Comment on "Machine learning enhanced analysis of EBSD data for texture representation"

Check for updates

Helmut Schaeben[1] ✉ & K. Gerald van den Boogaart[2,3]

Our concerns apply to the inadequate ways statistical distributions of crystallographic orientations are compared and occasionally confirmed to agree sufficiently well. The authors of "Machine learning enhanced analysis of EBSD data for texture representation"[1] suggest a method to replace an EBSD dataset of crystallographic orientations with a much smaller synthetic dataset preserving the texture. They claim that their "texture adaptive clustering and sampling" algorithm generates datasets of a few hundred crystallographic orientations, realizing an equivalent crystallographic orientation distribution as the initial dataset. To prove the principle and substantiate their claim of equivalent orientation distributions, the authors content themselves with (i) a visual inspection of the crystallographic pole density function, in fact, of three crystallographic "pole figures" and (ii) Kolmogorov–Smirnov tests for each of the three Euler angles of the crystallographic orientations individually. However, these criteria are insufficient to confirm equivalence of orientation distributions, they do not provide scientific evidence to substantiate the authors' claim that "texture adaptive clustering and sampling" generates crystallographic orientations in terms of their Euler angles representing the same texture.

Our dissent against (i) is that different distributions of crystallographic orientations may have several identical crystallographic pole figures. Our objection to (ii) is threefold.

- The stochastic independence of observations required by the Kolmogorov–Smirnov test is most likely to be violated by spatially referenced crystallographic orientations sampled with EBSD experiments scanning the crystallites of a material specimen. The individual Euler angles are presumably affected by spatially induced stochastic dependence acquired from the grain fabric.

- Failure to reject the null-hypothesis of agreement of two distributions of the Kolmogorov–Smirnov test must not be interpreted as proof of their equivalence. As with every statistical test of significance, the Kolmogorov–Smirnov test provides insightful inference in case of reasonable rejection of its null-hypothesis only. Since the replacement datasets are designed to be small, probabilities of false non-rejection tend to be large. Applying the Kolmogorov–Smirnov test to any original dataset and a replacement dataset consisting only of its median will always fail to suggest rejection of its null-hypothesis of equivalent distributions.

- Furthermore, the agreement of the marginal distributions for each of the three Euler angles does not generally imply agreement of their joint distribution, the exception being the case of joint stochastic independence of the three Euler angles.

Last but not least, we remind you that examples of "successful" application of some mathematical or statistical method of data analysis do not substitute the mathematical proof of its properties and performance. Mistaking examples as proof of principle is all the more dangerous if the means of validation by examples are inadequate themselves. Then, an example may be envisioned, where the algorithm goes astray, however unnoticed by the applied means of validation.

In the following we shall elaborate on our opposition to the authors' way to compare statistical distributions of crystallographic orientations in terms of counter examples.

Counter example 1: Different crystallographic orientation density functions with a few coincident pole figures

We assume orthorhombic crystal symmetry (*mmm*) with all axes of unit length ("pseudo cubic"). Then we define two different orientation density functions f_1, f_2 each composed of three equally weighted components labeled $\psi_{ij}, i = 1, 2, j = 1, \ldots, 3$, of de la Vallée Poussin type ψ_{dlVP} of a halfwidth of $b_\kappa = \pi/6$ centered at orientations $g_{11} = (\pi/2, 0, 0)$, $g_{12} = (\pi/2, \pi/2, \pi/2)$, $g_{13} = (0, \pi/2, 0)$ and $g_{21} = (\pi, 0, 0)$, $g_{22} = (\pi, \pi/2, \pi/2)$, $g_{23} = (\pi/2, \pi/2, 0)$, respectively, in terms of Euler angles according to the *zxz*-convention, i.e.,

$$f_1(g; \psi_{11}, \psi_{12}, \psi_{13}) = \tfrac{1}{3}\big(\psi_{\mathrm{dlVP}}(g; g_{11}, b_\kappa) + \psi_{\mathrm{dlVP}}(g; g_{12}, b_\kappa) + \psi_{\mathrm{dlVP}}(g; g_{13}, b_\kappa)\big)$$
$$f_2(g; \psi_{21}, \psi_{22}, \psi_{23}) = \tfrac{1}{3}\big(\psi_{\mathrm{dlVP}}(g; g_{21}, b_\kappa) + \psi_{\mathrm{dlVP}}(g; g_{22}, b_\kappa) + \psi_{\mathrm{dlVP}}(g; g_{23}, b_\kappa)\big).$$

