arXiv:2303.03772v1 [physics.data-an] 7 Mar 2023

人为因素：小角散射数据分析循环比对结果

Brian R. Pauw[1], **Glen J. Smales**[1]

Andy S. Anker[2], **DanielM. Balazs**[3], **Frederick L. Beyer**[4], **Ralf Bienert**[5], **Wim G. Bouwman**[6], **IngoBreßler**[1], **Joachim Breternitz**[7], **Erik S Brok**[8], **Gary Bryant**[9], **Andrew J. Clulow**[10],**Erin R. Crater**[11], **Fr´ed´eric De Geuser**[12], **AlessandraDel Giudice**[13], **J´erˆome Deumer**[14], **SabrinaDisch**[15], **Shankar Dutt**[16],**Kilian Frank**[17], **Emiliano Fratini**[18], **Elliot P. Gilbert**[19], **Marc Benjamin Hahn**[1], **James Hallett**[20], **Max Hohenschutz**[21], **Martin Hollamby**[22], **Steven Huband**[23], **Jan Ilavsky**[24], **JohannaK. Jochum**[25],**Mikkel Juelsholt**[26],**Bradley W. Mansel**[27],**Paavo Penttil¨a**[28], **Rebecca K. Pittkowski**[2], **Giuseppe Portale**[29], **Lilo D. Pozzo**[30], **Paulo RicardodeAbreu FurtadoGarcia**[31], **Leonhard Rochels**[15], **Julian M. Rosalie**[1], **Patrick E. J. Saloga**[32], **Susanne Seibt**[10], **Andrew J. Smith**[33],**Gregory N. Smith**[34],**Venkatasamy Annadurai**[35], **GlennA. Spiering**[36], **Tomasz M. Stawski**[5], **Olivier Tach´e**[37],**Andreas F. Th¨unemann**[1],**Kristof Toth**[38],**Andrew E. Whitten**[19], **JoachimWuttke**[39]

[1] 德国联邦材料研究与测试研究所 12205 柏林，德国

电子邮箱：brian.pauw@bam.de[2] 哥本哈根大学化学系，丹麦哥本哈根 2100[3] 奥地利科学技术研究所（IST Austria），Am Campus 1，克洛斯特新堡，3400，奥地利 [4] 美国陆军研究实验室，6300 Rodman Road，美国马里兰州阿伯丁试验场 21005[5] 德国联邦材料研究与测试研究所，Richard-Willst¨atter-Str. 11，12489 柏林，德国 [6] 代尔夫特理工大学[7] 柏林亥姆霍兹材料与能源中心（Helmholtz-Zentrum Berlin f¨ur Materialien und Energie GmbH）[8] FORCE Technology，Park Alle 345，2605 布隆比，丹麦[9] 皇家墨尔本理工大学理学院物理学系，澳大利亚墨尔本[10] 澳大利亚核科学与技术组织—澳大利亚同步加速器，800 Blackburn Road，Clayton 3168，维多利亚州，澳大利亚[11] 弗吉尼亚理工大学，美国弗吉尼亚州布莱克斯堡 24060[12] 格勒诺布尔-阿尔卑斯大学、法国国家科学研究中心、格勒诺布尔国立综合理工学院、SIMAP，F-38000 格勒诺布尔，法国[13] 罗马智慧大学“乌戈·希夫”化学系，P.le A. Moro, 5，I-00185 罗马，意大利

[14] Physikalisch-Technische Bundesanstalt，Abbestr. 2，10587 柏林[15] 科隆大学化学系，Greinstraße 4-6，50939 科隆，德国¨澳大利亚国立大学材料物理系，澳大利亚首都领地，澳大利亚¨慕尼黑大学物理学院和CeNS，Geschwister-Scholl-Platz 1，80539 慕尼黑，德国¨佛罗伦萨大学“乌戈·希夫”化学系和CSGI，Via della Lastruccia, 3，50019 塞斯托·菲奥伦蒂诺，意大利¨澳大利亚核科学与技术组织澳大利亚中子散射中心，New Illawarra Road，卢卡斯高地，新南威尔士州 2234，澳大利亚[16] 雷丁大学化学系、化学、食品与药学院，Whiteknights，PO Box 224，雷丁 RG6 6AD，英国[17] 亚琛工业大学物理化学研究所，Landoltweg 2，D-52056 亚琛，德国¨基尔大学化学与物理科学学院，英国¨华威大学物理系，CV4 7AL，英国[18] 先进光子源，阿贡国家实验室，9700S Cass ave，美国伊利诺伊州勒蒙[19] 海因茨·迈尔-莱布尼茨中心，慕尼黑工业大学，D-85748 加兴，德国[20] 牛津大学材料系，Parks Road，牛津，OX1 3PH，英国[21] 国家同步辐射研究中心，101 Hsin-Ann Road，新竹科学园区，新竹 30076，台湾，中华民国[22] 阿尔托大学生物制品与生物系统系，P.O. Box 16300，FI-00076 阿尔托，芬兰[23] 格罗宁根大学（荷兰）[24] 华盛顿大学化学工程系，Box 351750，美国华盛顿州西雅图 98195-1750[25] 巴西同步辐射光实验室（LNLS）—Giuseppe M¨aximo Scolfaro 10000，邮编：13083-100，坎皮纳斯，圣保罗州，巴西¨独立研究人员[26] 钻石光源有限公司，哈韦尔科学与创新园区，迪德科特，牛津郡，OX11 0DE，英国[27] ISIS中子与μ子源，科学技术设施委员会，卢瑟福·阿普尔顿实验室，迪德科特，OX11 0QX，英国[28] 迈索尔大学，NIE一级学院，迈索尔 - 570008[29] 弗吉尼亚理工大学化学系、大分子创新研究所（MII），美国弗吉尼亚州布莱克斯堡 24061[30] 巴黎-萨克雷大学，CEA，法国国家科学研究中心，NIMBE，91191，吉夫叙尔伊韦特，法国[31] 美国国家标准与技术研究院材料科学与工程部，美国马里兰州盖瑟斯堡 20899 ´于利希研究中心有限公司，JCNS-MLZ，Lichtenbergstraße 1，85747 加兴，德国

摘要。

开展了一项循环比对研究，以评估人为因素对小角散射数据分析的影响。向参与者提供了四个校正后的数据集，可供直接分析。所有数据集均测自含有球形散射体的样品，其中两个来自稀分散体系，另外两个来自粉末。46 位参与者中的大多数正确识别出了稀分散体系中的总体数量

，其中总体均值结果有一半在 1.5 % 以内，总体宽度结果有一半在 40 % 以内。由于结构因子增加了分析的复杂性，提交粉末数据集答案的人数少得多。在提交答案的参与者中，均值和宽度结果各有一半分别在 44 % 和 86 % 以内。这项轮次试验揭示了造成差异的若干原因，并提出了相应的解决方案。

关键词：循环比对、小角散射、纳米结构定量已提交

收件人：*Modelling Simul. Mater. Sci. Eng.*

1. 引言

科学方法的发展历来旨在消除人类因素和仪器偏差对我们认识自然界的影响。它已应用于众多领域，成效各异。循环比对（RR）实验等工具可用于评估这种成效。例如，将相同的物品送至不同实验室进行检测，就能量化各实验室结果的差异。理想情况下，只要所用手段和方法达到最低标准，观测结果及由此得出的结论就不应受到观察者或仪器的影响。如果确实如此，我们便可以确信结果不存在偏差，并且很可能准确地反映了所研究的物体或现象。

纳米材料分析领域中，若干值得关注的比对试验比较了不同技术的结果（前提是不同技术能够提供真正可比的终端参数）。然而，为了更准确地反映真实的人为或仪器差异，专注于单一技术甚至某项技术的单一方面的试验或许更为合适。这样的研究重点有助于我们找出造成实验室间差异的主要因素，最终目标是消除或尽量减少这类依赖关系。此类研究的显著案例见于原子探针断层成像 [1]，X射线衍射 [2, 3]，中子散射 [4]，中子粉末衍射 [5]，生物小角X射线散射（生物专用小角X射线散射） [6]，以及采用Brunauer、Emmett和Teller法测定表面积的研究 [7]。

