 anisms can be reliably applied, and users should take heed of those assessments. There seems
21 in this case a responsibility on both developers and users of code to be alert to issues of verifi-
22 cation and validation, but also an opportunity to share the burden.
23
24

25 There is a difficult path to navigate for both individual investigators and community
26 actors such as editorial boards, in avoiding the extremes of imposing unreasonably and perhaps
27 impossibly stringent criteria for V&V documentation, without causing or permitting confusion
28 produced by poorly executed computations. There is wide scope for both technical progress
29 and better understanding of the broader issues.
30
31

32 **VII. Open Source and Publicly Available Codes**

33

34 Software development and computational calculations are prominent research activities
35 in LTP science, yet surprisingly the community has not been driven to define clear standards
36 for the various steps of the workflow, including the publication of results. The term *results*
37 refers to both the outcome of the computational calculations and the code used in these calcu-
38 lations, since the code is necessary to reproduce, validate and confirm (or challenge) the calcu-
39 lations and the research findings [102]. Indeed, accessing the research software, as publicly
40 available or open-source code, is not only desirable to ensure the quality standards of the pub-
41 lished material, but it can also accelerate and inspire advances in the scientific work [103],
42 especially in a small community such as LTPs.
43
44

The term “open-source code” means any computer software under a license in which the copyright holder grants users the rights to use, analyze, modify and distribute the source code to anyone and for any purpose [104]. The term “publicly available code” is used here in relation to any software that contains or is derived (in whole or in part) from “free software”, “open-source software”, “copyleft” or similar licensing and distribution models; and/or when referring to any software redistributable at no charge for the purpose of making derivative works. Open-source codes are therefore publicly available codes, and this section addresses both, yet clearly distinguishing open-source codes as a special category of publicly available codes. We will start by briefly presenting some well-known open-source and publicly available codes in LTPs (including data generating codes), then making a case for open-source codes and finalizing with some recommendations for code development.

A. Examples of Open-Source and Publicly Available Codes in LTPs

Although the LTP community has never defined a systematic route for the development of open-source codes or the sharing of simulation tools, this practice has been adopted by several members and groups for several decades. Some examples are listed below.

The problem of calculating the electron energy distribution function by solving the electron Boltzmann equation is among the subjects that most stimulated the development and sharing of codes. ELENDIF [105] (presently not available), BOLSIG+ [67], EEDF [106], BOLOS [107], and LoKI [65, 66] adopt the classical two-term approximation for solving Boltzmann’s equation [108, 109, 110]. METHES [111] is a Monte Carlo collision code. Magboltz [112] uses a multi-term expansion (to the third order) of the electron distribution function with a Monte Carlo integration technique, and MultiBolt [113, 114] is a multi-term Boltzmann equation solver. Magboltz, BOLOS, METHES, MultiBolt and LoKI are open-source. Magboltz is a Fortran code with hardcoded data. BOLOS is a Python library using an algorithm similar to that adopted in BOLSIG+. METHES, MultiBolt and LoKI are written in MATLAB. BOLSIG+, BOLOS, LoKI, METHES and MultiBolt accept input files with electron scattering cross sections obtained from the LXCat open-access website [115].

There are two very popular freeware codes to solve global 0D plasma chemistry models. GlobalKIN [44, 116], available upon request, solves a multi-zone model for plasma kinetics and plasma-liquid-surface chemistry, with electron rate coefficients calculated from a two-term spherical harmonics expansion of the electron Boltzmann equation, at E/N values provided by a circuit model or a power waveform. ZDPlasKin [117] is a computational utility for complex plasma chemistry that adopts a two-step operation. First, a preprocessor is used to translate the

list of species, reactions, and rate coefficients from a user-friendly text format into a Fortran 90 module that interfaces to an ODE solver and the BOLSIG+ Boltzmann solver. Second, the compiled code calculates the time evolution of the densities of species and the reaction rates.

Often, plasma chemistry schemes involve several hundred (thousand) reactions, with very different impacts on the model results. The LTP community has started to adopt sensitivity analysis approaches to reduce these schemes, aiming for an easier definition of “reaction mechanisms”, corresponding to sets of reactions and rate coefficients validated against benchmark experiments. The open-source C++ code PumpKin (pathway reduction method for plasma kinetic models) is a tool for the post-processing of results from 0D plasma kinetics solvers [118]. The tool, compatible with the output format of GlobalKIN and ZDPlasKin, was developed to reduce complex plasma chemistry schemes, and can also analyze the production and/or destruction mechanisms of certain species of interest.

In a different approach, GlobalKIN was extended using the expert system Quantemol-P [119], for the automatic generation of the plasma chemistry which considers all possible gas phase reactions and the likely surface reactions, from a set of atomic and molecular species specified by the user. The reactions are sorted by importance, and the chemistry set is pruned by discarding unphysical reactions and reaction data. The system has been further extended, by adding the ability to generate electron-molecule collision data using Quantemol-N [120]. Quantemol-N is designed to treat low energy electron impacts, specifically those processes which lie below the ionization threshold of the species concerned, by running fully ab initio molecular R-matrix codes [121].

In the 90s, the freeware tool SIGLO-2D, a 2D User-Friendly Model for Glow Discharge Simulation [122], was often adopted in the numerical fluid modelling of radio-frequency discharges. The tool, which is no longer available, was validated using measurements of the spatial distribution of the plasma density in the Gaseous Electronic Conference (GEC) reference cell [123, 124], as discussed in Section II.

More recently, the Hybrid Plasma Equipment Model (HPEM) [96], available on request, is a comprehensive modelling platform developed for low pressure (< a few tens of Torr) plasma processing reactors. The HPEM with a primary version in 2D and less supported version in 3D adopts the hierarchical approach of a hybrid modelling, in which different physical processes on vastly disparate timescales are addressed in compartmentalized modules (e.g., electromagnetics, electron energy transport, fluid kinetics-Poisson, plasma chemistry, surface kinetics, radiation transport), iteratively combined using time-slicing techniques. The HPEM has been applied to a variety of reactor types, for example inductively coupled plasmas, reactive

1
2
3 ion etchers, electron cyclotron resonance sources, magnetron sputter and ionized metal physical
4 vapor deposition, remote plasma-activated chemical vapor deposition, and dust particle
5 transport. The HPEM is widely used in the semiconductor industry for plasma equipment and
6 process design.
7
8

9
10 The same group that provided HPEM has also developed, validated and released the
11 modelling platform nonPDPSIM, originally written to simulate plasma display panel (PDP)
12 cells, but whose application space has grown in scope [125]. nonPDPSIM is a 2D multi-fluid
13 hydrodynamics simulator in which transport equations for all charged and neutral species and
14 Poisson's equation are integrated as a function of time on an unstructured mesh capable of
15 capturing a large dynamic range in length scale.
16
17

18 Recently, the Idaho National Laboratory (INL) has built the open-source software de-
19 velopment framework MOOSE (Multiphysics Object Oriented Simulation Environment) [126].
20 MOOSE is designed to solve highly non-linear, coupled systems of equations across various
21 areas of physics and containing multiple physical models. MOOSE adopts a rigorous and well-
22 documented development strategy, with a comprehensive set of tools for testing, so that changes
23 to the code or MOOSE-based applications are only merged into the framework when the testing
24 ensures that the changes are compatible with the applications. With this modular structure,
25 users can either use the existing applications or develop new applications based on their needs.
26 The LTP community has used this framework to model atmospheric pressure plasma-liquid
27 interactions [127, 128], using the Zapdos-CRANE open-source package that compiles the
28 Chemical ReAction NEtwork (CRANE) module [129] into the plasma transport software
29 Zapdos [130]. An electromagnetic module, implementing the full set of Maxwell's equations
30 in the MOOSE framework, is currently being developed [131].
31
32

33 Another recent addition to the LTP simulation community is SOMAFOAM [6], a fluid-
34 based LTP software platform built on the popular open-source CFD software OpenFOAM
35 [132]. Like MOOSE, it is also modular, but built on a finite volume solver as opposed to finite
36 elements. As OpenFOAM has been a widely adopted and often modified tool for a number of
37 engineering communities and computational techniques (e.g., [133, 134]), the potential exists
38 for similar such development for LTP applications as well.
39
40

41 In recent years, improved and easier access to high performance computing resources
42 has contributed to the development of quantum codes for generating cross sectional data rele-
43 vant to the modelling of LTPs.
44
45

1 2 3 B. Making a case for open-source codes 4

5 The description of LTPs often involves solving multidimensional multiscale nonlinear
6 problems, which requires considerable investment to develop incredibly complex codes, some
7 mentioned in the previous section. This is a highly demanding scenario. Considering that it
8 has been estimated that software contains approximately 1-10 errors per thousand lines of code
9 [135, 136, 137], methods are needed to alleviate the developers' work and improve the quality
10 of the final product.

11 Adopting a review process during the development of research software facilitates de-
12 bugging and verification activities (tasks which are often underestimated in their scope), con-
13 tributing to raising code quality and correctness. A review process also brings the additional
14 advantages of providing continuity of the research work and in improved knowledge transfer
15 within and between research groups. The benefits of a code review process to enhance the
16 quality of the scientific work are probably better understood as the reader tries to answer the
17 following questions: Could you rewrite the same code built some years ago, reproducing the
18 same computational results that were then published? Can you easily find the old version of
19 the code used for that publication? How many versions of the software can you find simulta-
20 neously being used in your research group? Do you know the differences between these ver-
21 sions? What is the best / most recent version to distribute to a new student? Is there any related
22 documentation? Are you confident about the set of input data to use in the simulations?

23 A single reader may have answered no to several of these questions concerning their
24 own code and their own research group. The situation is likely not better at the community
25 level. The introduction of a code review process at the research group level could significantly
26 contribute to elevating research standards. By extension, quality would also likely be improved
27 by making the codes publicly available, ideally sharing the source code, verifying its correct-
28 ness, and evolving the software as part of a collective effort. Obviously, some codes have
29 specialized requirements and should be developed with less openness. However, in many other
30 cases, a shared code is an asset for the community when it comes to transparent reporting on
31 models and procedures.

32 The standards proposed for any kind of scientific result should also apply to codes. For
33 open-source codes, the review process would naturally take place during software development,
34 at least involving close colleagues within the same research group. However, in all cases veri-
35 fication should be encouraged as part of peer review before publication of the computational
36 results. Detailed information on verification process should be included in reporting (or as a
37 minimum as part of the documentation) to facilitate confirmation of the reported findings (see

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
Section VII).

20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
5
Focusing on software alone is not enough since codes need data to produce results. There seems to be wide acceptance about the need for proper data management, one method being adopting the FAIR principles [59]. This acceptance is demonstrated by an increased activity on the FAIR topic within the LTP community, where the focus is no longer solely on the data itself but also on exchange formats and web-based dissemination protocols. This evolution deepens the premise that the conduit between codes and data (or databases) should ideally be compatible between different codes and databases. Section IV addresses the standardization of data representation and handling, a first step to ensure interoperability and reusability of codes and data.

