TEST (2021) 30:59–63
https://doi.org/10.1007/s11749-021-00760-4

DISCUSSION

Comments on: Recent advances in directional statistics

Kanti V. Mardia[1,2]

Received: 13 January 2021 / Accepted: 30 January 2021 / Published online: 19 March 2021

I should first like to congratulate the authors on giving a very comprehensive review of what has happened since (Mardia and Jupp 1999) (we are grateful for making our book as its base) in directional statistics (DS); the task is not an easy one as the field has grown dramatically since then. In my discussion I will take a complimentary approach—a broad historical perspective starting from the beginning of the field. More so, 2020 was the Florence Nightingale's Bicentenary Year as she was born on 12 May 1820, and indeed, on the 1 December 2020 there was the "Florence Nightingale Bicentenary Celebration" in Oxford with an Exhibition. As expected, Florence Nightingale's rose diagrams to visualize data on a torus (trivariate) were highlighted in the talks and in the exhibition. In fact, about 160 years ago (1858) she gave us these visualizations to make the subject accessible to policy makers namely, Crimean War deaths through Rose Diagrams and others (see also Hedley 2020). Visualization of data on a torus (bivariate) has also appeared in Biochemistry—the celebrated Ramachandran plots which appeared in 1963; see, for example, Mardia (2013). I have extended its visualization (Mardia 2012) by opening the torus on a repeated 3×3 grid so that modes and antimodes can easily be seen.

Beginning of DS. Now, I will muse on the historical growth of the subject. I begin with the timeline in Fig. 1. Of course the papers by von Mises (1918) and Fisher (1953) are the pioneering papers for modelling data on the circle and the sphere, respectively, but I believe Watson-Stephens and their other collaborators woke up this sleeping beauty.

I came into the field totally by chance via the uniform scores in the context of a bivariate nonparametric test. Details are given in Mukhopadhyay (2002). I was fascinated by this field and the more I read, the more I could see that one needs to create a terminology and unify the material. When I moved to Hull, the late Toby Lewis encouraged me to write a monograph on the subject and, he gave me generous

This comment refers to the invited paper available at: https://doi.org/10.1007/s11749-021-00759-x.

✉ Kanti V. Mardia
k.v.mardia@leeds.ac.uk

1 University of Leeds, Leeds, UK

2 University of Oxford, Oxford, UK

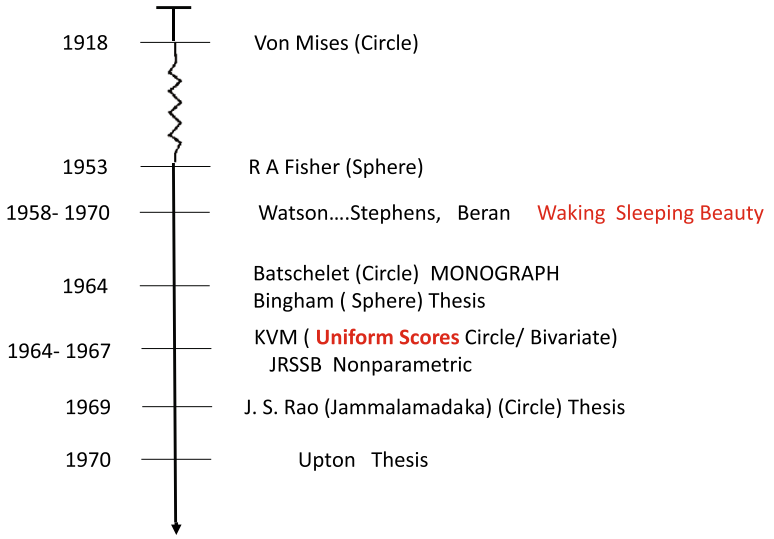

Fig. 1 Birth of directional statistics timeline (not to scale)

sabbatical leave and I submitted my manuscript of the book in 1971 accepting that there was no time to expand spherical chapters in this edition. On the other hand, I think The late Toby suggested to write the JRSSB discussion paper (Mardia 1975) which I completed in Leeds. It turned out there were many eminent discussants to the paper including the late Henry Daniel, the late David Kendall, David Cox, John Kingman, the late Toby Lewis, the late Edward Batchelet, J.S.Rao, Madan Puri, John Kent, Chris Bingham. The late Professor Henry Daniels while seconding the vote of thanks said "It is a young and growing field, not yet too specialized for amateurs to venture into, and it is therefore an admirable topic for discussion at one of our meetings.... As such it is to be warmly welcomed." Indeed, the field thus came to the attention of the main stream statisticians.