沿着这一思路，几年前开展了一项大规模比对试验，重点是采集纳米颗粒液体悬浮液的小角散射数据。小角散射是一种用于对大量样品中的纳米结构进行（可溯源）定量的方法。经过适当的数据校正后，可以获得有关散射体形态（“形状因子”）、其尺寸分布 ‡）或其堆积（“结构因子”）的信息。在大多数情况下，只要已知或假设另外两项信息，就能阐明这三项中的一项。在少数情况下，可以令人信服地从数据中提取其中两项或更多信息。

在上述轮次测试中，收集了来自众多实验室的小角X射线散射数据集，随后使用一套程序和一致的初始参数进行分析 [8]。这项实验清楚地表明，至少对于不具挑战性的样品，大多数实验室和仪器都能够采集到一致的数据；平均粒径的标准差约为百分之一，而总体尺寸分布宽度的标准差约为百分之十。接下来顺理成章的一步，就是探究人为因素会对数据分析产生何种影响。

影　　响研究人员对结果的影响，可以通过分发　a

‡ 虽然数据包含与散射体质量或体积成正比的信息，但对于分布较窄的情况，可以将其转换为数量加权分布，同时保持较低的不确定度。

待解读的数据集，并量化最终形态参数的变化幅度 [7, 2, 3]。尤其是在小角散射领域，数据分析可能成为一大障碍，因此预计本次小角散射数据分析轮测（DARR）的结果会有很大差异。

鉴于该领域中可能涉及的样品和分析种类繁多，挑战在于找到具有代表性的数据集，这些数据集应当：

（i）涵盖相关材料及数据集常见挑战的各种情形；（ii）质量较高，以尽量减少数据不确定性导致的结果差异；（iii）来自经过充分表征且深入了解的样品；（iv）附有与材料研究人员通常提供的名义信息量相同的补充信息

实验还受到答题空间方面实际限制的影响：需要设计一份机器可读的答题表，用于呈现所得形态参数以分析其差异，同时避免暗示预期结果或答案范围。答题表中的若干方面被有意含糊处理，旨在进一步探明小角散射从业者对一些常见但容易引起困惑的术语的理解（例如球形散射体的“平均尺寸”，并未说明指的是直径还是半径）。下文将进一步讨论答题表是否合适。

尽管始终受到上述考量的限制，但从46份提交结果中推断出的结果，仍然很好地揭示了小角散射作为材料科学工具所面临的挑战。在介绍数据集、方法和结果之后，文中通过简短讨论，结合结果分析，指出了该领域可能需要进一步关注的方向以及改进途径。此外，文章还用一段内容讨论了未来比对试验可能的改进方向。最后需要说明的是，尽管本文所呈现的结果必然有所局限，但匿名化结果已发布在 Zenodo [9]，上，同时还提供了用于校正、解读和可视化的 Jupyter 笔记本，希望能够据此发展出对结果的其他解读或扩展性解读。

2. 数据集描述

向愿意参与的参与者提供了四个一维散射数据集，分别代表两个稀溶和两个浓密的纳米颗粒体系。数据集以制表的三列 .dat 文件形式提供，其中包含 q、I(q) 和 $\sigma_I(q)$。其中，q 为以 nm^{-1} 为单位的散射矢量，I(q) 为以 $(m \cdot sr)^{-1}$ 为单位的散射强度，而 $\sigma_I(q)$ 表示强度的绝对不确定度（一个标准差）。此外，还向参与者提供了一封介绍数据集及后续任务的信函，以及一份用于以标准化格式记录结果的 Excel 表格（参见 Zenodo 存储库 [10]）。

四个数据集如图1所示，同时给出了模型拟合结果；这些拟合仅用于对数据集适用模型的建议。数据集1和2在

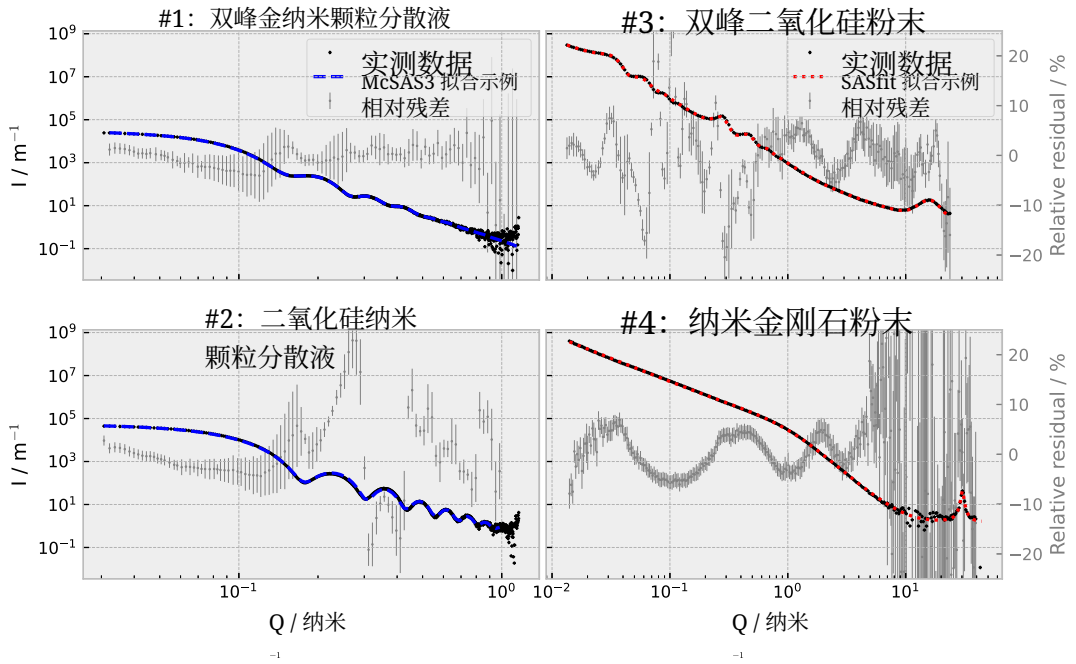

图 1. 本次比对试验选用的四组数据集（黑色），以及示例拟合和相对残差（浅灰色，对应右侧的次坐标轴）。示例拟合分别使用 McSAS3（左侧，蓝色）和 SASfit（右侧，红色）生成。这些拟合得到的尺寸分布如图 2 所示。

McSAS3中使用球形形状因子进行拟合，而数据集3和4则在SASfit中使用球形形状因子（采用对数正态分布）并结合适当的结构因子进行拟合（数据集3和4分别采用黏性硬球和质量分形），同时加入背景贡献和峰函数，以更好地描述广角数据。这些拟合结果也可在Zenodo 存储库中找到。 [9], 表1列出了部分拟合参数（每个数据集各总体的均值和宽度）。

数据集 1 和 2 来源于公开可用的测量数据 [11]，测量在 Physikalisch-Technische Bundesanstalt (PTB) 的 SAXS 光束线上进行，地点为 Berliner Elektronenspeicherring-Gesellschaft für Synchrotronstrahlung II mit beschränkter Haltung (BESSY II 有限责任公司) [12, 13], ；测量对象是为欧洲计量创新与研究计划（EMPIR）“nPSize”项目合成的参考样品。这些参考样品详见 [14]。

数据集 3 和 4 使用 MOUSE（超细结构探索方法优化） [15]以粉末形式进行测量。X射线由带有铜阳极的微焦点X射线管产生，随后通过多层光学元件将X射线束准直并单色化，使其波长约为 Cu Kα (λ =0.154 nm)。样品以少量粉末的形式装入薄型激光切割样品架中，并夹在两片低散射思高牌透明胶带之间固定。散射

辐射由真空中的 Eiger 1M 探测器（Dectris，瑞士）检测，探测器放置在距样品 55 至 2507 毫米的多个位置。随后使用 DAWN（数据分析工作台）软件包，按照标准化的完整二维校正流程对所得数据进行处理，并缩放至绝对强度，同时进行不确定度传播 [15, 16]。