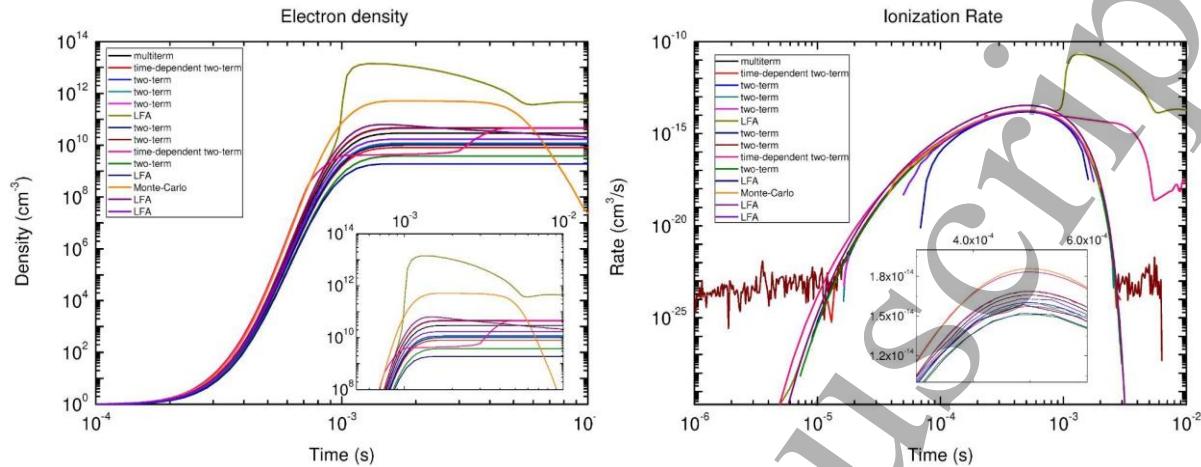
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
However, future progress could go well beyond these first steps, including collaborative development of a collection of open-source “foundational libraries” that could provide the basis for software developers to build computational tools for solving specific problems. Examples of existing collaborative frameworks close to our community are the open-source CFD software OpenFOAM [132] and the free software program Basilisk [138] for the solution of partial differential equations on adaptive Cartesian meshes. An additional advantage in adopting this type of framework is that they can be used as “docker containers”. These are open-source platforms that help a user to package an application and all its dependencies (including libraries, post-processing tools) into a container for the development and deployment of the software, thus freezing the version dependencies of the entire computational workflow.

20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
5
There are impressive advances being made using open-source modules, originally intended for other purposes, to develop LTP modeling platforms. SOMAFOAM, built on the OpenFOAM platform, has enabled computationally scalable, 2- and 3-dimensional simulations of low and high pressure LTPs [6]. Several open-source plasma models have been implemented using the MOOSE framework [126, 127, 128, 129, 130].

20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
5
The LTP community has been slow to adopt many of these practices. The reasons are due to a combination of factors. The lack of a common programming language, architecture or application dependent codes, restrictions by government and industry on distribution of codes funded by those sponsors, and lack of funding (and enormous effort) to update or rewrite codes are all barriers that would need to be overcome. Software development for engineering physics is often difficult to get funded and so research projects tend to focus on the final outcomes rather than the building of tools to reach these outcomes; in the short term.

20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
30
The LXCat workshop held at the 2016 Gaseous Electronics Conference [139], identified the need for a community wide activity on validation of plasma chemical kinetics in commonly

1
2
3
4
5
6
7
8
9
used gases, and proposed a round-robin exercise to assess the consistency in results of calculations from different participants in a simplified system. Subsequently, two rounds of exercises were attempted, in some cases revealing unanticipated disagreements in the computational calculations presented by the participants (see Figure 6).



25
26
27
28
29
30
31
32
33
34
35
Figure 6: (left) Time evolution of the electron density and (right) the ionization rate calculated by the participants of the 2017-2018 round-robin exercise for the modelling of a pure argon plasma with four species (Ar , Ar^* and Ar^+ and e) undergoing the following electron-impact collisions: elastic scattering with Ar , direct excitation, direct ionization and dielectronic recombination. The plasma is excited by applying an electric field pulse to the neutral gas at 0.1 bar pressure and 300 K temperature, for initial electron and ion densities of 1 cm^{-3} . The insert caption identifies the different model approximations that were used, where LFA refers to the “Local Field Approximation” in the Boltzmann-chemistry coupling.

36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
51
Although these differing results can in part be attributed to ill-definition of the detailed working conditions, there was also evidence of (i) different implementations of the same publicly available code and (ii) different interpretations or deployments of the physical models and the corresponding input data for in-house codes. Both had an impact on the results (see also Section IV). The sharing of codes or, at least, of the details about the numerical implementation of models (including algorithms, convergence criteria, closure / boundary conditions), could significantly improve the quality of computational predictions in LTPs. It could also nurture and support a new generation of researchers developing computational algorithms and models, a population that has continuously decreased in the last decade [140].

52 C. Recommendations for Code Development

53
54
55
56
57
58
59
In the following we provide some practical recommendations (see Figure 7) for developing research software that is open and adheres to FAIR principles [141, 142].

60
a) *Planning:* As in any other research activity, planning is essential for an effective and successful outcome. In the case of software, this involves several aspects: financial (securing

the necessary resources for the task), technical (deciding about the programming language to adopt), and scientific (formalizing the problem, preparing the numerical implementation and choosing numerical methods and algorithms). Ensuring the compatibility of the software with I/O databases and platforms relevant for the community should also be considered during the planning phase.

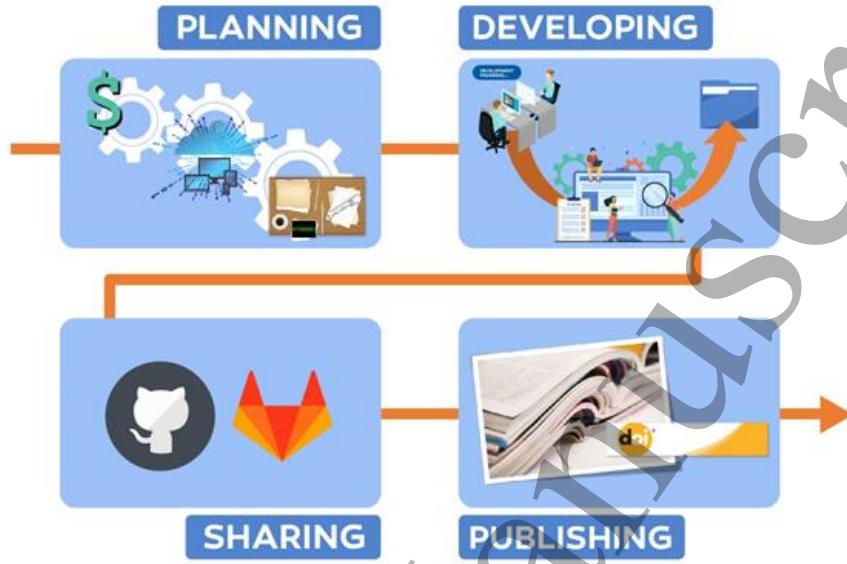


Figure 7: Schematic of recommendations for code development. Some illustrations (license free) were designed using resources from Freepik.com.

b) *Developing, reviewing and verifying:* Introducing a review process during code development should contribute to ensuring code correctness and enhancing the quality of the scientific work. Code review implemented as a discussion thread (under some versioning system like Git [143] or Mercurial [144]), with the ability to comment and suggest code changes, can work towards improving debugging outcomes, preserving code legacy, and providing a natural distribution of the workload during both development and maintenance phases.

Code review should also include verification procedures (e.g., checking that the code reproduces asymptotic limits, satisfies the conservation of quantities, and provides results similar to other codes, within numerical uncertainties) and regression tests (to confirm that new code changes do not affect the existing functionalities). More details and examples of the verification process are discussed in Section IV.

c) *Sharing:* Making codes publicly available is the next step towards software adhering to FAIR principles. Publishing the code on the research group website or, for open-source codes, making it accessible on source-code-repository-hosting platforms such as GitHub [145] or GitLab [146], should, when allowed by sponsors and government agencies, become

1
2
3 a standard practice for research software of interest to the community at large.
4

5 d) *Publishing*: Like any other research result, codes should also be published in a journal that
6 accepts software as primary submission material [147]. Moreover, even if the code is avail-
7 able in a public repository and because these web resources might not be permanent, as-
8 signing a digital object identifier to the version referenced in a published manuscript is also
9 highly recommended. This must be done carefully so that updates and corrections to codes
10 can be tracked and made available to the community.
11
12

13 **D. Outlook**

14 Many of the recommendations discussed in this section are not new. Indeed some were
15 raised decades ago [148]. However, implementing the previous recommendations is a chal-
16 lenging task that requires a culture change from researchers, publishers and funding institutions,
17 where openness is supported and sharing software is encouraged as part of the high-quality
18 standards in scientific research. The LTP community should use the main conferences in the
19 field as a forum to design and carry out collective efforts for the verification and benchmarking
20 of codes, following the recommendations listed in section VII.C, for developing research soft-
21 ware.
22
23

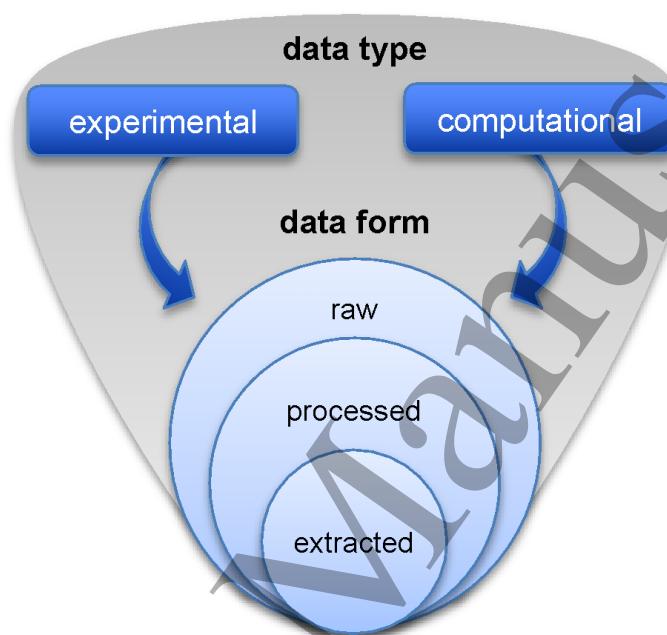
24 **VIII. Reporting**

25 Historically, reporting of data, how it has been analyzed, and how it is presented has
26 been heterogeneous for the LTP community. Most journals do not have specific standards on
27 how to report data, and when there are standards, they often vary from journal to journal. In
28 the spirit of a research enterprise that collectively advances our scientific knowledge and un-
29 derstanding of how the world works and the technology that drives it, accurately reporting data
30 in journal articles is essential. Accurately reporting data (a) makes clear the underlying infor-
31 mation, and how it was acquired, that leads to scientific conclusions, (b) enables readers to
32 assess the validity or quality of that information, and (c) empowers researchers to reproduce the
33 results to confirm said conclusions. From this perspective, reporting of data necessarily has
34 significant overlap with data availability, but takes it one step further to also include how the
35 data is analyzed, assessed, and presented in the published literature. For this reason, reporting
36 of data also often intersects with issues surrounding ethics and scientific misconduct.
37
38

39 **A. Defining Reporting and How it Applies to Plasma Research**

40 Here, we define reporting of data to broadly mean how data is presented in published
41 archival journal articles, including not only the main text but any accompanying appendices,
42
43

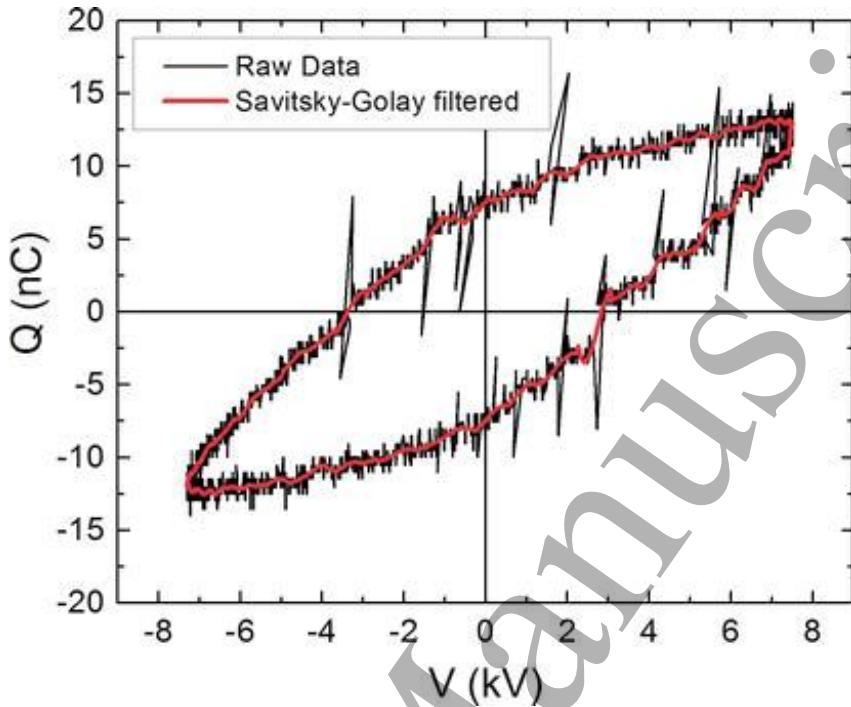
1
2
3 supporting information, or supplemental material. Here we emphasize the standards on reporting
4 of data. Standards for the dissemination of data are discussed in Section IV. Data is usually
5 presented in one of two ways – graphically or in tables – but it can also be reported as individual
6 values within the text of an article. Furthermore, the concept of data is vast, and data itself can
7 encompass almost innumerable forms. Here we classify data into two general types, experimen-
8 tal and computational, and in three general forms, raw, processed (or analyzed) and a third
9 category we term extracted data, as shown in Figure 8.
10
11
12
13
14



37
38 **Figure 8:** Different types of data and their sources and the different forms in which
39 they are reported.