Then, followed in 1975–1980: the Leeds Symposium on Directional Statistics under an EPSRC (then SRC) grant. Visitors for collaborating under the grant included The late C. G. Khatri, Barndorff-Nielsen, Rudy Beran, Kit Bingham, Tom Downs, John Kent, J. S. Rao, The late Geoff Watson, The late Michael Stephens, Madan Puri, The late Edward Batchelet, Peter Jupp (Post Doc 76-78).

One key contribution from my work with Khatri is "Matrix distributions on Stiefel manifold" (Khatri and Mardia 1977), which included the Fisher–Bingham distribution on the Stiefel manifold. There was some supplementary material for the paper such as the normalizing constant for the Fisher–Bingham distribution, which was given in our departmental report (Khatri and Mardia 1975)—the arXiv of those times. However, the report has not been accessible for subsequent work such as that of Kume et al. (2013) and on GvM_2 (Generalized bivariate von Mises distribution). Also during this period, the seeds were sown for Shape Analysis (Mardia et al. 1977); the work was motivated by the problem to test whether the triangles for a central place data are

equilateral. Motivated by this practical example, David Kendall took this topic much further later on.

Period 1980–2000. Then came the period of consolidation 1980–2000 with books, edited volumes, reviews. Namely, the books by Batschelet (1981) (circle), Fisher et al (1987) (sphere), Fisher (1993)(Circle), Mardia and Jupp (1999) (title changed to "Directional Statistics" from "Directional Data Analysis" in the 1st edition of Mardia (1972) to emphasize significant new material. Also, there was a change of publisher, and in the UK, the publication year is 2000.) Jammalamadaka and SenGupta (2001) (circle), also their R library (CircStats (2001)). Further there were edited volumes: Mardia (1988), Mardia (1989), Mardia (1992) which has a group photo of directional statisticians in 1989. A review of the subject appeared in Jupp and Mardia (1989) which has a slightly different emphasis than the current paper as our joint paper aimed to give a unified view of the theory of directional statistics, 1975–1988 after my discussion paper of 1975.

Period 2001–2010. The period 2001–2010 was beginning of new waves in the subject brought about by the relationship between shape analysis and directional statistics, and the exposure to Molecular Biology/Protein Bioinformatics. There were conferences, books, and further reviews of the subject but I regard that moving in to Molecular Biology was a major step forward during this period. Protein structure can be summarized in terms of dihedral angles and this part appears in the basic Biochemistry textbooks under the topic of Ramachandran plots. In fact, my introduction to Bioinformatics was by chance as described in Mukhopadhyay (2015) including my collaboration with Thomas Hamelryck on solving the probabilistically local structural prediction problem (Boomsma et al. 2008). Incidentally, (Mardia 2013) while highlighting statistical approaches to three key challenges in protein structural bioinformatics has given a list of open problems to be solved.

Period 2010–2020. More new waves arose in the period 2010–2020. The subject became further consolidated with the books by Ley and Verdebout (2017) and Ley and Verdebout (2018). Appearance of many software is as described in the paper. Beginning of Triennial Dedicated Conferences ADISTA starting from 2014 but Covid restrictions led the last November to "Virtual Symposium on Directional Statistics" organized by Karlsruhe Institute of Technology.

The manuscript I have of the paper has about 50% divided between the text and the references and the bib file has over 1700 references. But, as the authors are aware, it cannot be comprehensive. I will give just one example which might go into the bibliography. The estimation of the axis of a geometrical figure is truly directional and in particular for α helices having vital application in drug discovery. Mardia et al. (2018) and Alfahad et al. (2018) have given some methods of estimation which leads to a regional change point problem. This specific problem of estimating the axis of the curves is fundamentally directional as it is estimating a vector of the rotation matrix. Though there is a definite overlap between shape analysis and directional statistics, in shape analysis the rotation matrix is a nuisance parameter. Incidentally, Downs and Mardia (2000) is the precursor to Downs and Mardia (2002) and Downs (2003).

From an historical perspective, to me it seems important to distinguish between increment and step-change research though both are important for any subject to grow. I will give a few specific examples of the step-change advances which might

help the readers. Let me start with the Möbius transformation-based models of Downs and Mardia (2002). They were the first to filter from the fundamental mathematical paper on Möbius transformation of McCullagh (1996) from the Cauchy distribution to directional statistics, leading to several new important developments, I believe. Another such example is score matching estimators (Mardia et al. 2016), inspired by the Hyvärinen (2005) scoring rule, are introduced to directional distributions though the use for the manifolds is not as simple as we thought. The introduction of score matching approximation in Mardia et al. (2016) explores a totally new direction than Hyvärinen (2005).

It is well known that there is no bivariate distribution in the circular exponential family, which has got both marginal and conditional von Mises distributions as proved by Mardia (1975). Recently, Kato and Pewsey (2015) have introduced a bivariate Cauchy distribution, which has both marginal and conditional distributions of the same form, namely wrapped Cauchy. This is a remarkable result but is it if and only if?