2.1. 数据集 *1*：双峰金纳米颗粒

数据集 1 来自一种名为 nPSize 1 的材料。该样品设计为在水中含有两种总体已知浓度的球形金纳米颗粒（NP），其直径分别为（30 和 60）nm，数量比为 1:1。根据这一数量比，体积分数比约为 1:8。换言之，较小的总体（总体 1）所占体积分数为 1/8，因此对散射图样的贡献远小于较大的总体（总体 2）。需要指出的是，在实际测量中，通常并不知道总体的模态。因此，并未向参与者明确揭示第二个总体的存在，而只是通过答题表的设计略作暗示。

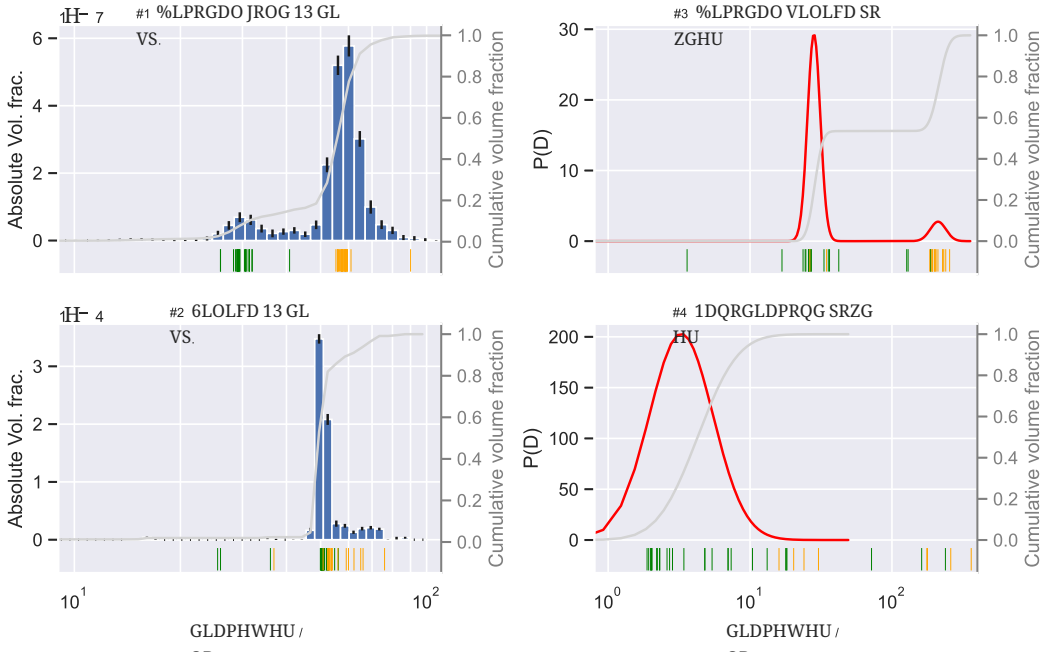

图 2. 使用 McSAS3（左）和 SASfit（右）得出的数据集 1 至 4 的体积加权分布示例。McSAS 直方图柱表示该柱所覆盖范围内的积分体积分数，SASfit 结果则表示体积加权概率函数。两者均在次坐标轴上以灰色绘制累积分布函数（CDF）。下方的标记表示参与者提交的总体 1（绿色）和总体 2（橙色）均值。

图 2 给出了该散射图样的一种可能解决方案，使用

McSAS3 [17]给出了该散射图样的一种可能解。此示例解显示存在两个总体。分别在直径范围为 $20 \leq D$(nm) ≤ 40 和$40 \leq D$(nm) ≤ 80内分析这些总体后，可得总体1和总体2的体积加权总体均值及宽度（标准差），如表1所示。作为比较，还列出了使用SASfit [18]进行示例拟合得到的数值。请注意，我们不能断言这些解是“正确”的，但它们可作为示例。

表1. 使用 McSAS 和 SASfit 对数据集 1 至 4 进行示例拟合得到的均值和宽度，涵盖总体 1 和 2。请注意，下标 v 表示体积加权， n 表示数量加权值。有可用数据时，括号内为不确定度。

数据集	总体	软件	结果（纳米）			
			μ_v	SD_v	μ_n	SD_n
1	P1	McSAS	30.8(4)	3.8(4)	-	-
		SASfit	31.2	5.0	29.1	4.7
	P2	McSAS	59.0(2)	6.2(4)	-	-
		SASfit	58.8	5.3	57.4	5.2
2	P1	McSAS	52.2(6)	4.4(1)	-	-
		SASfit	50.7	1.1	50.6	1.1
	P2	SASfit	56.1	8.5	52.4	7.9
3	P1	SASfit	28.5	3.0	27.6	2.9
	P2	SASfit	214.8	27.4	204.5	26.1
4	P1	SASfit	4.26	2.20	1.92	0.99

2.2. 数据集 *2*：二氧化硅纳米颗粒

第二个数据集是同一系列中的 nPSize 10，其中散射体由标称直径为 60nm[14]、标称为单峰的二氧化硅窄分布组成。近期讨论表明，该样品也可能含有少量稍大粒径的总体（参见表1）。McSAS3 的示例拟合（以及此处未展示的 SASfit 拟合）确实表明存在少量宽分布颗粒，而且相当多的参与者也得出了相同结论。

2.3. 数据集 *3*：*AS-40* 与 *250* 纳米二氧化硅粉末的混合物

数据集 3 由实验室内部制备，将两种球形二氧化硅材料按 1:1 的质量比混合而成。较小的二氧化硅球体通过冷冻干燥 Ludox AS-40（Sigma-Aldrich，直径约 22 nm）制得；较大的球体则采用 Stöber 工艺合成：将正硅酸四乙酯（TEOS，Sigma-Aldrich，98%）加入乙醇（Sigma-Aldrich，96%）、水和氢氧化铵溶液（ACS 试剂，约 28%）的混合液中，并在室温下搅拌 24 小时。随后将所得悬浮液离心，用乙醇洗涤，并在 60 ℃ 下干燥过夜。图 2 展示了 SASfit 的示例分布，均值和宽度详见表1。

2.4. 数据集*4*：纳米金刚石粉末

数据集4的测量对象是从柏林的 PlasmaChem GmbH 购得的纳米金刚石商品样品，目录编号为 PL-D-G02。这些球状金刚石颗粒以干粉形式供应，标称直径为（4 至 6）nm。PlasmaChem 网站上提供了该材料的一些技术细节和证明其用途的更多参考资料 [19]。图 2 展示了 SASfit 的示例分布，均值和宽度详见表1。

2.5. 电子显微图

图 3 展示了构成四个数据集的散射体的电子显微图。数据集 1 和 2 的图像由法国原子能和替代能源委员会`a l'´energie atomique et aux ´energies alternatives (CEA) 测量，并存放于 Zenodo 存储库 [20]。

数据集 3 和 4 的样品图像使用现场配备的电子显微镜拍摄。这两组样品的实验细节如下：将粉末分散于乙醇中，超声处理至少 5 分钟，制备 TEM 样品。取一至三滴所得悬浮液，滴加到覆有花边状或多孔碳膜的铜 TEM 栅格上。使用 JEOL 2200FS 仪器进行 TEM 观察，工作电压为 200 kV。采用明场 TEM 模式采集图像，并使用高对比度光阑增强颗粒的可见度；但对于数据集 4 中直径约为 10 nm 的纳米金刚石，则改为获取高分辨率 TEM 图像。

3. 数据读入与校正

这 46 份有效提交的答题表在进行比较之前，必须经过若干处理步骤。作者信息和报告的分析结果分别读入，以便进行匿名化。评估数据按以下顺序进行了校正：

(i) 对部分工作表进行了人工校正，以确保条目位于正确的列中，便于读取；删除多余信息；将小数逗号改为小数点；在与作者沟通后补充缺失信息，等等。人工校正这一步所需的频率可能高于预期。(ii) 对软件包名称进行排序，并缩短为足以识别它们的最简名称。对于仅使用过一两次的软件包，将其归入“其他”，以免损害匿名化效果。(iii) 将加权类别限定为“体积”、“数量”或“未定义”。(iv) 根据第一总体均值的取值范围，判断作者（最有可能）是否：