40
41 The differences between experimental and computational data are fairly straightfor-
42 ward. Was the information generated from a physical process itself or by virtue of numerically
43 solving mathematical descriptions of the physical process? The differences between raw and
44 processed data are similarly straightforward. The raw data are the data as produced by the data
45 acquisition equipment. Processed data changes the raw data to make it more easily readable or
46 interpretable such as filtering spurious or low signal-to-noise-ratio data. For example, a spec-
47 trum produced via optical emission spectroscopy (OES) will include line intensities on an ab-
48 solute scale related to the number of photon counts by the detector – this is raw data. Manipu-
49 lating the data to normalize it or subtract the baseline noise, would be characterized as processed
50 data. In electrical dielectric barrier discharge (DBD) measurements, the raw current data may
51 include the capacitive current, but it is often presented ‘processed’ to only include the discharge
52 conduction current. Similarly, raw charge-voltage data used to produce a Lissajous plot may
53
54
55
56
57
58
59
60

1
2
3 be smoothed by a filter, as illustrated in Figure 9. Computational data can also be processed.
4
5 For example, in particle kinetic simulations, such particle-in-cell Monte Carlo collision (PIC-
6 MCC) simulations, raw calculated data is often averaged over several integration-times in order
7 to suppress the random statistics inherent to the method.
8
9



32 **Figure 9:** Example of filtered charge-voltage data, also known as a Lissajous plot,
33 from a dielectric barrier discharge. Reproduced from Ref. [149].
34
35

36 Both experimental and computational data can also be further analyzed to extract phys-
37 ical quantities, which adds another layer to data reporting. For example, OES spectra can be
38 analyzed to extract characteristic plasma properties including electron temperature, electron
39 density, and relevant excited state temperatures [150, 151, 152]. Similarly, DBD measurements
40 can be analyzed to extract the power deposited in the discharge [149, 153] or parameters such
41 as the number of filaments per cycle [154]. Within the context of computational data, the gov-
42 erning equations are typically solved numerically for quantities including energy or velocity
43 distribution functions, species concentrations, momenta, and energies, electric potential, and
44 magnetic and electric fields. However, other parameters can also be extracted from the com-
45 puted data, such as transport coefficients and production/destruction rates from computed dis-
46 tributions [155], and ionization and chemical conversion efficiency from the computed concen-
47 trations [156]. Critically, depending on the model adopted in constructing the governing equa-
48 tions, additional processing might be needed to extract physical quantities.
49
50
51
52
53
54
55
56
57
58
59
60

1 2 3 **B. Uncertainty, Error, and How it Relates to Reporting of Data**

4
5 An essential element of reporting data – both experimental and computational – is
6 clearly explaining the confidence that one has in the data, where confidence is defined here as
7 the statistical probability that a certain fraction of the same experiments or computations would
8 produce the same data. Conventionally, the confidence in the data is related to how accurate it
9 is relative to the *true* value that exists independent of the measurement [157]. For experimental
10 data, reporting uncertainties based on both precision (or repeatability of experiments) and bias
11 (inherent accuracy of the measurement system, instruments, and method) effects is typically
12 expected, whereas for extracted data, uncertainty propagation analysis should be used. Further,
13 for small data sets, Student's *t*-analysis should also be used. On the computational side, accuracy
14 of data is related to both the verification and validation of the computational code [158] as
15 discussed in Section VI, with emerging techniques also focusing on incorporating uncertainty
16 quantification into computational simulations [159]. Both areas have begun to receive appre-
17 ciable attention in the LTP community [101, 160, 161].
18
19

20 **C. Emerging Trends and Standards in the Broader Research Community**

21 Over the past decades, a number of research communities have raised concerns about
22 standardizing how their data are reported. In the life and social sciences, there has been signif-
23 icant focus and momentum toward more effective and accurate data reporting in an effort to
24 mitigate intentional and unintentional shading of the results to support a particular outcome, to
25 produce more reproducible and replicable research, and to eliminate bias in the scientific
26 method [162, 163, 164]. Physical sciences, such as chemistry and physics, have also seen an
27 increasing trend in this heightened awareness [165, 166, 167, 168, 169, 170], such that collec-
28 tively it has become apparent that the entire scientific community is moving toward more open-
29 ness and transparency in data [171, 172, 173]. FAIR Data Principles, where the four principles
30 are Findability, Accessibility, Interoperability, and Reusability, have, in particular, started to
31 attract widespread endorsement and acceptance [59]. An initial draft of a reporting standard
32 for data in the field of LTP has been introduced with the plasma metadata schema, Plasma-
33 MDS [45], which aims to support the implementation of the FAIR Data Principles.
34
35

36 Depending on the nature of the research and the types of data that are being reported,
37 different concerns arise. For example, in the life and social sciences, statistical analysis is par-
38 amount, and thus ensuring the statistics and data are presented well is essential. There has been
39 a noticeable trend in the graphical presentation of data to move away from simpler *x-y scatter*
40 plots, which show single points with error bars (or confidence intervals) that are often not well-
41
42

defined and ambiguously open to interpretation, to scatter plots that include the raw data with the statistical information overlaid [174]. (See, for example, Figure 10.) In the physical sciences, concerns often arise based on how extracted quantities are determined and presented, where inconsistencies within a field make it difficult to both assess the validity of the claims in the paper and to compare to prior literature. For example, in areas such as photocatalysis [167] and photovoltaics [175, 176], accurately determining and reporting parameters such as conversion efficiency and stability of new materials or designs is fraught with difficulties. Given that there are presently no standards for how these quantities should be defined, it is that much more important that reporting make clear the method that was used.

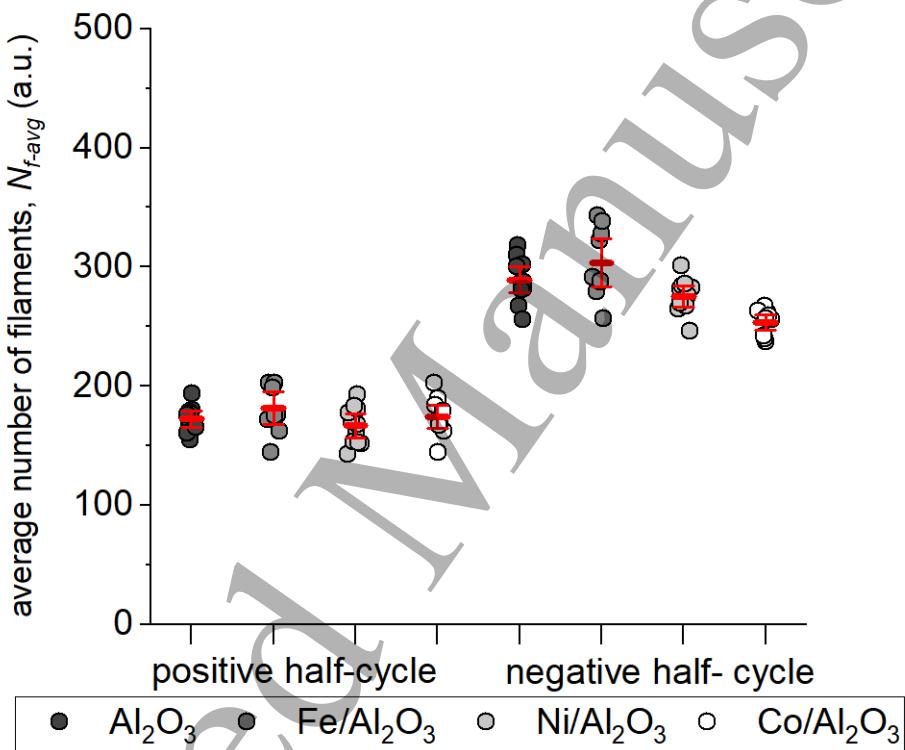


Figure 10: Example of data being presented as the raw data (points) with the statistical information overlaid (red bars for the respective averages and 95% confidence interval error bars). Data are from DBD experiments in packed beds with different materials and reflect the average number of filaments per half-cycle for different material configurations. Reproduced from Ref. [154].

While a number of editorials have been written over the past decade [177], most scientific journals do not have explicit standards on reporting of data within their papers. While policies on availability of data, such as the Institute of Physics (IOP) Publishing data availability policy [178] that governs *Plasma Sources Science and Technology*, are common for many journals, few have policies on how data must be presented and/or requirements on providing information on the nature of the data and how uncertainty was analyzed. For example, the IOP

1
2
3 policy includes unbinding rules such as “Authors are encouraged to share their data but not
4 required to. If your data will not be made publicly available, the journal requires that any data
5 required to support or replicate claims made in an article should be made available to the jour-
6 nal’s editors, reviewers and readers without undue restriction.” and “Sharing research data as
7 supplementary information files is discouraged.” [178]

8
9
10
11 There are exceptions, though. The *Science* family of journals has a set of ‘Research
12 standards’ as a part of their editorial policies, and these include specific guidance on statistical
13 analysis and how data are reported [179]. The level of detail is greater than the norm, with two
14 examples of policies being:

15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
● “Data pre-processing steps such as transformations, re-coding, re-scaling, normalization, truncation, and handling of below detectable level readings and outliers should be fully described; any removal or modification of data values must be fully acknowledged and justified.”
● “Methods used for conducting statistical tests (e.g., *t*-test, Wilcoxon signed rank test, Wald test of regression coefficient) and for constructing confidence intervals (e.g., normal-based 95% CI: mean 2SD, likelihood ratio-based interval) should be clearly stated. Mention methods used in the Materials and methods and then provide the individual test name in the figure legend for each experiment.”

50
51
52
53
54
55
56
57
58
59
60 Similarly, the American Institute of Aeronautics and Astronautics (AIAA) family of journals has an editorial policy that specifies what is expected in demonstrating both experimental and numerical accuracy when reporting data [180]. However, as noted in other fields, editorials and editorial policies do not always have the desired improvement in reporting [181, 182, 183]. To overcome the limited effectiveness, some publishers, with *Nature* being the most prominent, require authors to provide detailed information that affects data reporting in the form of checklists that must be included at the time of journal submission [184], with more stringent requirements for specific areas such as photovoltaics [185] and reports of lasing [186].