Since the two sets of crystallographic orientations are related by a rotation of $\pi/2$ about the *z*-axis, they are not crystal symmetrically equivalent. Once the orientation density functions have been computed, their pole density functions for the crystal forms of Miller indices $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$, $(0, 1, 1)$, $(1, 1, 1)$, $(1, 1, 3)$, $(1, 0, 2)$, $(1, 2, 3)$ are calculated. Then, the two

[1]Geoscience Mathematics and Informatics, TU Bergakademie Freiberg, Freiberg, Germany. [2]Applied Stochastics, TU Bergakademie Freiberg, Freiberg, Germany. [3]Helmholtz Institute Freiberg for Resource Technology, Helmholtz Zentrum Dresden Rossendorf e.V. (HZDR), Freiberg, Germany. ✉e-mail: schaeben@tu-freiberg.de
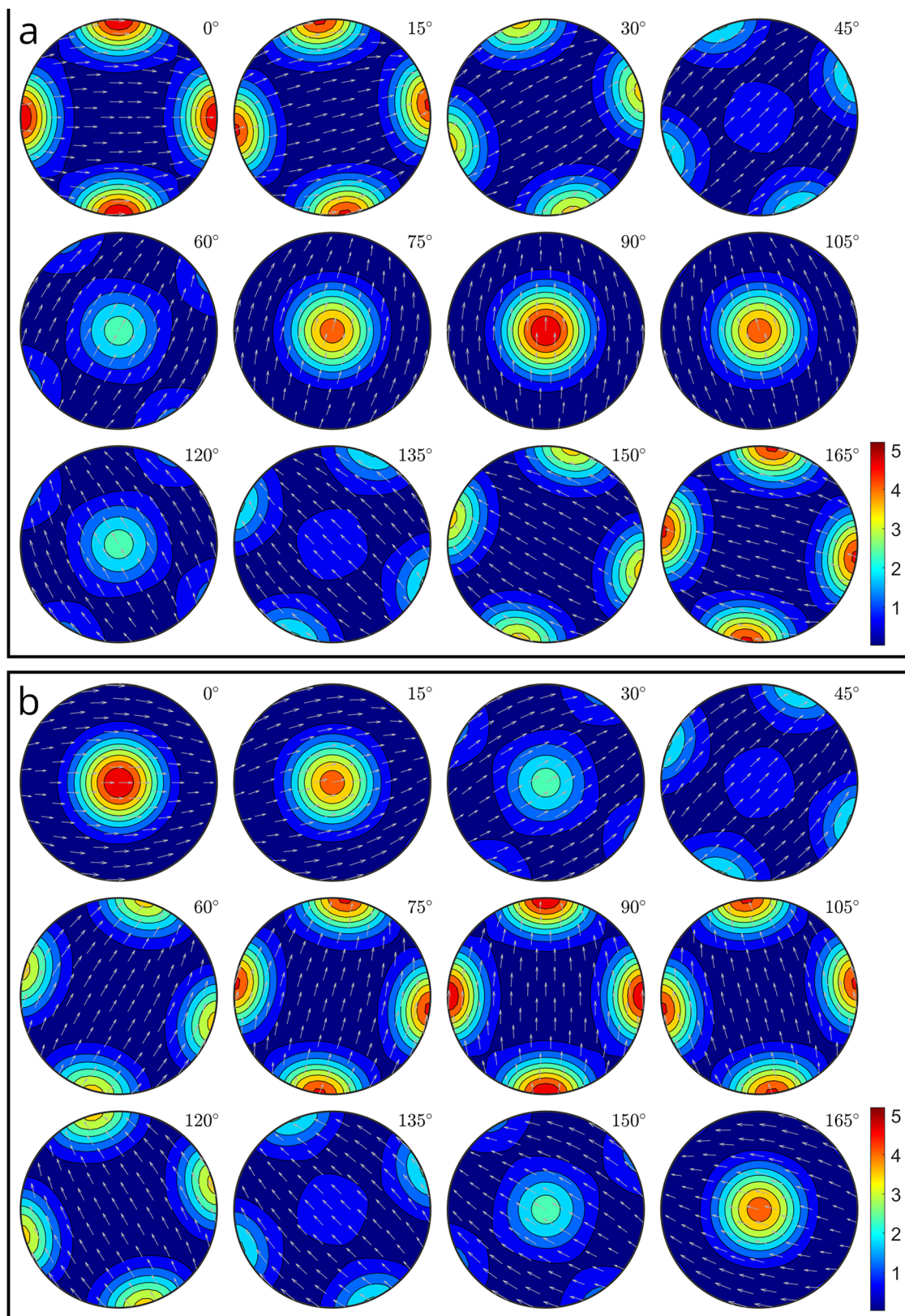
Fig. 1 | σ-sections of two different crystallographic orientation density functions. a σ-sections of orientation density function f_1. **b** σ-sections of orientation density function f_2.