人为　　因素：小角散射数据分析循环比对结果 1　　0

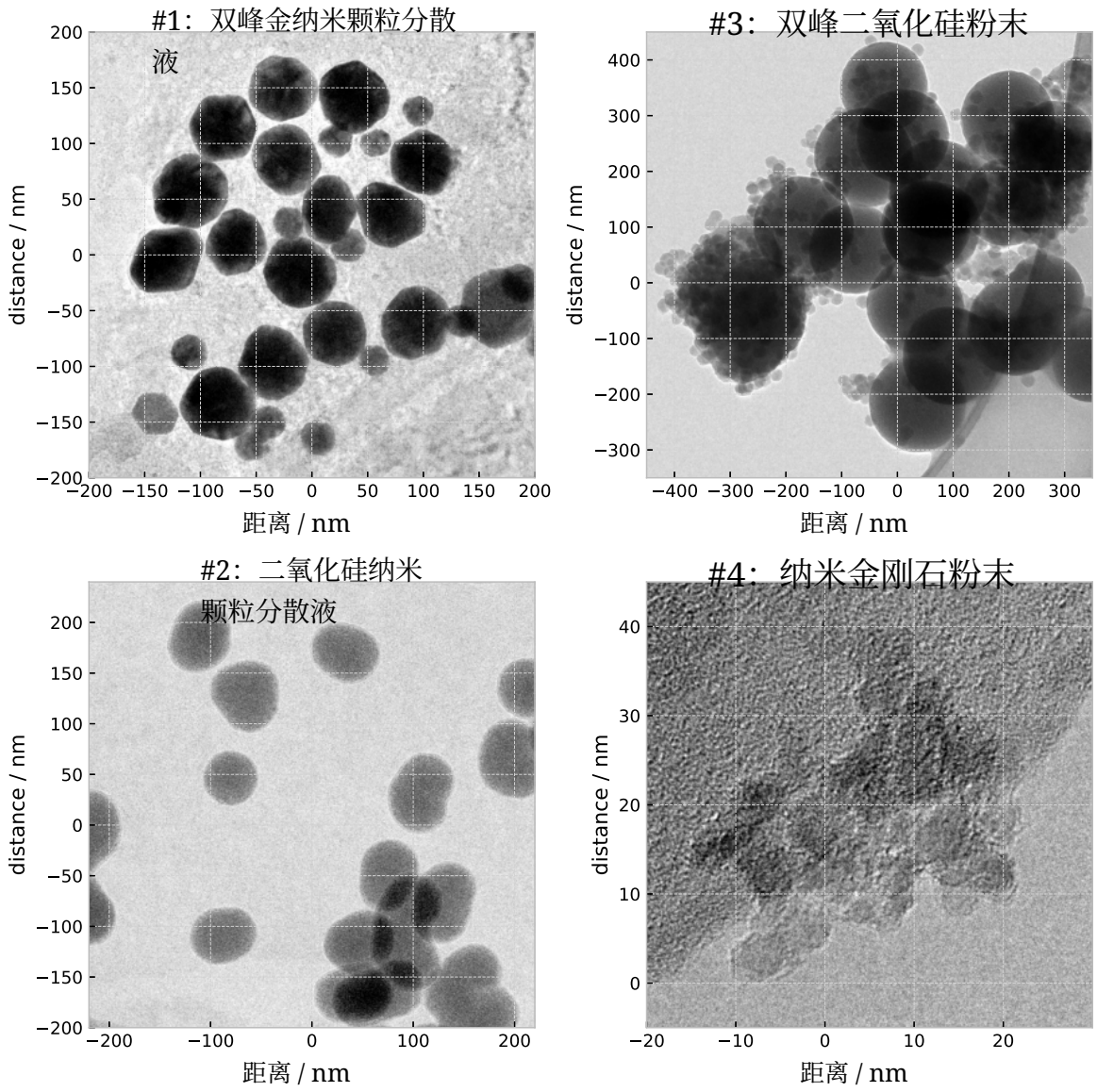

图 3. 构成四个数据集的散射体的电子显微图。数据集 1 和 2 的图像来自 nPSize项目代码库 [20]。数据集 3 和 4 的样品图像是在现场拍摄的。

(a)（72 %）：将“尺寸”理解为“半径”（校正因子设为 2）；(b)（4 %）：遗漏了数据集单位信息（额外校正因子为 10）；(c)（4 %）：错误地校正了数据集单位，或以 $^{\circ}$Angstr¨om 为单位报告信息（额外校正因子为 0.1）；(d)（40 %）：使用了 SasView 软件包，但报告的是 SasView“多分散性”单位下的尺寸分布宽度，而非以标准差表示的总体宽度（将校正因子设为平均半径）。其他软件包也可能以其他方式报告分布宽度，但目前尚不清楚，因此未作校正。

(v) 由于提交条目数量有限，因此未应用离群值检验来进一步排除提交的数值。(vi) 最后，对拼接数据集中的条目进行随机排序，使其不再按录入顺序排列。

4. 结果与讨论

4.1. 总体统计

并非所有参与者都对所有数据集进行了拟合，而且参与者在数据集中识别出的总体数量也不尽相同。答题表不可避免地透露出样本中可能存在多个总体，但没有指出哪些样本可能如此。图4显示了针对每个数据集，提交了某个总体数值的参与者人数，以及他们使用了什么软件来检测该总体。

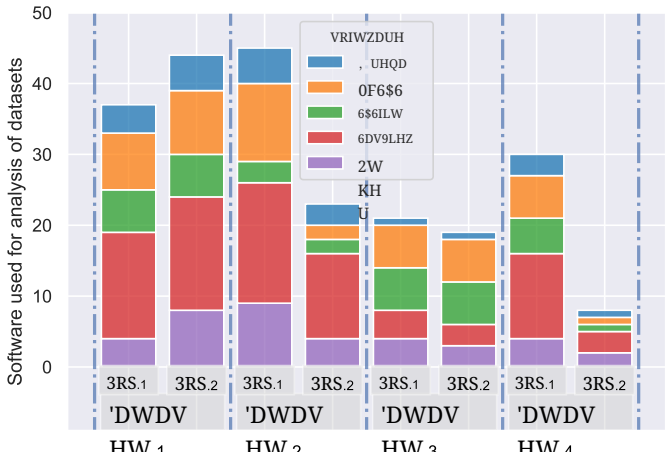

图4. 各数据集中每种总体的条目数量，可从每个柱状图的总高度读出，并按软件细分。数据集 1 和 3 来自名义上具有双峰分布的样品。

大多数参与者在数据集 1 和 3 中识别出了双峰总体，约有一半的参与者在数据集 2 中添加了第二个总体，而几乎没有人认为数据集4中存在第二个总体。与数据集 1 和 2 相比，数据集 3 和 4 更具挑战性，只有约一半的参与者提交了这两个数据集的结果。

这些分析所用的软件包主要有四种，按字母顺序排列§：“Irena” [21], 、“McSAS” [22] 和/或“McSAS3“ [17], 、“SasFit” [18]和“SasView” [23]。请注意，有些软件包（如 Irena）提供多种优化方法。其他用户使用了较不常见的软件，其中包括（按字母顺序排列）：

§ 当软件名称为缩写时，本文未将其拼写完整，原因是这样更便于阅读，而且这些软件通常以其名称而非完整拼写的反向缩略词为人所知。

- autosaxs（内部软件），
- 通过 BioXTAS RAW 使用 GNOM [24],
- PySAXS [25],
- SAXS数值反演 [26]。
- XSACT [27],

以及若干内部开发的代码库。

鉴于数据集带来的挑战多种多样，分析同样可以受益于借助软件包中实现的不同方法。因此，图4展示了每个数据集所使用的软件包。从中可以看出，各软件包的使用分布相当均衡。这一方面表明生态系统十分健康，但另一方面也使比较变得复杂，因为每个软件包报告分布宽度、体积分数和拟合优度等参数的方式可能各不相同。

4.2. 稍等片刻：数量加权与体积加权

需要注意的是，各种软件包通常会根据数量加权分布或体积加权分布来呈现总体统计量。例如，平均尺寸可以表示为按数目计的平均尺寸，也可以表示为按体积计的平均尺寸（或者更准确地说，按质量计）。