49 50 D. Opportunities and Challenges for the Plasma Science Community

51
52
53
54
55
56
57
58
59
60 The plasma science community as a whole is incredibly heterogeneous, with the low-temperature, non-equilibrium (non-thermal) community quite distinct from the high-energy and density (thermal) community. Even if just narrowing to LTPs, the types of data are incredibly vast because the field itself is so diverse. With much of current LTP development being application-oriented, data is often focused on a specific application. Even considering measuring only classical plasma properties, there is a wide variety of intrusive and non-intrusive electrical

1
2
3 and optical diagnostics that could be used (as discussed in Section III). It would be challenging
4 to define standards on data reporting for all but a few of the most widely used techniques (e.g.,
5 optical emission spectroscopy or Langmuir probes). As many of the desired quantities are not
6 measured directly and require extracting information from directly measured data, it is difficult
7 to develop standards that encompass the entire field. A good example of this, for example, is
8 the measurement of delivery of reactive species to a substrate, where the community has not
9 formed a consensus for the definition of ‘dose’, as discussed in Section IX.
10
11

12 Computational efforts have similar challenges, as many different research groups utilize
13 their own in-house codes, whether those be particle- or fluid-based. There is no clear roadmap
14 for the verification and benchmarking of these codes, and no consensus on how to validate these
15 codes against an experimental standard, although recently there have been efforts to standardize
16 these [62, 95, 97, 187, 188] (and Section VI for more information). Given it is clear that a one-
17 size-fits-all set of reporting standards is not possible, it is instructive to look where there are
18 opportunities to standardize reporting for the LTP community.
19
20

21 The purpose of reporting new research findings is to *advance our collective understanding*
22 and *knowledge about natural and physical phenomena*. Essential to this goal is the ability
23 for others to confirm reported findings. Reporting data (as supplementary material, eventually
24 archived in well-established repositories with persistent web links) and how they were *deter-
25 mined* and *analyzed* as transparently as possible should be an essential element of scientific
26 reporting in archival journal articles. Data accessibility recommendations, such as the afore-
27 mentioned FAIR data principles, are therefore important and complementary to reporting stand-
28 ards.
29
30

31 These principles extend to detailed descriptions of the methods and techniques used,
32 including any experimental instruments that introduce inherent bias uncertainty. For example,
33 if spectroscopy results are presented in any form – either as raw data or results extracted from
34 raw data – then the resolution of the spectrometer, the integration time, the grating, the calibration
35 process, and other pertinent device specifications should be included in the paper. Infor-
36 mation about grid resolution, time steps, and convergence criteria in computational simulations
37 similarly should be included. Being explicit about statistical analysis of experimental data,
38 verification of computational simulations, and similar data analysis techniques should be stand-
39 ards in journal articles or supplementary material. Figures and their captions should be presented
40 *unambiguously*. That is, it should be clear what is being presented, in what units, and what the
41 uncertainties are so that the conclusions drawn from the figure are well informed. Ultimately,
42 these recommendations are all in alignment with the principle of making data reusable, and
43
44

1
2
3 resources such as the plasma metadata schema Plasma-MDS [45] that provide a framework for
4 including a wide variety of information should become commonplace in the plasma community.
5
6

7 In areas of plasma science and engineering that overlap with other fields where there
8 are established standards, those standards should be followed. For example, plasma catalysis
9 is a rapidly expanding field that generally applies a plasma to a catalyzed reaction. Existing
10 standards for reporting catalytic performance [166, 189, 190], such as using turnover frequency
11 to report the intrinsic catalyst activity, should be adopted when possible. When those standards
12 need to be modified, such as to distinguish between plasma-only and plasma-catalyst effects
13 [191], that should be done in a way that is transparent and clear to the reader. Similar principles
14 can be applied to other emerging areas such as plasma medicine, plasma agriculture, and plasma
15 combustion.
16
17

18 It is critical to reference and report information from original sources. For example, in
19 many plasma simulation codes – both fluid and particle – energy-dependent collision-cross sec-
20 tions can be extracted from data repository websites, such as LXCat (the Plasma Data Exchange
21 Project) [115]. However, when reporting usage of information, authors should point readers to
22 the *original* source, and not the repository. Apart from the due recognition, the reason is also
23 to ensure findability and accessibility of data, since online material hosted at web addresses
24 have no guarantee of perpetuity. Similar sentiments hold for review articles that collect infor-
25 mation but are not the original source of data. Like in several other ethical issues above, the
26 responsibility for complying with guidelines lies with all parties involved – authors, managers
27 of data repositories, journal publishers, editors, and reviewers. Indeed, these obligations should
28 be clearly stated in the rules for referencing of the repositories (e.g., for LXCat [192]), in the
29 policy of the journals about data citation (e.g., for IOP journals [193]), and also in the guidelines
30 for the reviewers [194], who are empowered to scrutinize the material submitted.
31
32

33 The goals of transparency and reproducibility are difficult to argue against, and so the
34 process adopted to achieve these goals cannot be overpowering and burdensome. If so, the
35 process will not be adopted, and these goals will not be met. In the end, judgment may prevail.
36 Some of the most defining discoveries of the LTP field were based on what today we might
37 consider single point measurements made in a single experiment. Being able to accommodate
38 new discoveries and innovation, while being transparent within a system of standards and re-
39 quirements is vital to the advancement of the field.
40
41

IX. Plasma Dose

Plasma dose is a measure of plasma produced reactivity. Plasma dose is a concept that is intended to minimize uncertainties in comparing the results between experiments and laboratories when, for example, treating surfaces with different plasma devices. If the surfaces are treated with the same plasma dose, then in principle the activating species delivered to the surface are the same. With the same activating species delivered to the surface, attention can focus on the consequences of that treatment on the surfaces. There have been several proposed definitions for what plasma dose implies, most of which have originated in plasma medicine [195, 196, 197, 198, 199]. These concepts attempt to define plasma dose as the amount of exposure of a biological sample by a particular plasma device to produce a specific biological outcome. For example, plasma dose might define the amount of exposure of an *E.coli* culture to an argon plasma-jet that is required to reduce the viable population by a given amount.

The use of plasma dose (PD) for plasma medicine is motivated by the use of doses that are standard in radiation chemistry [200, 201, 202]. Here, a dose is a given fluence (time integration of flux) of x-rays or a given energy deposition by electron or proton beams. The successful use of dose in radiation chemistry is, in part, a result of the delivery of the activation energy being essentially independent of the final application and of the ambient environment. This independence results from the activation energy – the electron or x-ray beam – being generated by a stand-alone accelerator with there being essentially no feedback from the object being treated. The use of dose in radiation chemistry is typically separate from an assessment of the outcome of using that dose.

There is a tendency in plasma-biomedical applications to associate PD with a particular biological outcome. In spite of its utility, defining PD in terms of a biological outcome is perhaps too high-a-bar to pass in defining dose – at least at this early stage of discussion. Such a definition would require consistency and standards in how the sample is prepared (e.g., initial density of colony-forming-units (CFU), type and depth of culture medium), the geometry of the setup (plasma source-to-sample distance, gas residence time, vortexing or recirculation of gas), environmental factors (ambient humidity and sample temperature) and method of measuring surviving cells, to name only a few. Having a standard that describes all of these properties for the purpose of calibrating a protocol or set of equipment would be extremely helpful. However, that is a different issue than PD.

In the context of this discussion, we describe plasma dose in less ambitious terms by not associating PD with a particular outcome. Doing so would tie PD to a specific application which then would require additional standards for how that application is configured. To be as

universal as possible, PD should be as independent of the final application as possible, in analogy to the practice in radiation chemistry.

For the purpose of this discussion, PD is defined as the net fluence (time integral of flux) of a reactive species (or sum of species) delivered by a plasma device through a conceptual surface. The plasma dose would then be defined for individual species [Ar ions – PD(Ar^+); or ozone – PD(O_3),] or collectively for a set of species [reactive oxygen species – PD(ROS); or oxidation-reduction potential PD(ORP)]. PD requires units and so for the purpose of discussion, we define the unit of PD as being the fluence that would expose, on the average, each site on a surface to one reactive species. With the surface site density approximately 10^{15} cm^{-2} , the unit of plasma dose, PD_u, would be 10^{15} cm^{-2} . This could be achieved by delivering a flux of $10^{15} \text{ cm}^{-2}\text{s}^{-1}$ for 1 s, or a flux of 10^{14} cm^{-2} for 10 s. So for this discussion, 1 PD_u(O_3) is delivery of 10^{15} cm^{-2} O_3 molecules over a specified period of time.

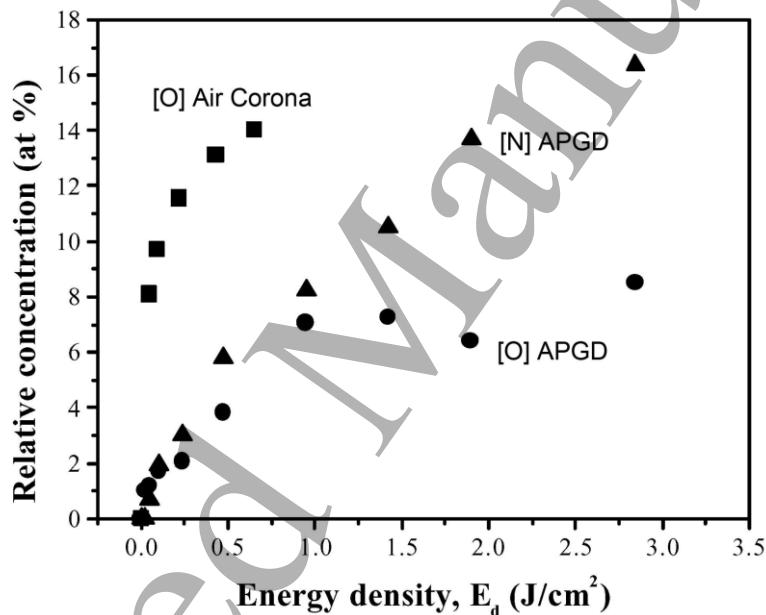


Figure 11: Surface concentrations of oxygen [O] and nitrogen [N] as a function of J/cm^2 of plasma energy incident onto the surface after treating biaxially-oriented polypropylene (BOPP) using an air corona and an atmospheric pressure glow discharge (APGD) sustained in nitrogen. The repetition rate was 1 kHz. Adapted from Ref. [205].

The concept of a fluence-based dose is already implemented in plasma surface treatment of polymers. For this application, treatment is characterized by the J/cm^2 delivered to the surface [203, 204] although the range of operating conditions for which this dose-driven description applies is limited. (See Figure 11.) Characterizing treatment of polymers by J/cm^2 is typically done for atmospheric pressure plasma treatment in a specified ambient – for example, air or He. We also note that in plasma-based chemical conversion and the emerging field of

1
2
3 plasma-catalysis, there is also a concept of energy delivery, which is termed the specific energy
4 input (SEI) and is defined as the amount of energy deposited by the plasma per volume or
5 molecule of gas, typically expressed in J/cm³, kJ/L, or eV/molecule [206]. This measure of
6 plasma activity is distinct from the concept of dose, as it reflects the excitation of the plasma
7 molecules themselves as opposed to the fluence of excitation delivered by a plasma to a surface.
8
9

10
11 The recent discussion of PD originated in the context of atmospheric pressure plasmas
12 for treating biological systems. At atmospheric pressure, advection typically dominates over
13 diffusion, and so the directionality of the flux constituting the dose is more easily defined at
14 atmospheric pressure (again, in analogy to radiation chemistry) as opposed to low pressure
15 where diffusion dominates. However, in principle, PD is (or should be) independent of pres-
16 sure, flow-rate or other ancillary conditions – *fluence is fluence*. In spite of the universality of
17 fluence, it will be necessary for plasma device manufacturers to be very explicit in how they
18 define their device's capabilities in terms of delivering PDu. Such a specification for a plasma
19 jet might be – the plasma device will deliver 10 PDu of O₂(¹Δ) every 20 seconds through a 2
20 mm × 2 mm window centered on axis and located 1 cm from the orifice of the device when
21 using the standard operating conditions. Similar specificity would be required by researchers
22 in reporting exposure of a surface by a given PDu of O₂(¹Δ).
23
24

25 Such descriptions of PD by manufacturers or researchers say nothing about what else
26 might be delivered along with the specified PDu(O₂(¹Δ)). Doses of ions, photons or RNS may
27 be delivered along with the PDu(O₂(¹Δ)). These other doses may affect the final outcome of
28 treatment, but do not affect the fact that the plasma device delivers a particular PDu(O₂(¹Δ)).
29 A specified PDu of O₂(¹Δ) at 1 cm from an orifice also says nothing about the PDu(O₂(¹Δ)) at
30 0.9 cm or 1.1 cm from the orifice.
31
32

33 It is true that a PDu(X) likely varies within a given system, which then requires care in
34 where the PDu(X) applies. For example, a flowing low pressure plasma delivering 10 PDu(O)
35 at 1 cm above a surface likely also delivers nearly the same 10 PDu(O) at 0.9 cm and 1.1 cm
36 above the surface. The rate of gas phase reactions of O atoms is low at low pressure and so the
37 flux of O atoms would likely be nearly the same at all of these locations. At atmospheric pres-
38 sure, the gradient in O atoms is much steeper due to the increased rate of reaction of O atoms.
39 As a result, PDu(O) at 1 cm would likely be different than at 0.9 cm and 1.1 cm, and could be
40 very different.
41
42

43 This working definition of PDu implies some aspect of linearity of the outcome of the
44 delivered dose, which is typically the case in radiation chemistry and polymer treatment. The
45
46

1
2
3 same PDu can be achieved with different combinations of convective speed, density and expo-
4
5 sure time. A dose of 10 PDu($O_2(^1\Delta)$) over 10 seconds with a convective speed of 10^3 cm/s
6 corresponds to a density of $O_2(^1\Delta)$ of 10^{12} cm $^{-3}$. 10 PDu($O_2(^1\Delta)$) delivered over 1 second with
7 a convective speed of 10^2 cm/s corresponds to a density of $O_2(^1\Delta)$ that is 100 times larger. If
8 there are non-linear aspects of the plasma chemistry or surface treatment, the same 10
9 PDu($O_2(^1\Delta)$) delivered for these two conditions might result in different outcomes. These po-
10
11 tentially different outcomes emphasize that to have any utility, PD should not be tied to out-
12
13 comes. PD would best be conceived as a measure of delivery of a given radical, ion or photon.