orientation density functions are plotted as σ-sections, Fig. 1, and their pole density functions for the 8 crystallographic forms listed above are plotted as equal-area projections of the upper hemisphere onto the unit disk (Fig. 2).

While the orientation density functions are different, the first six polfigures are identical, only the (102)- and (123)-pole figures are different. This example may appear as a very special case of ambiguity. However, it is reminded, that any two crystallographic orientation
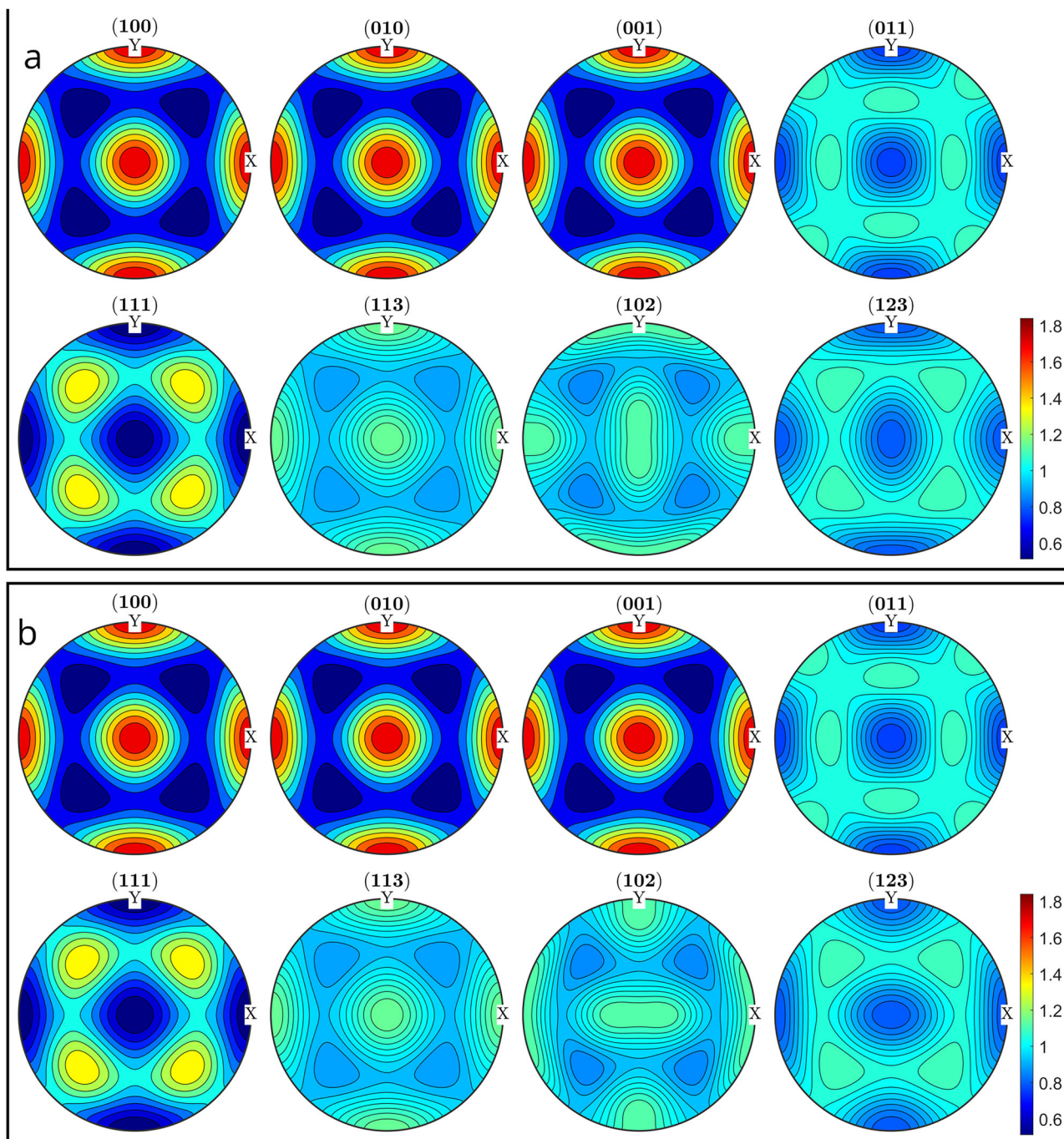
Fig. 2 | Pole density functions of several crystallographic forms of two different crystallographic orientation density functions. a Pole figures of orientation density function f_1. **b** Pole figures of orientation density function f_2.