我们可以以加权样本矩的定义为基础，用数学方式表达这些量。由此可以利用表2中的公式定义总量（零阶原始矩）、均值（一阶原始矩）、方差（二阶中心矩）以及任何更高阶的（中心）矩$k > 2$ （其不确定性逐渐增大）：

表 2. 定义由 N 个贡献组成的总体的数量加权与体积加权矩 m。下标 n 和 v 分别表示数量加权和体积加权。 D_i 是物体（样品） i 的直径， $v_{f,\,i}$ 是同一物体的体积分数，

k	含义	数量加权	体积加权
$k = 0$	积分值	$m_{0,\mathrm{n}} = 1$	$m_{0,\mathrm{v}} = v_f = \sum_{i=1}^{N} v_{f,i}$
$k = 1$	样本均值	$\mu_{\mathrm{n}} = \frac{1}{N}\sum_{i=1}^{N} D_i$	$\mu_{\mathrm{v}} = \frac{1}{v_f}\sum_{i=1}^{N} D_i v_{f,i}$
$k \geq 2$	方差、偏度、峰度 , etc.	$m_{k,\mathrm{n}} = \frac{1}{N}\sum_{i=1}^{N}(D_i - \mu_{\mathrm{n}})^k$	$m_{k,\mathrm{v}} = \frac{1}{v_f}\sum_{i=1}^{N}(D_i - \mu_{\mathrm{v}})^k\, v_{f,\,i}$

由此可得宽度 σ ，它由方差 m_2 计算得出：

$$\sigma = \sqrt{m_2} \tag{1}$$

此外，还可以通过以下方式选择性地校正抽样偏差，以获得无偏矩：

$$m_{k,\mathrm{unbiased}} = m_k \frac{N}{N-1} \tag{2}$$

抛开数学不谈，这两种加权方式在实际中的区别在于，体积加权均值总是大于数目加权平均值，而且分布越宽，差异越大（其他矩也不可互换，因为它们定义了不同的总体分布）。尽管先前的研究发现，小角散射数据集中的信息与体积加权分布高度吻合 [28], ，但解析拟合方法仍可使用数目加权参数对尺寸分布建模，只不过随着分布变宽，较小尺寸端的不确定性会不断增加。由于解析建模软件包通常不会显示这种分布不确定性，用户可能会误以为整个范围内的数目加权分布都具有同等精度。例如，本次循环比对中数据集 1 的结果（见下文）表明，有些软件包所显示的数值代表什么，可能存在误解；也可能是体积加权参数与数目加权参数本质上有所不同这一点并未得到明确说明。

4.3. 数据集 *1* 示例

请注意，由于各条目包含大量信息，此处只能展示和讨论其中一部分。建议读者自行下载匿名化结果及配套的 *Jupyter* 笔记本，以便进一步研究 *[9]*。

数据集 1 或许是最简单明了的，因此非常适合作为讨论实际结果的起点。图5试图在一张图中尽可能展示相关信息。由于尺寸分布相当狭窄，体积加权平均值和数目加权平均值足够接近，可以在同一张图中展示。

可视化结果显示，各条目之间的跨度约为 ± 7 %，按所报告散射体尺寸均值计算的95%置信区间为此；值得注意的是，50%的条目与中位数均值相差不到1.5%。所报告的宽度差异则大得多：50%的条目与中位数宽度相差在44%以内。这很可能是由于各软件包对分布宽度的报告方式不一致。尽管说明要求将宽度报告为标准差，但许多数值远在这一规定的合理范围之外，通常徘徊在0到1之间。由此可得出的一个初步结论是，由于报告方式不一致，所报告的宽度基本无法用于比较。因此，在没有可接受的解决方案之前，本文余下部分将主要关注所报告的总体均值。

另一个值得关注的方面是各种软件包结果的聚类。例如，McSAS 的结果形成了一个簇，位置略偏于均值右侧。该簇稍稍偏右，可能是因为结果的体积加权影响了总体均值。不过，这一解释并非普遍适用，因为报告的数量加权值和体积加权值分布在整个范围内。

最后，　显然，一些软件

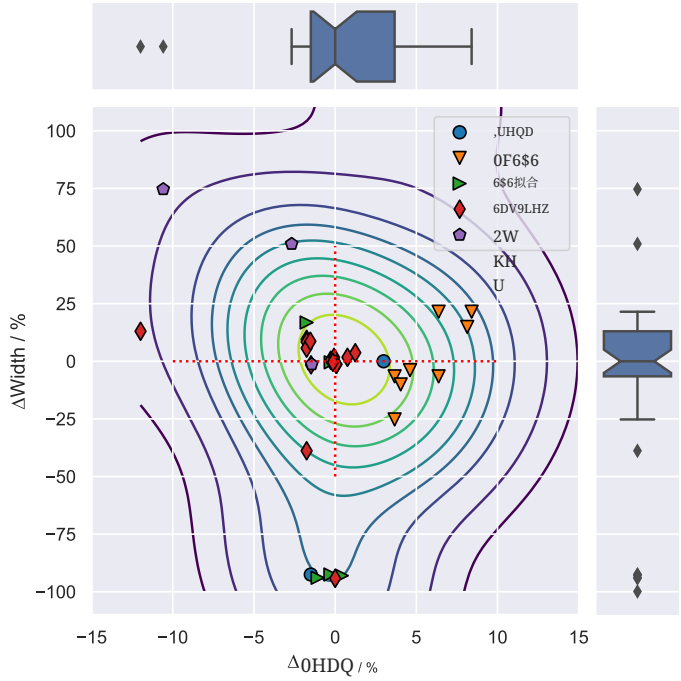

图5. 数据集 1 中最小总体的均值和宽度相对于中位数的偏差（以百分比表示），按软件区分。二维核密度估计（核密度估计）以等高线表示，中位数以红色虚线十字标出，缺口箱线图显示相应均值或宽度值的四分位数，须线则表示 95%置信区间。

包的用户在所报告总体值（即均值和宽度）所采用的加权方式方面存在知识差距。SasView 用户中有 44% 表示这些值采用体积加权，而 51% 的用户则报告称这些值代表数量加权值（另有少数用户采取模棱两可的态度，完全没有报告加权方式），这一点便证明了上述情况。对于 SasFit，这两个比例分别为 52% 和 43%。需要说明的是，SasView 和 SASfit 报告的总体统计量均采用数量加权，但 SASfit 默认以体积加权形式绘制分布，从而进一步造成混淆。这似乎是一个用户界面（UI）问题，因为 McSAS 和 Irena 的用户似乎不存在这种混淆：其中分别有 97% 和 78% 的用户报告称这些值代表体积加权值。

4.4 全部数据集的发现

比较各数据集中每个总体的提交总体均值的相对离散度后，可以明显看出，较具挑战性的粉末类样品呈现出大得多的离散度（图6）。这可归因于数据集 3 中显著结构因子的存在所带来的复杂性，以及数据集4中近似分形的宽分布与其底层结构因子的共同影响。所选结构因子或结构因子（局部）体积分数参数的变化，都会显著影响得出的均值。