14
15 Condition dependent outcomes while keeping a scaling parameter constant are not
16 unique to PDu. Consider one of the most fundamental scaling parameters in LTP science –
17 E/N, electric field/gas number density, measured in terms of Townsends ($1\text{ Td} = 10^{-17}\text{ Vcm}^2$).
18 Operating two devices with the same E/N only implies that the rate coefficients for electron
19 impact processes of the ground state will be the same (assuming conditions akin to the local
20 field approximation). However, the outcome of those identical rate coefficients could be very
21 different for systems operating with the same E/N. Consider operating an oxygen plasma at
22 10^4 V/cm at 1000 Torr compared to 1 V/cm at 0.1 Torr. The outcome of exciting, dissociating
23 an ionizing oxygen at the same E/N will be very different in these two systems. Even for the
24 same pressure and same E/N, the outcome would be different for different current densities.

25
26 At the time that the concept of E/N was first proposed, there was also an understanding
27 of the limits of describing a system using E/N. Those limits include that excitation must be
28 dominantly by electron impact from the ground state, partial ionization should be small to re-
29 duce the influence of electron-electron collisions, gradients in density should be small com-
30 pared to the distance over which the measurements are made and the time for electrons to come
31 into equilibration with the applied electric field should be short compared to the time of interest.
32 Today, experiments often violate the basic assumptions that enable comparing systems operat-
33 ing at the same E/N. For example, the purpose of ns-pulsed plasmas is to overcome these very
34 limitations. Nevertheless, E/N remains a hugely valuable method of comparing plasma sys-
35 tems, even if the precision of those comparisons is not high.

36
37 A manufacturer specifying that a device delivers a given PDu(O) or a researcher report-
38
39 ing that treatment of a surface was performed with a given PDu(O) requires that the measure-
40 ments of O densities and convective speeds be properly performed. Having said that, PD should
41 be independent of the manner of measuring PD. Verifying that the measurements have been
42 done properly is discussed in Section III. It should also be recognized that measuring dose is

1
2
3 by no means easy and for some species, there is no consensus on the best approach to do so.
4 Benedikt et al., for example, recently outlined some of the best approaches for measuring flux,
5 but also some of the continuing challenges [207]. The entire plasma community would greatly
6 benefit from advances in measurement techniques for quantifying dose of different particles
7 and convergence toward accepted standards.
8
9

10 The concept of plasma dose is very appealing and, in some ways, is necessary to prom-
11 ulgate the use of plasmas to non-plasma experts. Such examples come from the biomedical
12 radiation chemistry community. Patients are treated by electron-, proton- and x-ray beams
13 based on delivering a dose (e.g., J/cm^3) while the practitioners are not necessarily experts in the
14 technology delivering those doses. For this reason alone, standards for PD (and PDu to measure
15 that dose) should be developed that are independent of the final application. However, the
16 expectation that PD can be used to precisely compare LTP systems may be too great a task for
17 today. It may be better to think of PD in the same manner as E/N – a concept that strictly
18 applies to idealized systems and which becomes less precise as those systems diverge from
19 those idealized conditions. PD would nevertheless retain high value in spite of the lack of
20 precision, just like E/N, in providing guidance, design principles and enabling initial compari-
21 son between systems.
22
23

24 **X. Technology Transfer**
25
26

27 This section addresses the transfer of research and development results to practical ap-
28 plications. The transfer of basic research to technological applications has long played a major
29 role in LTP physics, with an early example being technology invented by Siemens in 1857 for
30 ozone generation by means of silent discharges. Nevertheless, technology transfer in this re-
31 search area still faces barriers that often prevent the direct use of research results for the estab-
32 lishment of a new technology. These barriers are different in the various applications of LTPs
33 in industry and the healthcare field. These challenges, common practices and requirements in
34 transferring (or translating) experimental techniques, devices and codes to specific applications
35 are discussed here. Methods are suggested for how technology transfer can be supported and
36 efficiently organized in the future. (See Figure 12.)
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

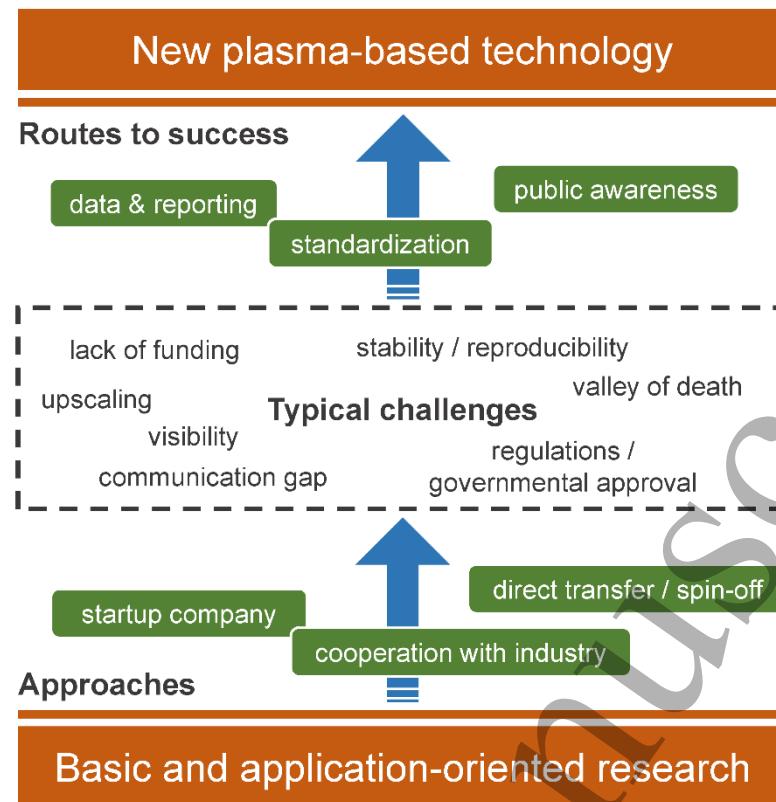


Figure 12: Common approaches, typical challenges and identified routes to success for technology transfer in the field of plasma science and technology.

Basic research takes place primarily at universities and national laboratories, which face particular challenges in the transfer of fundamental knowledge. University technology transfer in LTPs face problems similar to other fields of knowledge, with some accrued difficulties in identifying the relevant marketplace for applications that are often at the intersection of several disciplines. These issues strongly depend on the specific technology, and the perspectives and outcomes can vary by country or region (US, Europe, Asia) due to cultural, organizational and financial reasons. In spite of these differences, university technology transfer typically follows common standards and comparable challenges.

- (i) Modifying the desired outcomes of the research plan in terms of moving from fundamental-oriented to application-oriented, transforming the research results into a “product” with commercial added-value, and perhaps redefining scientific output with patents replacing papers in scientific journals
- (ii) Funding expectations, in terms of the initial investment needed to launch a startup company, and/or to develop a prototype that can be commercialized with a higher Technology Readiness Level (TRL) that satisfies possible regulatory requirements
- (iii) Maintaining the patent or the portfolio of patents that protects the invention;

1
2
3 (iv) Making decisions about the best route for the commercial exploitation of the invention,
4 concerning both the management of the intellectual property (opting for a controlled dis-
5 closure under a patented protection or using confidentiality agreements), and its exploitation
6 (under a startup company with the participation of the inventors, by licensing the research
7 product at certain royalty rates, or by selling the invention preferably keeping access to the
8 evolution of knowledge).

9
10 In general, technology transfer works particularly well when industrial partners are in-
11 volved in research and development from the outset. This is common practice in material and
12 surface processing industries, where plasma technology has been well established for decades
13 [208]. However, bridging the gap between laboratory scale solutions and industrial require-
14 ments remains challenging. Reproducibility of plasma processes and plasma uniformity on
15 large scales are among the issues in this regard and extensive consulting work occurs after
16 transferring technologies if devices do not work as they were originally expected. This chal-
17 lenge can be heightened by lack of full disclosure. Proprietary knowledge is often not fully
18 shared even if non-disclosure agreements are in place. In this respect, long-term partnerships
19 between research institutions and companies are advantageous before the start of a joint devel-
20 opment project as a trust relationship has likely already been established. Unlike for medical
21 and biomedical applications, standards and standard operating procedures do not play a partic-
22 ularly critical role in the field of plasma-based materials processing.

23
24 Challenges of a completely different kind occur in the healthcare field, where plasma
25 technology is not yet widely deployed, although research has been conducted in this area for
26 more than a decade [209, 210, 211]. The major issue is the governmental approval (e.g., the
27 Medical Device Regulation (MDR) in Europe), which is required to use plasmas for medical
28 applications. Particular challenges include patient safety with respect to high-voltage security
29 and electromagnetic compatibility, which follows the international standard IEC 60601-1
30 “Medical electrical equipment - Part 1: General requirements for basic safety and essential per-
31 formance”. Two-fold challenges arise from the fact that plasmas are complex and not (yet) well
32 known by the inspection and certification agencies. On the one hand, it is difficult to introduce
33 a secure and well-tested plasma device for a specific application when the response of the reg-
34 ular may be “what is plasma?”. On the other hand, it might happen that new plasma devices
35 are approved just because the first device also using plasma technology received approval with
36 the regulator assuming all plasmas are alike.

37
38 These difficulties can be overcome by means of testing standards. One attempt to unify
39 requirements for plasma medical devices has been made by introducing the DIN SPEC 91315

[212]. National and international standards are consensus-based and should be developed with the participation of all stakeholders. As the field of plasma medicine is still relatively young, the DIN SPEC was chosen as a first attempt, which is a special standardization defined as specification, and can be developed by a temporarily appointed committee advised by DIN e.V. (Berlin, Germany). The DIN SPEC “General requirements for medical plasma sources” describes basic criteria for the characterization of medical plasma devices, referring to international or national standards whenever appropriate, and creates a basis for physical and biological characterization of plasma and its effect on cells.

The DIN SPEC was first published in June 2014 with the intent to develop it further into a DIN / international standard with community consensus. Although that has yet to be accomplished, multiple studies have been published following the DIN SPEC protocols [213, 214, 215, 216, 217]. With a set of standard characterization ranging from gas temperature of the plasma to patient leakage current, inactivation of defined microorganisms and viability tests of eukaryotic cells to some simple chemical measurements in plasma treated liquid, the DIN SPEC can provide safety for users (investigators, patients, physicians), and accelerate technology transfer. In general, complete knowledge of all formal reporting requirements in the respective country is needed to ensure sufficient data management and proper documentation of procedures to meet the respective regulatory standards.

A promising way of technology transfer in both industrial and biomedical applications of LTPs is the establishment of spin-off companies. The main advantage of this approach is that the required scientific understanding and detailed knowledge of the intended application is directly available from the inventors themselves. Obstacles here are often the lack of start-up funding and, of course, the transition phase for any technological innovation. This does not exclude plasma technology. There is always the challenge to survive the “valley of death”, which is the transition from a laboratory prototype to a manufacturable and maintainable product.