density functions which differ in their odd Fourier coefficients only, have identical crystallographic pole density functions for all crystal forms[2,3]. For more details, the reader is referred to MTEX documentation[4].

Counter example 2: Different crystallographic orientation density functions with coincident marginal distributions of their three Euler angles

Visualizing the marginal distributions of individual Euler angles (φ_1, Φ, φ_2) of the crystallographic orientations sampled from the orientation density functions f_1 and f_2, respectively, reveals almost perfect matches of the histograms of the marginal distributions, Figs. 3 and 4.

The corresponding univariate cumulative frequency distribution functions are very close, Fig. 5. Their comparison in terms of Kolmogorov–Smirnov tests is illusive because (i) non-rejection of its null-hypotheses must not be interpreted as their acceptance and (ii) univariate marginal distributions do not generally define a unique joint multivariate distribution, i.e., identical univariate marginals may stem from different joint multivariate distributions.

Conclusion

Visual inspection of pole figures may be deceptive because they are images of the even part of crystallographic orientation density functions only. It may lead to false conclusions. Referring to the three
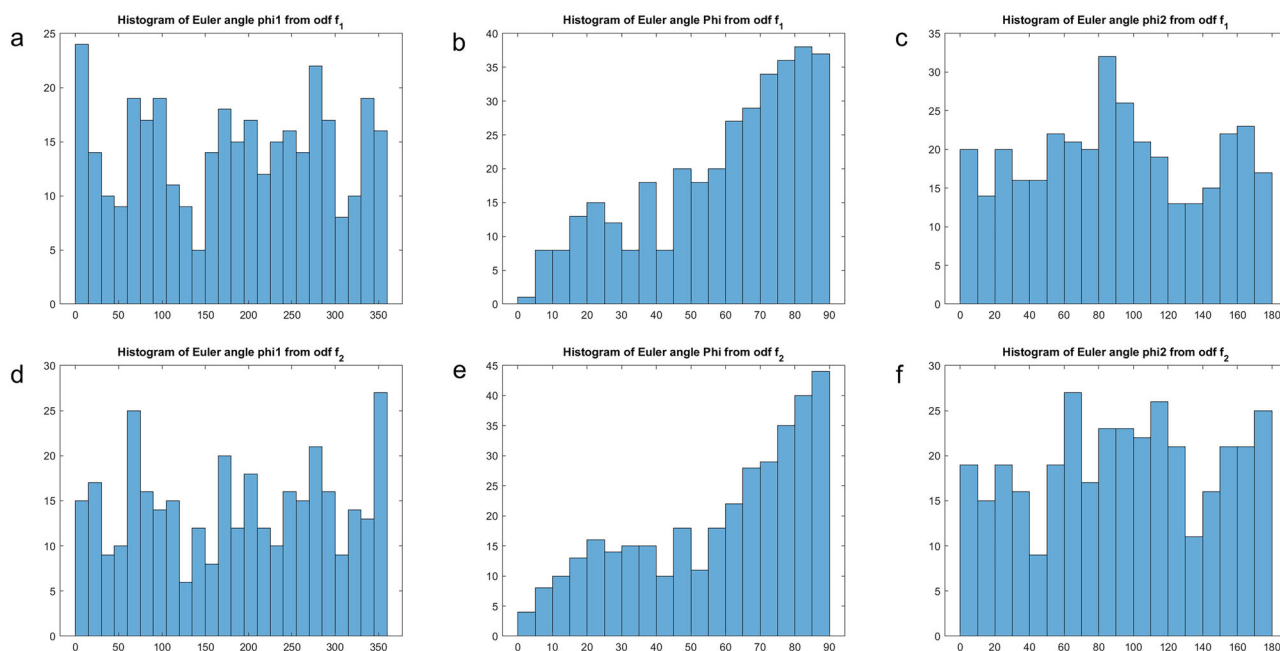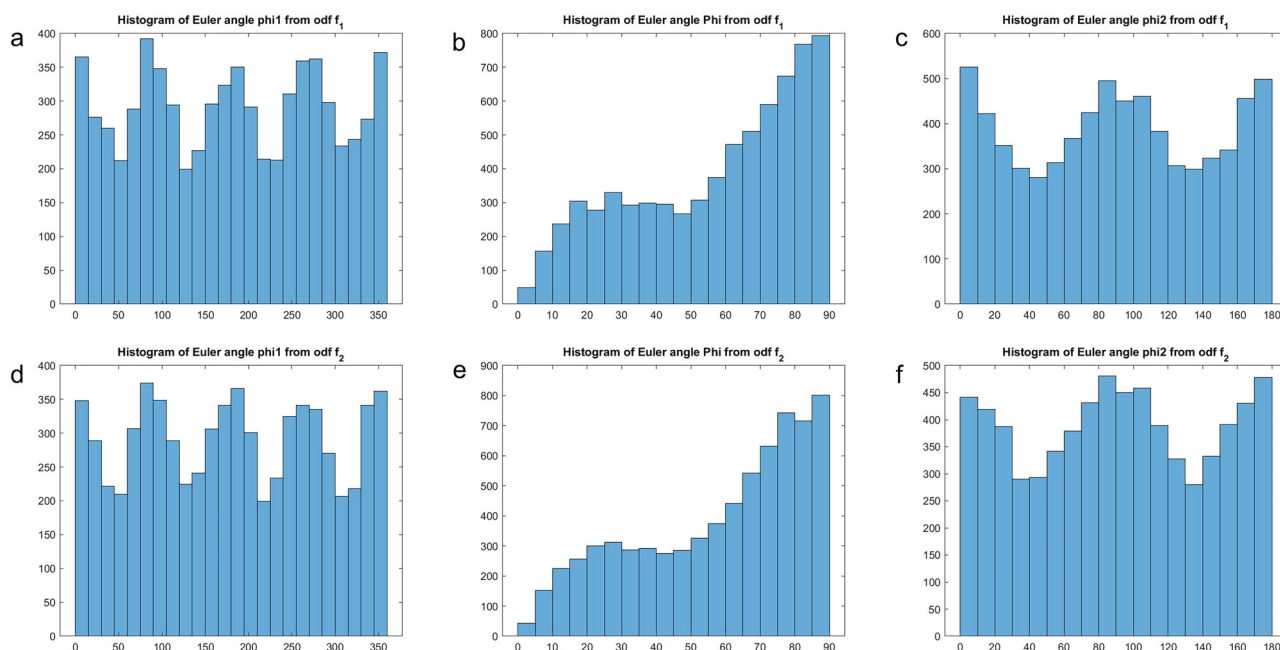