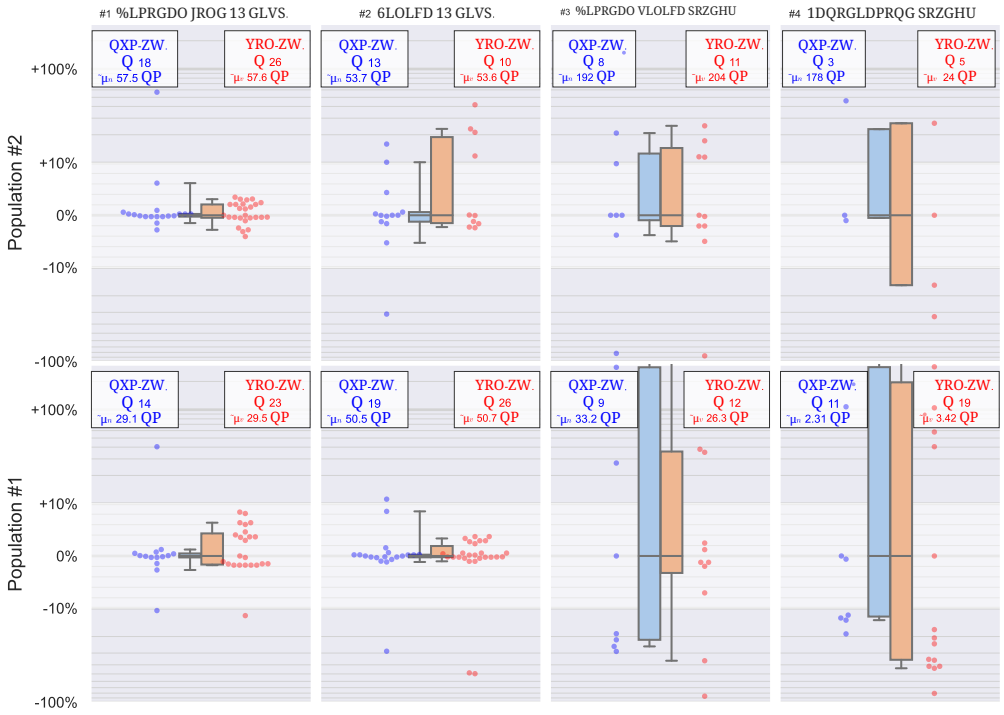

图6。 提交给 DARR 的均值精度。该图以各数据集各总体中报告均值相对于中位数均值的偏差百分比表示，并按数量加权（蓝色）和体积加权（红色）区分；两者分别以各自的中位数均值为参照 $\bar{\mu}_n$ 和 $\bar{\mu}_v$。数据以箱线图表示（显示四分位数，须线表示95%置信区间），并附有各个数据点。图中采用对称对数刻度，较浅的阴影区域表示从 $-10 \leq \bar{\mu}\% \leq 10$ 开始的线性区域。每个数据集的两个总体均在插入框中标出了其中位数均值的绝对值和条目数量。与分散样品相比，较具挑战性的粉末样品所提交数值的离散程度更大。

同样，粒径分布模型的差异也会对最终结果产生影响。

这意味着，可以分别区分数据集 1 和 2 的条目偏差，以及数据集 3 和 4 的条目偏差。完成这一处理后（图7），我们可以得出结论：对于低浓度分散液，总体均值测定值的95%置信区间可以控制在约百分之十以内。然而，参与者给出的宽度结果并不那么一致，这可能是由于前述各软件包在报告方式上存在差异；就目前情况而言，结果的差异可能超过100%。对粉末样品结果进行同样的分析后发现，结果分布过于宽泛，无法使用，这表明盲测

不同实验室之间的粉末结果比对可能尚不可靠。或许可以通过就此类样品的分析方法达成一致来改善这一状况（参见第6段）。

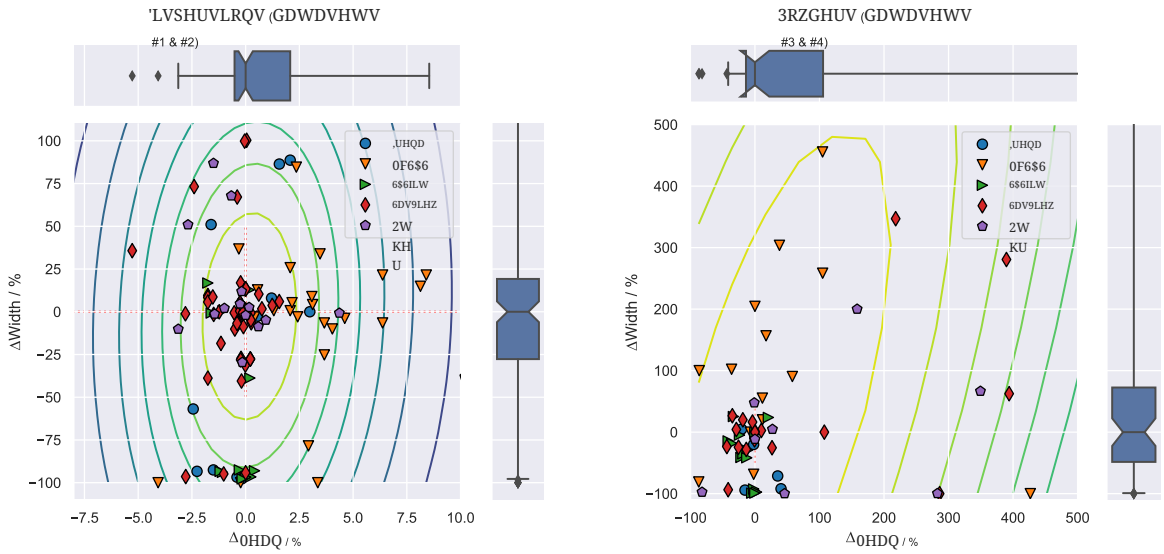

图7. 数据集 1 和 2（左侧）以及数据集 3 和 4（右侧）中两个总体的所有数据点，显示为相对于总体均值中位数的偏差（横轴）和相对于总体宽度中位数的偏差（纵轴）。数据点按软件分组。二维核密度估计以等高线表示，中位数以红色虚线十字标出，缺口箱线图显示相应均值或宽度值的四分位数，须线表示95%置信区间。

4.5. 体积分数

参与者被要求在可用时填写各总体的体积分数信息。为确定这些数值，数据已缩放至绝对单位 ‖。

虽然体积分数有明确的定义，但结果并未体现这一点，提交的数值分布范围大得无法使用，尤其是分散体的数值（图8）。这种差异的来源，以及由此确定的修正途径，并不明了。唯一但不尽如人意的结论是，需要开展统一协调工作，并对体积分数的计算和呈现方式进行交叉核查。

5. 关于自我评估经验

提供了多个信息字段，至少可以间接反映参与者的经验。这些字段包括从业年限、其工作中小角散射所占的比例

‖ 不过，对于金刚石粉末和双峰二氧化硅样品，可能使用了不正确的厚度，因为采用的是根据其X射线吸收确定的材料表观厚度，而不是其所在实际容器的厚度。

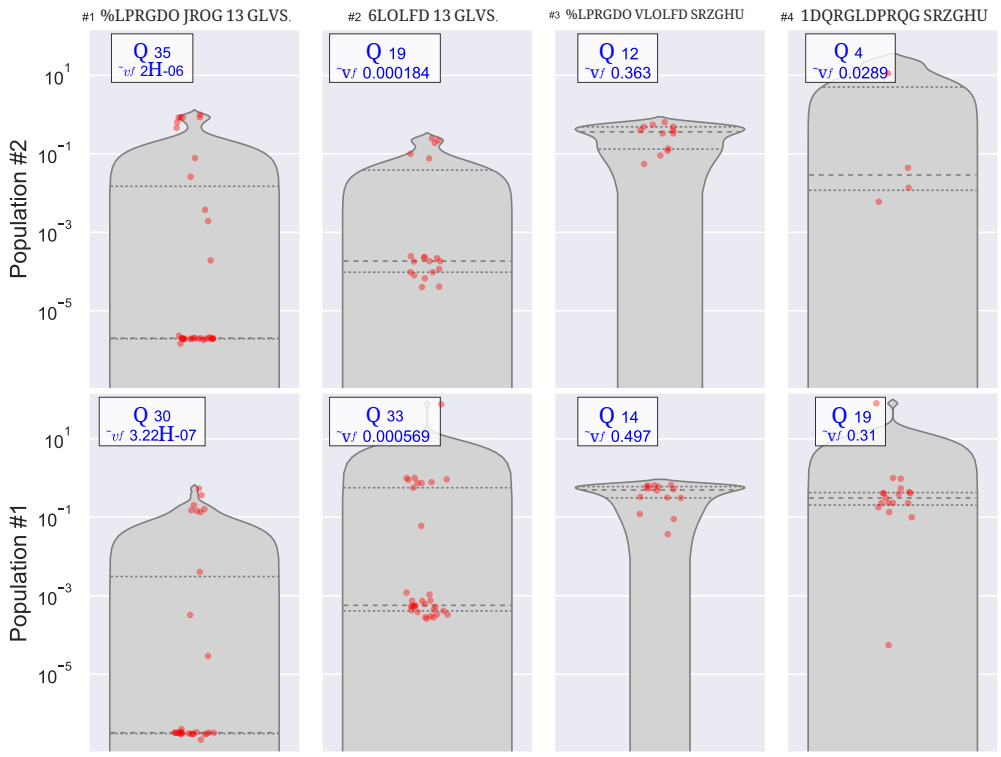

图8. 参与者报告的各数据集中各总体的体积分数。小提琴图以虚线显示四分位数，红点表示各项结果。每幅图的文本框中除了列出报告的体积分数数量外，还显示中位体积分数。

职业生涯，以及对自身知识水平的评估（以1到10为量表）。虽然这些信息只能非常粗略地量化每个人从事散射工作的经历，但我们可以尝试从 ¶中获得一些见解。正如预期的那样（图9），自我评估的知识水平与从事小角散射工作的累计年数存在弱相关性；该累计年数根据其职业生涯中从事小角散射工作的百分比乘以职业生涯年数计算得出。换句话说，在该领域工作得越久、投入越多，他们对自身工作知识的估计就越高。