Other issues encountered on the way to establishing new applications of plasma technology are visibility and economics. As a niche or merely supporting technology, it is often difficult to be perceived as an alternative to established technologies that have a large installed base. Resource consumption (e.g., energy, gas, electrode materials) and possible toxic by-products (e.g., toxic species, fluoride exhaust) are barriers to success. If the goal is replacement of an existing technology, collaboration with interdisciplinary partners with knowledge of current production technologies and product developments is beneficial. In all cases the specific regulatory, technological, infrastructural and operational conditions need to be considered. Life-

cycle assessment processes and cyber-physical systems like digital twins will become more and more relevant [218]. It has to be demonstrated that the plasma technology has a positive impact on the complete process chain of the production engineering. Moreover, the advantages in the product life cycle up to the recycling has to be addressed. In summary, existing obstacles to technology transfer can be minimized by

- Development of standardized tests and operational procedures for industrial and medical plasma applications.
- Establishment of data management and documentation standards that consider existing regulations and standard operating procedures.
- Expanding the development of standard and scalable plasma sources for various applications.
- Broad education about the benefits and risks of plasma technology in society.

Many of the topics discussed above (e.g., standard plasma sources, data and mechanisms, reporting, FAIR data principles) thus play a key role, especially for efficient technology transfer.

XI. Concluding Remarks

The goals of transparency, reproducibility and collaboration in LTP research are both fundamental to the field and a challenge to implement. The broad intellectual diversity of LTPs and their allied fields, the large variety of applications, and membership from the physical, biological and engineering sciences make rigidly applying a stringent set of standards difficult and probably not advisable. That said, working towards these goals while also embracing innovation should be a guiding principle for the field. This review has discussed several methods to achieve these goals, through proposed standards and best practices, while also acknowledging the role of judgement so that innovation is encouraged and not impeded.

The use of standard plasma sources such as the GEC reference cell and the COST jet have improved our understanding of fundamental plasma processes. The use of standard plasma sources should be a tool towards improving fundamental understanding of plasma processes and speeding translation of scientific findings to applications, while not hindering innovation. Rather, they can act as “calibration sources” to contextualize results, for diagnostics and modeling, and to facilitate new source development and scaled up processes. More community-driven designs for reference plasma sources and plasma-surface treatment systems are encouraged as they will likely promote fundamental understanding, benefiting the whole LTP community.

In contrast to other disciplines, recommended protocols for standard measurement techniques are less developed in plasma science, and for LTPs in particular. This situation is largely due to the challenging and strongly diverse nature of LTPs, often requiring diagnostic techniques specifically adapted to the distinct plasma environment and particular plasma parameters of interest. Key issues to consider in a proposed framework for diagnostic standards to support transparency, reliability and transferability of experimental measurements are: identifying and classifying the diagnostic technique, the regime of validity of the technique, analysis technique(s) or requirements for a range of plasma environments, the range of equipment and potential calibration procedures, and input data required for analysis of the data including uncertainty assessment.

The demand for improved and more widely applied plasma models will continue to drive the demand for data on key atomic and molecular processes. While sophisticated methods, both experimental and theoretical, will continue to provide important results for key processes, limitations of these methods both in terms of expense and capability will mean that other, cheaper and more approximate methods will continue to be used. Community based best practices for applying those approximate methods will work towards better comparisons of models. Machine learning is likely to provide a fruitful means of filling the many gaps in data provision for plasma modelling, a process that will also benefit from some standards or best practices.

Developing research software should consider the options of open-source and adopting FAIR principles early in the process. A possible roadmap involves planning; development, review and verification; sharing; and publishing. These practices could significantly contribute to elevating research standards, improving the quality of codes and computational predictions, reinforcing the conduit between codes and data (or databases), and promoting the transparent reporting on models and procedures.

Establishing the correctness of both computer codes and the mathematical models that they express has been a matter of increasing concern to many technical communities in recent years. Many LTP scientists share these concerns. However, the problems that arise are not simple. For many computational methods in wide use in the LTP community, canonical methods of demonstrating correctness (“verification”) do not yet exist. Even when appropriate techniques are available, a comprehensive verification effort can consume significant resources that could be otherwise used. Model testing by comparison with experiments (“validation”) may be even more time consuming and expensive. Consequently, authors, editorial boards, and other actors need to carefully assess what level of commitment to V&V activities is appropriate for

any particular calculation. The answer may vary widely, depending on the nature of the calculation and the role of the calculation in the chain of scientific argument. For instance, in exploratory scientific research, a rather light V&V effort may be appropriate. In engineering prediction, the full machinery of formal V&V may be needed. Future progress likely requires both technical improvements and better understanding by the community of the broader issues that arise.

As scrutiny of scientific findings continues to grow among the general public as well as public institutions and governments, proper reporting of research findings in archival journal articles takes on an added urgency. The LTP community is fortunate that there has not been a high-profile retraction or reproducibility scandal in the recent past that diminishes public trust in the field and its findings. While journals and journal publishers have a role to play in establishing standard practices for data reporting, the LTP (and broader scientific) community at large, including authors, reviewers, editors, and those that use data, plays perhaps the most important role in setting expectations for proper data reporting.

The transfer from basic and applied research to new plasma-based technologies faces several challenges that differ for industrial plasma applications and the healthcare sector. While the former is primarily hampered by communication and upscaling issues, the latter is often constrained by the requirements of governmental approvals and device regulations. Wider adoption of standards and best practices for the development and application of plasma sources, diagnostic methods, plasma models and simulation codes, data storage, and reporting on all of these aspects would be of great benefit to technology transfer in both fields.

Perhaps the most challenging of standards is correlating plasma source operation with a plasma dose – a specified fluence of an excited state, ion or radical delivered by the plasma source. Ideally, the doses of reactivity delivered by plasma sources should be part of their specifications. We now generally lack this correlation and lack the ability to describe plasma sources in terms of the dose of individual reactive species delivered by the sources. As a result, applications are strongly coupled to a particular and sometimes unique plasma source. Overcoming this limitation, then opens the possibility of having interchangeable plasma sources, or using different plasmas sources to treat small surface areas and large surface areas. Defining doses and processes requirements using dose might stimulate commercial development of a wide range of plasmas sources using the dose standard.

Acknowledgements

The authors acknowledge support from the following agencies for their efforts on this paper, and acknowledge the following individuals for their contributions:

L. L. Alves: FCT-Fundaçao para a Ciéncia e a Tecnologia (UIDB/50010/2020 and UIDP/50010/2020).

M. Becker: Federal Ministry of Education and Research (BMBF) under the grant mark 16QK03A; and S. Hamaguchi, T. Gerling, and M. Thomas for helpful input on the technology transfer section.

J. van Dijk: The Dutch Science Foundation NWO and the Netherlands eScience Center (project “passing XSAMS”).

T. Gans: UK Engineering and Physical Sciences Research Council EP/K018388/1.

D. B. Go: US National Science Foundation (PHY-1804091 and PHY-2206420), US Department of Energy (DE-SC0021083, DE-EE0009103, DE-FE0031862), US Air Force Office of Scientific Research (FA9550-18-1-0157), and US Army Research Office (W911NF-17-1-0119).

K. Stapelmann: US National Science Foundation (PHY-2107901), US Department of Energy (DE-SC0021329, DE-NE0009167), US National Institute of Health (R01EB029705), and the US Department of Agriculture (2020-67017-31260).

M. M. Turner: Science Foundation Ireland (16/IA/4628,16/SP/3809).

M. J. Kushner: US Department of Energy (DE-SC0020232), US National Science Foundation (1902878,1903151) and the US Army Research Office Multidisciplinary University Research Initiative (W911NF-20-1-0105 and W911NF-18-1-0240).

Conflict of Interest

The authors have no conflicts to disclose.

Data Availability

The data used in preparation of this article appear in the article or in the cited references. Requests for additional data should be directed to the corresponding author.

1
2
3
4
5 **References**
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1. <https://www.merriam-webster.com/dictionary/standards>
2. https://en.wikipedia.org/wiki/Technical_standard
3. P. J. Hargis et al. *Rev. Sci. Instrum.* **65**, 140 (1994).
4. S. Rauf, *Plasma Sources Sci. Technol.* **29**, 095019 (2020).
5. K. Siari, S. Rebiai, H. Bahouh and F. Bouanaka, *Phys. Rep.* **46**, 667 (2020).
6. Abhishek Kumar Verma and Ayyaswamy Venkatraman, *Comp. Phys. Comm.* **263**, 107855 (2021).
7. J. Golda et al. *J. Phys. D Appl. Phys.* **49** 084003 (2016).
8. F. Riedel, *Plasma Sources Sci. Technol.* **29**, 095018 (2020).
9. T. Murakami, K. Niemi, T. Gans, D. O'Connell and W. G. Graham, *Plasma Sources Sci. Technol.* **23**, 025005 (2014).
10. S. Schröter et al., *Phys. Chem. Chem. Phys.* **20**, 24263–86 (2018).
11. A. Brisset et al., *Journal of Physics D: Applied Physics* **54**, 285201(2021).
12. Y. Gorbanev, J. Golda, V. Schulz-van der Gathen and A. Bogaerts, *Plasma* **2**, 316 (2019).
13. D. Shaw, A. Wes, J. Bredin and E. Wagenaars, *Plasma Sources Sci. Technol.* **25**, 065018 (2016).
14. A. Privat-Maldonado, et al., *Cancers* **10**, 394 (2018).
15. A. Yayci et al., *Chem. Cat. Chem.* **12**, 5893 (2020).
16. J-W. Lackmann et al., *Sci. Rep.* **8**, 7736 (2018). <https://doi.org/10.1038/s41598-018-25937-0>
17. M. Laroussi, et al., *IEEE Transactions on Radiation and Plasma Medical Sciences* **6**, 127 (2020).
18. J-W. Lackmann et al., *PLoS One* **14**, e0216606 (2019)
19. K. Wende et al., *RSC Adv.* **10**, 11598 (2020).
20. P. Ranieri et al., *Applied Sciences* **10**, 2025 (2020).
21. G. Willems, J. Benedikt and A. von Keudell , *J. Phys. D Appl. Phys.* **50** 335204 (2017).
22. R. Engeln, B. Klarenaar and O. Guaitella, *Plasma Sources Sci. Technol.* **29**, 063001 (2020).
23. P. Bruggeman, D. C. Schram, M. G. Kong and C. Leys, *Plasma Process. Polym.* **6**, 751 (2009).
24. A. Dogariu, B. M. Goldberg, S. O'Byrne and R. B. Miles, *Phys. Rev. Appl.* **7**, 024024 (2017).
25. B. M. Goldberg, T. Hoder and R. Brandenburg, *Plasma Source Sci. Technol.* **31**, 073001 (2022).
26. Nonthermal Plasmas for Materials processing, Plasma Diagnostics (Chapter 3), J F Frie-
drich & J Meichsner (2022), <https://doi.org/10.1002/9781119364757>
27. N. St. J. Braithwaite, *Plasma Sources Sci. Technol.* **6** 133 (1997).
<https://doi.org/10.1088/0963-0252/6/2/006>
28. V. A. Godyak and V. I. Demidov, *J. Phys. D: Appl. Phys.* **44** 233001 (2011).
<https://doi.org/10.1088/0022-3727/44/23/233001>
29. J. Benedikt et al, *J. Phys. D: Appl. Phys.* **45**, 403001 (2012). <https://doi.org/10.1088/0022-3727/45/40/403001>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