Fig. 3 | Histograms of marginal distributions of 350 Euler angles sampled from two different crystallographic orientation density functions. Top row: Histograms of the marginal distributions of 350 Euler angle φ_1 (**a**), Φ (**b**), and φ_2 (**c**) of crystallographic orientations sampled from the orientation density function f_1;

Bottom row: Histograms of the marginal distributions of 350 Euler angle φ_1 (**d**), Φ (**e**), and φ_2 (**f**) of crystallographic orientations sampled from the orientation density function f_2.

Fig. 4 | Histograms of marginal distributions of 7000 Euler angles sampled from two different crystallographic orientation density functions. Top row: Histograms of the marginal distributions of 7000 Euler angle φ_1 (**a**), Φ (**b**), and φ_2 (**c**) of crystallographic orientations sampled from the orientation density function f_1;

Bottom row: Histograms of the marginal distributions of 7000 Euler angle φ_1 (**d**), Φ (**e**), and φ_2 (**f**) of crystallographic orientations sampled from the orientation density function f_2.

marginal distributions of individual Euler angles, e.g., comparing their cumulative frequency distributions, does not generally provide additional information to distinguish different crystallographic orientation density functions, because marginals do not generally define the joint distribution. Thus, the authors' proof of principle and validation of

their TACS approach are invalid. Users may be better off with any of the conventional methods, e.g. methods based on kernel density estimation[5].