可以预料，这些指标中有些会与结果接近中位数结果的程度高度相关，前提是中位数接近目标值。图10表明，即便存在这种相关性，相关程度也很弱。这令人遗憾，因为这意味着我们并不能

¶ 缺失的信息包括参与者的专业领域、他们多年来日常工作中小角散射组分的变化、尤其是分析经验等。



图9. 小角散射知识水平与小角散射累计经验的关系，显示出较弱的相关性，符合预期。

随着经验增加而自动变得更好。一种解释是，参与者的真正才华受到软件报告功能局限的阻碍，同时也受到该领域中那些徒增精神负担、毫无必要的二分法的束缚（参见第6段）。因此，要作为一个群体不断进步，我们显然不能只是变老。

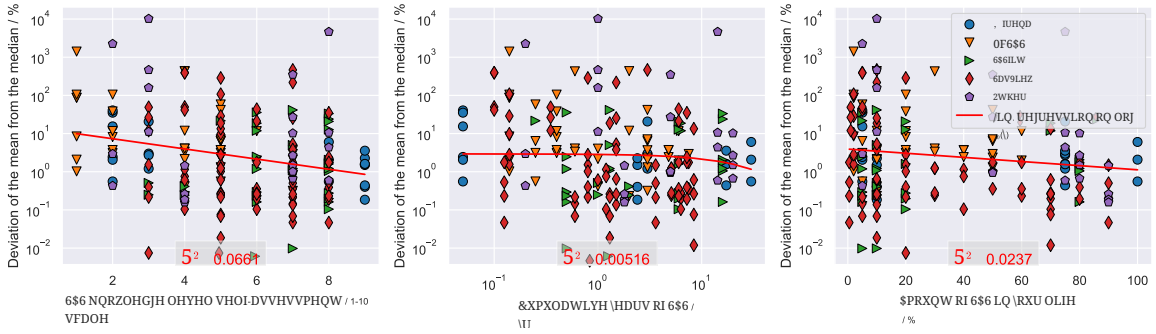

图10. 若干经验指标（小角散射知识、小角散射累计年限以及小角散射在生活中的占比）与接近中位数结果之间仅有微弱的相关性证据。遗憾的是，结果显示二者几乎没有相关性。

6. 可立即采取的改进措施

除分散体的总体均值外，本研究发现其余总体参数的离散程度很大，这凸显了人为因素可能给整个小角散射数据解释过程带来显著的不确定性。本研究估算的这种不确定性，其量级远大于仅由数据采集和校正所引起的不确定性 [8, 15, 29]。散射数据解释中的一些困难，在研究人员甚至尚未开始分析时就已出现：他们面临着五花八门的单位、非标准化的数据格式，

以及定义不明确的数据实践。要求不熟悉相关领域的用户深入了解各种冗余单位及其使用情况，掌握每种分析方法的陷阱与不足，并理解它们在各类拟合软件中的实现差异，这无疑抬高了入门门槛。更进一步的问题是，社区中的许多人似乎并未意识到这道门槛的存在（更糟的是，有人甚至将其视为一种成人礼），而这些都很容易导致不尽如人意的解释。

为缓解这一问题，我们的社区应避免因缺乏约束和明确定义而主动令用户困惑。换言之，仪器负责人应与软件开发人员合作，就一套一致的单位和定义达成共识，并且自身也必须遵循这些规范。其次，应制定通用指南（或可由 CanSAS 或国际晶体学联合会工作组制定），说明如何应对常见样品类型的数据分析难题，而不是仅仅依赖局部的知识传递。最后，软件包的用户应花些时间阅读软件文档，并理解软件所呈现的数值。反过来，软件文档也可以编写一些易于理解的章节 [30]。

因此，可以通过以下方式立即改善实验室间结果一致性：

(i) 逐步统一各个软件包所报告的信息，例如，以分布矩（总值、均值、方差、偏度和峰度）的形式呈现通用总体信息，并报告每个总体的信息。(ii) 提供用户指南，说明如何处理标准散射分析问题，并为稀溶体系和浓体系提供稳健的模型建议及调整方法。此外，还应提供用于粗略估算交叉核查和结果验证的其他方法，例如，确保尺寸与 q 范围等相匹配。(iii) 在拟合中引入并使用切实合理的数据不确定度估计（例如，使用 [15]中采用的方法），使可靠的数据点比不可靠的数据点具有更大的权重；顺带一提，这也将带来 (iv) 提供一致且可比较的拟合优度度量，以评估拟合。若不确定度估计良好，这些拟合优度度量也将具有意义。(v) 去除陈腐、耗时却又不必要地令人困惑的二分法，以及因此必须进行单位换算所带来的错误风险 [+]。以下列举的内容并不全面：

- 应定义 Q 的默认单位（例如 nm^{-1}）
- 应定义 I 的默认单位（例如 $(m \cdot sr)^{-1}$）
- 尺寸应始终指物体的全长（直径），而不是有时对某些形状指其半长（半径）。这样，

[+] 尽管任何群体如何就两个选项之一达成一致，本身就是一个冲突重重的棘手问题

在转换到其他散射体形状时，可避免出现因子二偏移错误。

- 总体信息应采用体积加权，以更贴切地反映散射图样所含的信息；也可以采用数量加权，但无论采用哪种方式，都必须明确并反复注明。

7. 未来循环比对实验设计的建议

没有任何实验是完美的，本次循环比对也不例外。未来的迭代可考虑以下改进建议。

要汇聚更广泛的群体，并深入了解一致性的演进（或退步），一种方法是定期开展规模较小的循环比对研究。做法可以很简单，例如每学期提供一个数据集。这样还有一个额外好处，就是逐步建立数据和拟合示例库。

进一步区分用户与软件各自产生的影响，将有助于确定不确定性的主要来源。为此，某些循环比对研究可以规定使用特定的软件包或特定的模型。

对拟合质量进行交叉评估，或许能够评估人们对于何为“良好拟合”的共识程度。这也可能有助于确定可作为比较基准的目标拟合或最佳拟合。在此基础上，每位参与者都应保留重现某一拟合所需的全部元数据，以便在建立共识拟合后更好地确定产生分歧的来源。

8. 结论

一项循环比对研究已开展，用于研究个体研究人员对小角散射图样分析数值结果的影响。在分析之前，需要对若干结果进行校正，以弥补数据输入和结果报告过程中因遗漏或错误应用单位换算而产生的领域特定二分差异。

结果显示，对于由低浓度球状散射体分散液组成的样品，所确定的总体均值分布范围较窄，其中一半的数据项落在中位数均值的±1.5%以内。对于浓缩粉末更具挑战性的散射图样，分布范围从较大到极大不等，其中一半的数据项落在中位数均值的±44%以内。这可能是因为结果还受到结构因子模型和体积分数选择的影响。对于这两类样品，所确定的总体宽度差异都非常大，而且由于各种软件包报告的参数存在差异，这些结果显然无法直接比较（尽管名义上的答案格式明确规定宽度应以标准差的形式表示）。最后，对于某些软件包究竟是将拟合参数报告为体积加权值还是数量加权值，仍存在相当大的混淆。

此外，尽管参与者认为自己使用该方法的时间越长，知识水平就越高，但这与更接近

中位数均值并没有很强的相关性。因此，第 6 段提出了其他建议（即除了积累多年的专业经验之外），这些建议有助于提高所得结果的可比性，尤其是宽度和体积分数的可比性。实施其中一部分措施必将对提高小角散射所得科学结果的可比性产生积极影响。