30. U Fantz, *Plasma Sources Sci. Technol.* **15**, S137 (2006). <https://doi.org/10.1088/0963-0252/15/4/S01>
31. M. Gherardi, et al., *Plasma Sources Sci. Technol.* **24**, 064004 (2015).
32. A. Okigawa, et al., *Japanese Journal of Applied Physics* **36**, 4605 (1997).
33. *Laser Spectroscopy - Basic Principles*, Wolfgang Demtröder (2008), <https://doi.org/10.1007/978-3-540-73418-5>
34. H F Döbele et al., *Plasma Sources Sci. Technol.* **14**, S31 (2005). <https://doi.org/10.1088/0963-0252/14/2/S05>
35. J. E. M. Goldsmith and J. E. Lawler, *Contemporary Physics* **22**, 235 (1981). <https://doi.org/10.1080/00107518108231524>
36. J. Benedikt et al., *Plasma Sources Sci. Technol.* **25**, 045013 (2016). <https://doi.org/10.1088/0963-0252/25/4/045013>
37. S. Schröter et al., *Phys. Chem. Chem. Phys.*, **20**, 24263 (2018). <https://doi.org/10.1039/C8CP02473A>
38. S. Schröter et al., *Plasma Sources Sci. Technol.* **29**, 105001 (2020). <https://doi.org/10.1088/1361-6595/abab55>
39. K Niemi et al., *Appl. Phys. Lett.* **103**, 034102 (2013). <https://doi.org/10.1063/1.4813817>
40. K Niemi et al., *Plasma Sources Sci. Technol.* **14**, 375 (2005). <https://doi.org/10.1088/0963-0252/14/2/021>
41. G. Stancu, *Plasma Sources Sci. Technol.* **29**, 054001 (2020). <https://doi.org/10.1088/1361-6595/ab85d0>
42. C. Drag et al., *Plasma Sources Sci. Technol.* **30**, 075026 (2021). <https://doi.org/10.1088/1361-6595/abfbcb>
43. A. Hurlbatt, et al. (2017), *Plasma Process Polym.*, **14**: 1600138 , <https://doi.org/10.1002/ppap.201600138>
44. A. M. Lietz and M. J. Kushner, *J. Phys. D: Appl. Phys.* **49**, 425204 (2016). <https://doi.org/10.1088/0022-3727/49/42/425204>
45. S. Franke et al., *Sci Data* **7**, 439 (2020). <https://doi.org/10.1038/s41597-020-00771-0>
46. L.C. Pitchford et al. *Plasmas. Plasma Process. Polym.* **14**, 1600098 (2016) <https://doi.org/10.1002/ppap.201600098>
47. E. Carbone et al. *Atoms* **9**, 16 (2021). <https://doi.org/10.3390/atoms9010016>
48. <http://phys4entrydb.istp.cnr.it/Phys4EntryDB/>
49. M. L. Dubernet, et al., *J. Quant. Spectrosc. Radiat. Transf.* **111**, 2151 (2010). <https://doi.org/10.1016/j.jqsrt.2010.05.004>
50. Damien Albert et al. *Atoms* **8**, 76 (2020). <https://doi.org/10.3390/atoms8040076>
51. J. Tennyson, et al. *Plasma Sources Sci. Technol.* **26**, 055014 (2017). <https://doi.org/10.1088/1361-6595/aa6669>
52. <http://www.ppl.eng.osaka-u.ac.jp/ICDDPS3/>
53. L C Pitchford, *J. Phys. D: Appl. Phys.* **46**, 330301 (2013). <https://doi.org/10.1088/0022-3727/46/33/330301>
54. P. Koelman et al. *Plasma Process. Polym.* **14**, 1600155 (2016) <https://doi.org/10.1002/ppap.201600155>
55. <https://doi.org/10.5281/zenodo.6699264>
<https://doi.org/10.5281/zenodo.6400857>
56. S Tadayon Mousavi et al 2021 *Plasma Sources Sci. Technol.* **30**, 075007 (2021).

1
2
3 57. <https://www-amdis.iaea.org/xsams/index.html>
4 58. <https://www.inptdat.de/>
5 59. M. D. Wilkinson, et al. Scientific Data **3**, 160018 (2016).
6 <https://doi.org/10.1038/sdata.2016.18>
7
8 60. <https://math.nist.gov/MatrixMarket/>
9
10 61. For an overview, see <https://www.cfd-online.com/Links/refs.html#validation>
11
12 62. M. M. Turner et al. Physics of Plasmas **20**, 013507 (2013).
13 <https://doi.org/10.1063/1.4775084>
14
15 63. <https://www.landmark-plasma.com>
16
17 64. B. Bagheri et al. Plasma Sources Sci. Technol. **27**, 095002 (2018).,
18 <https://doi.org/10.1088/1361-6595/aad768>
19
20 65. A. Tejero-del-Caz et al., Plasma Sources Sci. Technol. **28**, 043001 (2019).
21
22 66. A. Tejero-del-Caz et al., Plasma Sources Sci. Technol. **30**, 065008 (2021);
23
24 67. G. J. M. Hagelaar and L. C. Pitchford, Plasma Sources Sci. Technol. **14**, 722 (2005).
25 <http://www.bolsig.laplace.univ-tlse.fr/>
26
27 68. R. Piron et al., High Energy Density Physics **23**, 38 (2017).
28 <https://doi.org/10.1016/j.hedp.2017.02.009>
29
30 69. G. P. Smith et al., <http://combustion.berkeley.edu/gri-mech/version30/text30.html>
31
32 70. <https://www.w3.org/2001/XMLSchema>
33
34 71. <https://json-schema.org/>
35
36 72. M.-L. Dubernet et al., Astronomy and Astrophysics **553**, A50 (2013).
37 <http://doi.org/10.1051/0004-6361/201220630>
38
39 73. <https://www-amdis.iaea.org/xsams/schema/v0.1.1/statesAtoms.xsd>.
40
41 74. <http://www.vamdc.org/activities/research/software/specview/>
42
43 75. <http://cassis.irap.omp.eu/>
44
45 76. <http://www.vamdc.org/activities/research/software/>
46
47 77. X. Wang, M. Wang, P. Biolsi and M. J. Kushner, J. Vac. Sci. Technol. A **39**, 033003
48 (2021).
49
50 78. K. Bartschat and M.J. Kushner Proc. Nat. Acad. Sci., **113**, 7026 (2016).
51
52 79. E.A. Belli, J. Candy and R.E. Waltz, Phys. Rev. Letts., **125**, 015001 (2020).
53
54 80. G. Colonna, et al., Plasma Sources Sci. Technol., **24**, 035004 (2015).
55
56 81. O.H. Crawford, A. Dalgarno and P.B. Hays Mol. Phys., **13**, 181 (1967).
57
58 82. S.-I. Chu and A Dalgarno Phys. Rev. A, **10**, 788 (1974).
59
60 83. Y. Itikawa, J. Phys. Soc. Jpn. **36**, 1121 (1974).
61
62 84. J.J. Munro, S. Harrison, J. Tennyson and M.M. Fujimoto J. Phys. Conf. Series, **388**,
63 012013 (2012)
64
65 85. Y.-K. Kim, Phys. Rev. A, **64**, 032713 (2001).
66
67 86. Y.K. Kim and M.E. Rudd, Phys. Rev. A, **50**, 3945 (1994).
68
69 87. V. Graves, B. Cooper and J. Tennyson, J. Chem. Phys. **154**, 114104 (2021).
70
71 88. K. Goswami, A.K. Arora, A. Bharadvaja and K.L. Baluja, Euro. Phys. J. D, **75**, 228
72 (2021).
73
74 89. H. W. Drawin, Z. Physik **225**, 470 (1969). <https://doi.org/10.1007/BF01392774>
75
76 90. H. W. Drawin, Z. Physik **225**, 483 (1969) <https://doi.org/10.1007/BF01392775>

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

91. W. H. Miller, *J Phys. Chem.* **102**, 793 (1998).
92. L. Hatton, *IEEE Computational Science and Engineering* **4**, 27 (1997).
93. W. L. Oberkampf and C. J. Roy, "Verification and validation in scientific computing," Cambridge University Press (2010).
94. T. Dauxois, *Physics Today* **61**, 55 (2008).
95. M. M. Turner, *Plasma Sources Sci. Technol.* **25**, 054007 (2016).
96. M. J. Kushner, *J. Phys. D: Appl. Phys.* **42**, 194013 (2009).
97. M. M. Turner, *Plasma Processes and Polymers* **14**, 1600121 (2017).
98. M. Surendra, *Plasma Sources Sci. Technol.* **4**, 56 (1995).
99. T. Charoy et al, *Plasma Sources Sci. Technol.* **28**, 105010 (2019).
100. Z. Lj. Petrovic et al, *J. Phys. D: Appl. Phys.* **42**, 194002 (2009).
101. M. M. Turner, *Plasma Sources Sci. Technol.* **24**, 015003 (2016).
102. A. G. Smart, *Physics Today* <https://doi.org/10.1063/PT.6.1.20180822a> (2018)
103. R. Bast, *Nature Physics* **15**, 728 (2019).
104. A. M. St. Laurent, "[Understanding Open Source and Free Software Licensing](#)", O'Reilly Media. p. 4. [ISBN 9780596553951](https://www.oreilly.com/library/view/open-source-and-free-software/9780596553951/) (2008)
105. W. L. Morgan and B. M. Penetrante, *Comput. Phys. Commun.* **58**, 127 (1990).
106. N. A. Dyatko, I. V. Kochetov, A. P. Napartovich and A. G. Sukharev, EEDF: The Software Package for Calculations of the Electron Energy Distribution Function in Gas Mixtures (2015). <https://doi.org/10.13140/RG.2.1.3675.1528>; <https://lxcat.net/download/EEDF/>
107. <https://github.com/aluque/bolos>
108. W. P. Allis, *Handbuch der Physik* volume 21, edited by S Flügge (Berlin: Springer) pp 383–444 (1956).
109. H. A. Lorentz, The motion of electrons in metallic bodies I *Proc. Royal Netherlands Academy of Arts and Sciences (KNAW)* vol 7 (Amsterdam: KNAW) pp 438–53 (1905).
110. L. L. Alves, A. Bogaerts, V. Guerra and M. M. Turner, *Plasma Sources Sci. Technol.* **27**, 023002 (2018).
111. M. Rabie and C. M. Franck, *Comput. Phys. Commun.* **203**, 268 (2016).
112. S. F. Biagi, *Nucl. Instrum. Methods Phys. Res. A* **421**, 234–40 (1999).
113. J. Stephens, *J. Phys. D: Appl. Phys.* **51**, 125203 (2018).
114. J. Stephens, *Phys. Plasmas* **25**, 103502 (2018).
115. <https://lxcat.net>
116. D.S. Stafford and M.J. Kushner, *J. Appl. Phys.* **96**, 2451 (2004).
117. S. Panchesnyi, B. Eismann, G.J.M. Hagelaar and L.C. Pitchford, ZDPlasKin <http://www.zdplaskin.laplace.univ-tlse.fr> (University of Toulouse, LAPLACE, CNRS-UPS-INP, Toulouse, France (2008))
118. A.H. Markosyan, A. Luque, F. J. Gordillo-Vázquez and U. Ebert, *Comput. Phys. Commun.* **185**, 2697 (2014).
119. J.J. Munro and J. Tennyson, *J. Vac Sci. Technol. A* **26**, 865-9 (2008).
120. J. Tennyson et al., *J. Phys.: Conf. Ser.* **86**, 012001 (2007).
121. L. A. Morgan, J. Tennyson and C. J. Gillan, *Comput. Phys. Comm.* **114**, 120 (1998).
122. J. P. Boeuf and L. C. Pitchford, *Phys. Rev. E* **51**, 1376 (1995).