To compare distributions of crystallographic orientations, it takes distributions of crystallographic orientations.
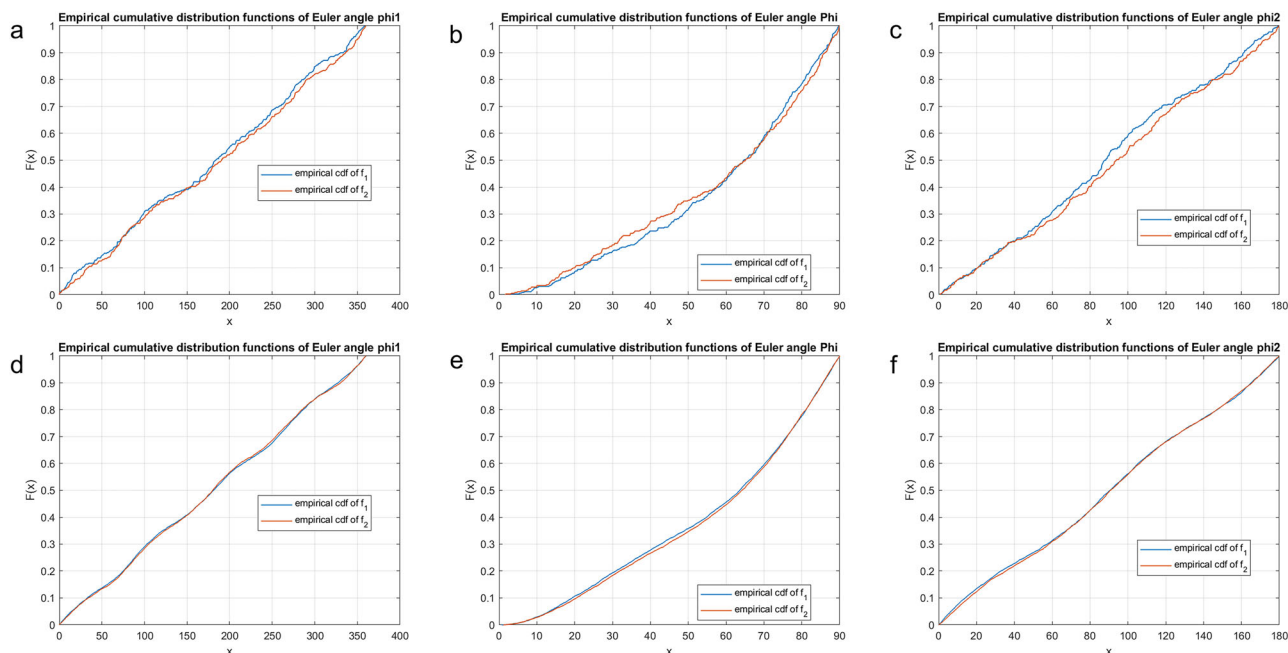
Fig. 5 | Comparison of empirical cumulative distribution functions of Euler angles for two different crystallographic orientation density functions. Empirical cumulative distribution functions of Euler angles φ_1 (**a**), Φ (**b**), and φ_2 (**c**) of 350 crystallographic orientations and φ_1 (**d**), Φ (**e**), and φ_2 (**f**) of 7000 crystallographic orientations sampled from the crystallographic orientation density functions f_1 and f_2, respectively.

References

1. Wanni, J., Bronkhorst, C. A. & Thoma, D. J. Machine learning enhanced analysis of EBSD data for texture representation. *npj Comput. Mater.* **10**, 133 (2024).
2. Matthies, S. On the reproducibility of the orientation distribution function of texture samples from pole figures (Ghost phenomena). *Phys. Stat. Sol. (b)* **92**, K135–K138 (1979).
3. Wagner, F. *Contribution a l'étude des modifications de texture par déformation et transformation dans les matériaux de basse symétrie.* Thèse Docteur ès Sciences, Université de Metz, France, 13–14 (1983).
4. Hielscher, R. (2024) https://mtex-toolbox.github.io/PoleFigure2ODFAmbiguity.html (accessed 9 February 2025).
5. Schaeben, H., Bachmann, F. & Fundenberger, J. J. Construction of weighted crystallographic orientations capturing a given orientation density function. *J. Mater. Sci.* **52**, 2077–2090 (2017).

Acknowledgements

Author contributions

H.S.: Identified the error in the annotated manuscript, pointed out the argument on non-provability by the statistical argument, provided the calculations using MATLAB/MTEX, and wrote the first version of the manuscript. K.G.v.d.B.: Provided the original counterexample, added the arguments on the incorrect use of statistics, and improved and expanded the manuscript accordingly. The authors confirm that all authors have read and approved the manuscript.

Funding

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Helmut Schaeben.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.