9. 数据可用性

数据以及用于数据清理、分析和绘图的 Jupyter 笔记本，可在 Zenodo 开放访问数据存储库中获取 [10, 9]。我们强烈鼓励对这些数据进行进一步分析和拓展。

10. 致谢

K.T. 感谢 NIST-NRC 博士后奖学金计划提供的支持。本文提及某些商业设备、仪器、材料或软件，旨在充分说明实验步骤。此类说明并不意味着 NIST 对其进行推荐或认可，也不意味着所提及的材料或设备必然是该用途的最佳选择。本研究部分经费由欧洲计量创新与研究计划（EMPIR）项目 No. 17NRM04 提供。

11. 参考文献

[1] Dong Y, Etienne A, Frolov A, Fedotova S, Fujii K, Fukuya K, et al. 使用最大分离距离法对实验数据中的团簇分析开展原子探针断层成像实验室间研究。《显微镜与微量分析》。2019;25(2):356-66。可访问：https://www.cambridge.org/core/product/identifier/S1431927618015581/type/ journal_article。[2]Madsen IC, Scarlett NVY, Cranswick LMD, Lwin T. 国际晶体学联合会粉末衍射委员会定量物相分析循环比对结果：样品 1a 至 1h。Journal of Applied Crystallography。2001;34:409-26。[3] Scarlett NVY, Madsen IC, Cranswick LMD, Lwin T, Groleau E, Stephenson G, et al. 国际晶体学联合会粉末衍射委员会定量物相分析循环比对结果：样品 2、3、4、合成铝土矿、天然花岗闪长岩和药品。Journal of Applied Crystallography。2002;35:383-400。[4] Rennie AR, Hellsing MS, Wood K, Gilbert EP, Porcar L, Schweins R, et al. 通过聚苯乙烯乳胶样品的循环比对测量，了解小角中子散射仪器及数据归约。Journal of Applied Crystallography。2013;46(5):1289-97。[5] Whitfield PS. 使用中子粉末衍射对复杂样品进行定量物相分析。重新审视粉末衍射委员会定量物相分析循环比对中的样品 #4。《粉末衍射》。2016;31:192-7。[6] Trewhella J, Vachette P, Bierma J, Blanchet C, Brookes E, Chakravarthy S, et al. 循环比对方法详细评估了生物分子小角散射数据的可重复性，并生成用于基准评估的共识曲线。《晶体学报》D辑。2022;78(11):1315-36。可访问：https://doi.org/10.1107/S2059798322009184。[7] Osterrieth J, Rampersad J, Madden DG, Rampal N, Skoric L, Connolly B, et al. 根据 BET 方程计算的表面积可重复性如何？ChemRxiv。2022:16。

[8] Pauw BR, K¨astner C, Th¨unemann AF. 纳米颗粒尺寸分布定量：小角 X 射线散射实验室间比对结果。Journal of Applied Crystallography。2017;50:1280-8。[9]Pauw BR, Smales GJ, Rosalie JM. 小角散射数据分析循环比对：匿名化结果、图表和 Jupyter 笔记本 [数据集]。Zenodo；2023。Zenodo 上的数据集。可访问：https://doi.org/10.5281/zenodo.7509710。[10] Pauw BR, Smales GJ. 小角散射数据分析循环比对：提供给参与者的原始[数据集]。Zenodo；2023。Zenodo 上的数据集。可访问：https://doi.org/10.5281/ zenodo.7506365。[11] Deumer J, Gollwitzer C. npSize SAXS 数据 PTB [数据集]。Zenodo；2022。Zenodo 上的数据集。可访问：https://doi.org/10.5281/zenodo.5886834。[12] Krumrey M, Ulm G. PTB 四晶单色器光束线上的高精度探测器校准。《核仪器与方法物理研究 A 辑：加速器、光谱仪、探测器及相关设备》。2001;467-468:1175-8。[13] Wernecke J, Gollwitzer C, M¨uller P, Krumrey M. 真空内 PILATUS 1M 探测器的表征。《同步辐射杂志》。2014;21:529-36。[14]Bartczak D, Hodoroaba VD. 具有非球形、非单分散尺寸分布及准确纳米颗粒浓度的候选参考物质的开发与验证报告（交付成果 D3） [报告]。Zenodo；2022。Zenodo 上的报告。可访问：https://doi.org/10.5281/zenodo.5886834。[15] Smales GJ, Pauw BR. MOUSE 项目：获取可溯源的宽范围 X 射线散射信息的严谨方法。Journal of Instrumentation。2021;16:P06034。[16]Pauw BR, Smith AJ, Snow T, Terrill NJ, Th¨unemann AF. 模块化小角 X 射线散射数据校正流程。Journal of Applied Crystallography。2017;50:1800-11。可访问：https://doi.org/10.1107/S1600576717015096。[17] Pauw BR, Bressler I. McSAS3：一种用于分析分散型 X 射线散射数据的蒙特卡罗方法；2022。最后查阅日期：2022-08-29。可访问：https://github.com/BAMresearch/ McSAS3。[18] Kohlbrecher J, Breßler I. SASfit 的更新：将解析表达式和数值模型拟合到小角散射图样。Journal of Applied Crystallography。2022;55(6)。可访问：https://doi.org/10.1107/S1600576722009037。[19] GmbH P. 纳米金刚石，特纯，G02 级；2022。最后查阅日期：2022-09-01。可访问：https://shop.plasmachem.com/nano-diamonds-and-carbon/32-112-nanodiamonds-extra-pure-grade-g02.html#/weight-1g。[20] Metrology P. nPSize CEA 图像作为二维数组 [数据集]。Zenodo；2021。Zenodo 上的数据集。可访问：https://doi.org/10.5281/zenodo.5886834。[21] Ilavsky J, Jemian PR. Irena：用于小角散射建模与分析的工具套件。Journal of Applied Crystallography。2009;42(2):347-53。可访问：https://doi.org/10. 1107/S0021889809002222。[22] Bressler I, Pauw BR, Th¨unemann AF. McSAS：用于从散射图样中提取模型参数分布的软件。Journal of Applied Crystallography。2015;48(3):962-9。可访问：https://doi.org/10.1107/S1600576715007347。[23] Developers S. 用于小角散射分析的 SasView；2022。最后查阅日期：2022-11-22。可访问：https://www.sasview.org。[24] Hopkins JB, Gillilan RE, Skou S. BioXTAS RAW：改进小角 X 射线散射数据归约与分析的免费开源程序。Journal of Applied Crystallography。2017;50(5):1545-53。可访问：https://doi.org/10.1107/S1600576717011438。[25] Tach´e O, Spalla O, Thill A, Carriere D, Testard F, Sen D. pySAXS：用于 SAXS 数据处理的开源 Python 软件包及图形用户界面；2015–2017。最后查阅日期：2021-06-19。可访问：http: //iramis.cea.fr/en/Phocea/Vie_des_labos/Ast/ast_sstechnique.php?id_ast=1799。[26] Bender P

磁性纳米颗粒的分布。New Journal of Physics. 2017;19(7):073012。可访问：https://dx.doi.org/10.1088/1367-2630/aa73b4。[27] Xenocs。XSACT：X射线散射数据处理与分析；2022年。最后检查日期：2022-11-22。可访问：https://www.xenocs.com/saxs-products/xsact-software/。[28] Pauw BR、Pedersen JS、Tardif S、Takata M、Iversen BB。通过蒙特卡罗方法从小角散射数据中进行尺寸分布反演的改进与注意事项。Journal of Applied Crystallography。2013;46(2):365-71。[29] Schavkan A、Gollwitzer C、Garcia-Diez R、Krumrey M、Minelli C、Bartczak D 等。悬浮液中金纳米颗粒的数浓度：与实验室方法相比，SAXS 和 spICPMS 作为可溯源方法。Nanomaterials。2019;9(4):20。[30] Wuttke J、Cottrell S、Gonzalez MA、Kaestner A、Markvardsen A、Rod TH 等。可持续数据处理软件协作开发指南。Journal of Neutron Research。2022;24:33-72。