1
2
3 123. L. J. Overzet and M. B. Hopkins, *J. Appl. Phys.* **74**, 4323 (1993).
4
5 124. L. J. Overzet and M. B. Hopkins, *Appl. Phys. Lett.* **63**, 2484 (1993).
6
7 125. S. A. Norberg, E. Johnsen and M. J. Kushner, *Plasma Sources Sci. Technol.* **24**, 035026
8 (2015).
9
10 126. D. Gaston, C. Newman, G. Hansen and D. Lebrun-Grandie, *Nuclear Engineering and De-
sign* **239**, 1768-1778 (2009).
11
12 127. A. D. Lindsay, D. B. Graves and S. C. Shannon, *J. Phys. D: Appl. Phys.* **49**, 235204
13 (2016).
14
15 128. K. Shan, C. Davide, D. Corey and S. Shannon, *APS Annual Gaseous Electronics Confer-
ence 2019*, abstract id.DT3.002;
16 <https://ui.adsabs.harvard.edu/abs/2019APS..GECDT3002K/abstract>
17
18 129. S. Keniley, C. Icenhour, A. Lindsay, D. Curreli, D. Panici, L. Harbour and D. Schwen
19 (2022); <https://doi.org/10.5281/zenodo.6780208>
20
21 130. <https://github.com/lindsayad/zapdos>
22
23 131. C. Icenhour, A. Lindsay, D. Green, R. Martineau and S. C. Shannon, *APS Annual Gase-
ous Electronics Conference 2018*, abstract id.ET1-002; [http://meetings.aps.org/Meet-
ing/GEC18/Session/ET1.2](http://meetings.aps.org/Meet-
ing/GEC18/Session/ET1.2)
24
25 132. H. Jasak, A. Jemcov, Z. Tukovic, OpenFOAM: A C++ library for complex physics sim-
ulations. In *International workshop on coupled methods in numerical dynamics* (Vol. 1000, pp. 1-20). IUC Dubrovnik Croatia, (2007). <https://www.openfoam.com/>
26
27
28 133. C. Fiorina, I. Clifford, S. Kelm, and S. Lorenzi, *Nuclear Engineering Design*, **387**, 111604
29 (2022).
30
31 134. C. White, et al., *Comput. Phys. Comm.* **224**, 22 (2018).
32
33 135. Z. Merali, *Nature* **467**, 775 (2010).
34
35 136. B. Boehm, "Foundations of Empirical Software Engineering: The Legacy of Victor R.
Basili" (H. D. Rombach and M. V. Zelkowitz, eds.), Springer (2005)
36
37 137. L. Hatton and A. Roberts, *IEEE Trans. Softw. Eng.* **20**, 785 (1994).
38
39 138. <http://www.basilisk.fr/>
40
41 139. "LXCat discussion" at the Gaseous Electronics Conference, Bochum, Germany (2016);
42 <http://www.gec2016.de/lxcat/index.html.en>
43
44 140. I Adamovich et al, *J. Phys. D: Appl. Phys.* **50**, 323001 (2017).
45
46 141. G. Wilson, et al., *PLoS Comput. Biol.* **13**, e1005510 (2017).
47
48 142. R. C. Jiménez, M. Kuzak, M. Alhamdoosh et al, *F1000Research* **6**, 876 (2017).
49
50 143. <https://git-scm.com/>
51
52 144. <https://www.mercurial-scm.org/>
53
54 145. <https://github.com/>
55
56 146. <https://gitlab.com/>
57
58 147. N. Chue Hong, "In which journals should I publish my software?";
59 <https://go.nature.com/2JfeoCA>
60
61
62 148. K. V. Roberts, *Comp. Phys. Commun.* **1**, 1 (1969).
63
64 149. F. Peeters, and T. Butterworth, "Electrical Diagnostics of Dielectric Barrier Discharges",
65 in *Atmospheric Pressure Plasma - from Diagnostics to Applications*. London, United
66 Kingdom: IntechOpen, 2018 [Online]. Available: [https://www.intechopen.com/chap-
ters/63194](https://www.intechopen.com/chap-
ters/63194) doi: 10.5772/intechopen.80433

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

150. X.M. Zhu and Y.K. Pu, *J. Phys. D* **43**, 403001 (2010).
151. T. Belmonte et al., *Plasma Source Sci. Technol.* **24**, 064003 (2015).
152. A.R. Hanna and E.R. Fisher, *J. Vac. Sci. Technol. A* **38**, 020806 (2020).
153. R. Brandenburg, *Plasma Source Sci. Technol.* **26**, 053001 (2017).
154. F.A. Herrera, et al., *J. Phys. D* **52**, 224002 (2019).
155. A.H. Markosyan, J. Teunissen, S. Dujko, and U. Ebert, *Plasma Source Sci. Technol.* **24**, 065002 (2015).
156. A. Bogaerts, et al., *Plasma Source Sci. Technol.* **26**, 063001 (2017)
157. P.F. Dunn and M.P. Davis, *Measurement and Data Analysis for Engineering and Science*. CRC Press (2017).
158. P.J. Roache, *Verification and Validation in Computational Science and Engineering*. Hermosa (1998).
159. R.G. McClaren, *Uncertainty Quantification and Predictive Computational Science*. Springer International Publishing (2018).
160. A. Fierro, et al., *European Physical Journal D*, **75**, 1 (2021)
161. A. Berthelot and A. Bogaerts, *Plasma Source Sci. Technol.*, **26**, 115002 (2017).
162. C. D. Chambers, *Cortex*, **66**, A1 (2015).
163. J. M. Wicherts, M. Bakker and D. Molenaar, *PloSOne*, **6** e26828 (2011).
164. Open Science Collaboration, 2015. Estimating the reproducibility of psychological science. *Science*, 349, aac4716 (2015), <https://doi.org/10.1126/science.aac4716>
165. X. Chen et al., *Nature Physics* **15**, 113 (2019).
166. F. Schüth, M. D. Ward and J. M. Buriak, *Chem. Mater.* **30**, 3599 (2018). <https://doi.org/10.1021/acs.chemmater.8b01831>
167. J. M. Buriak, P. V. Kamat, and K. S. Schanze, *ACS Appl. Mater. Interfaces* **6**, 11815 (2014). <https://doi.org/10.1021/am504389z>
168. U. Riebesell, V. J. Fabry, L. Hansson, and J. P. Gattuso, “Guide to best practices for ocean acidification research and data reporting”, Publications Office of the European Union, 2011. <https://data.europa.eu/doi/10.2777/66906>
169. H. A. Le Phuong, C. F. Blanford, and G. Szekely, *Green Chemistry*, **22**, 3397 (2020).
170. P. Vischia, *Reviews in Physics*, **5**, 100046 (2020). <https://doi.org/10.1016/j.revip.2020.100046>
171. J. E. Russell, *Nature* **496**, 7 (2013). <https://doi.org/10.1038/496007a>
172. G. Boulton, *Nature*, **530**, 281 (2016). <https://doi.org/10.1038/530281c>
173. M. Baker, *Nature*, **533**, 452 (2016). <https://doi.org/10.1038/533452a>
174. G. B. Drummond and B. D. Tom, *Advances in Physiology Education*, **35**, 365 (2011)
175. G. P. Smestad et al., *Solar Energy Matls. Solar Cells*, **92**, 371 (2008) <https://doi.org/10.1016/j.solmat.2008.01.003>
176. E. J. Luber, and J. M. Buriak, *ACSNano* **7**, 4708 (2013).
177. <https://journals.aps.org/prapublications/PhysRevA.83.040001>
178. <https://publishingsupport.iopscience.iop.org/iop-publishing-data-availability-policy/>
179. <https://www.sciencemag.org/authors/science-journals-editorial-policies>
180. <https://www.aiaa.org/publications/books/Publication-Policies/Editorial-Policy-State-ment-on-Numerical-and-Experimental-Accuracy>
181. J. Diong, A. A. Butler, S. c. Gandevia, and M. E. Héroux, *PloS one*, **13**, e0202121 (2018)..

1
2
3 182. A. Stevanovic et al., PLoS One **10**, e0128061 (2011).
4 183. M. Héroux, M., J. Neurophysiology, **116**, 1536 (2016).
5 184. <https://www.nature.com/documents/nr-editorial-policy-checklist.pdf>
6 185. <https://www.nature.com/nphys/for-authors/preparing-your-submission>
7 186. <https://www.nature.com/documents/nr-lasing-reporting.pdf>
8 187. S. O'Connor, et al., IEEE Transactions on Plasma Science **49**, 1724 (2021).
9 188. F. Riva, C. F. Beadle, and P. Ricci, P., 2017. A methodology for the rigorous verification
10 of particle-in-cell simulations. Physics of Plasmas, **24**, 055703 (2017).
11 189. D. P. Debeckeret al., ACS Sustainable Chem. Eng. **9**, 4936 (2021). [
12 https://doi.org/10.1021/acssuschemeng.1c02070]
13 190. E. L. Clark, et al., ACS Catalysis **8**, 6560 (2018).
14 191. P. Barboun et al., ACS Sustainable Chemistry & Engineering, **7**, 8621 (2019).
15 192. https://us.lxcat.net/instructions/how_reference.php
16 193. <https://publishingsupport.iopscience.iop.org/iop-publishing-data-availability-policy/>
17 194. [https://publishingsupport.iopscience.iop.org/wp-content/uploads/2018/05/Reviewer-
18 guide-V11.pdf](https://publishingsupport.iopscience.iop.org/wp-content/uploads/2018/05/Reviewer-guide-V11.pdf)
19 195. H. Cheng, J. X. Xu, X. Li, D. W. Liu and X. P. Lu, Physics of Plasmas **27**, 063514 (2020).
20 196. Th. Von Woedtke, S. Reuter, K. Masur and K.-D. Weltmann, Phys. Rep. **530**, 291 (2013).
21 197. H. M. Joh, S. J. Kim, T. H. Chung and S. H. Leem, Appl. Phys. Lett. **101**, 053803 (2012).
22 198. X. Dai, Z. Zhang, J. Zhang and K. Ostrikov, Plasma Proc. Polym. **17**, e1900178 (2020).
23 199. W-O. Ji, M-H Lee, G-H. Kim and E-H. Kim, Sci. Reports **9**, 19837 (2019).
24 200. D. R. Spitz et al., Radiotherapy Oncology **139**, 23 (2019).
25 201. G. Blain et al., Rad. Res. **198**, 218 (2022).
26 202. B. Ambrozic et al., Chem. Sci. **10**, 8735 (2019).
27 203. M. Strobel et al., Plasma Polymer **8**, 61 (2003).
28 204. T. Jacobs et al., Surf. Interface Anal. **42**, 1316 (2010).
29 205. S. Guimond et al., Plasma Polymer **7**, 71 (2002).
30 206. R. Snoeckx and A. Bogaerts, Chemical Society Reviews **46**, 5805-5863 (2017).
31 207. J. Benedikt, H. Kersten, A. Piel, Plasma Sources Science and Technology **30**, 033001
32 (2021)
33 208. D. B. Graves, IEEE Transactions on Plasma Science, **22**, 31 (1994).
34 https://doi.org/10.1109/27.281547.
35 209. G. Fridman, et al. Plasma Processes Polymers **5**, 503 (2008).
36 https://doi.org/10.1002/ppap.200700154
37 210. M. G. Kong, "Plasma medicine: an introductory review." New Journal of Physics **11**,
38 115012 (2009).
39 211. Th. von Woedtke, S. Reuter, K. Masur and K.-D. Weltmann, Phys. Reports. **530**, 291
40 (2013). <https://doi.org/10.1016/j.physrep.2013.05.005>
41 212. M. Mann et al., Clinical Plasma Medicine **4**, 35 (2016).
42 213. V. Hahn, R. Brandenburg, and T. von Woedtke, (2018). DIN SPEC 91315: A First At-
43 tempt to Implement Mandatory Test Protocols for the Characterization of Plasma Medical
44 Devices. In: H. R. Metelmann, T. von Woedtke, and K-D. Weltmann (eds) Comprehen-
45 sive Clinical Plasma Medicine. Springer, Cham. [https://doi.org/10.1007/978-3-319-
47 67627-2_35](https://doi.org/10.1007/978-3-319-
46 67627-2_35)

1
2
3 214. A. Lehmann, et al., *Clinical Plasma Medicine* **7**, 16 (2017).
4 215. E. Timmermann, et al. *Journal of Physics D: Applied Physics* **54**, 025201 (2020).
5 216. R. Tiede, S. Emmert, and G. Isbary, (2018). Treatment of Ulcerations and Wounds. In: H.
6 R. Metelmann, T. von Woedtke, and K-D. Weltmann (eds) *Comprehensive Clinical*
7 *Plasma Medicine*. Springer, Cham. https://doi.org/10.1007/978-3-319-67627-2_7
8 217. T. Bernhardt et al., *Oxidative Medicine and Cellular Longevity*, **2019**, 3873928 (2019).
9 <https://doi.org/10.1155/2019/3873928>
10 218. F. Tao et al., *Engineering* **5**, 653 (2019). <https://doi.org/10.1016/j.eng.2019.01.014>.
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60