The heart of Torus-PCA (Eltzner et al. 2018) lies in the introduction of torus deformation schemes. Jona-Lasinio et al. (2012) in using the wrapped normal random field have constructed an innovative (simple) truncation rule depending on data variability. Wang and Gelfand (2013) and others have taken the projected normal distribution of Mardia (1972) to a greater height. I have given a few step changes and am sure there will be some future publications on these lines by others.

What of 2021 and beyond—My thoughts. I agree with the comment in the paper "Given this background, predicting how the field might develop over the next 20 years is essentially impossible." Availability of the bibliography by the authors is another helpful development; this keeps on growing. Could there be a depository of DS material including the "dynamic" bibliography, open access papers, conference presentations, historical documents. Courses on DS were given in Leeds University M.Sc. courses a number of years ago and further for some years some basic ideas were introduced in the second year Maths undergraduate and I hope this happens more and more. Can some ideas filter into the High Schools?

I believe there will be separate monographs on the subtopics of DS. We hope that some basic ideas of DS appear in statistical text books. We expect that at least in Biochemistry text books, the bivariate circular distributions are quoted along with the Ramachandran plots.

We may like to see the subject much more in the public domain—two historical examples are the work of Florence Nightingales and another the postal stamp of Poland issued in 1975 with circular histogram. May be we will soon see Florence Nightingale's rose diagrams in relation to Covid-19 in BBC TV presentations?

References

Alfahad M, Kent J, Mardia KV (2018) Statistical shape methodology for the analysis of helices. Sankhya Ser A 80(5):8

Boomsma W, Mardia KV, Taylor CC, Ferkinghoff-Borg J, Krogh A, Hamelryck T (2008) A generative, probabilistic model of local protein structure. Proc Natl Acad Sci USA 105(26):8932–8937

Downs TD (2003) Spherical regression. Biometrika 90(3):655–668
Downs T, Mardia KV (2000) A family of spherical regression models. In: Kent JT, Aykroyd RG (eds) The statistics of directions, shapes and images. Leeds University Press, Leeds, p 223
Downs TD, Mardia KV (2002) Circular regression. Biometrika 89(3):683–698
Eltzner B, Huckemann S, Mardia KV (2018) Torus principal component analysis with applications to RNA structure. Ann Appl Stat 12(2):1332–1359
Fisher RA (1953) Dispersion on a sphere. Proc R Soc Lond Ser A 217:295–305
Hedley A (2020) Florence Nightingale and Victorian data visualisation. Sigificance 17(2):26–30
Hyvärinen A (2005) Estimation of non-normalized statistical models by score matching. J Mach Learn Res 6:695–708
Jona-Lasinio G, Gelfand A, Jona-Lasinio M (2012) Spatial analysis of wave direction data using wrapped Gaussian processes. Ann Appl Stat 6(4):1478–1498
Jupp PE, Mardia KV (1989) A unified view of the theory of directional statistics. Int Stat Rev 57(3):261–294
Kato S, Pewsey A (2015) A Möbius transformation-induced distribution on the torus. Biometrika 102(2):359–370
Khatri CG, Mardia KV (1975) The von Mises-Fisher matrix distributions. Research Report No 1, Department of Statistics, University of Leeds, Leeds. https://www1.maths.leeds.ac.uk/~sta6kvm/reprints/Khatri-Mardia-1975.pdf
Khatri CG, Mardia KV (1977) The von Mises-Fisher matrix distribution in orientation statistics. J R Stat Soc B 34:95–106
Kume A, Preston SP, Wood ATA (2013) Saddlepoint approximations for the normalizing constant of Fisher-Bingham distributions on products of spheres and Stiefel manifolds. Biometrika 100(4):971–984
Ley C, Verdebout T (2017) Modern directional statistics. Chapman and Hall/CRC, Boca Raton
Ley C, Verdebout T (eds) (2018) Applied directional statistics, Modern Methods and Case Studies. Chapman and Hall/CRC, Boca Raton
Mardia KV (1972) Statistics of directional data. Academic Press, London
Mardia KV (1975) Statistics of directional data (with discussion). J Roy Stat Soc B 37(3):349–393
Mardia KV (1988) Directional Data Analysis (Editor). A special Issue of the Journal of Applied Statistics, 15(2)
Mardia KV (1989) Statistics in Earth Sciences (Editor). A special issue of the Communications in Statistics 1989 A10, No.15
Mardia KV (1992) The Art of Statistical Science (Editor). In: honor of G. S. Watson, Wiley
Mardia KV (2012) Statistics of some topics in turbulence and bio-molecular fields. In: Mardia KV, Gusnanto A, Riley AD, Voss J (eds) New statistics and modern material sciences. Leeds University Press, Leeds, pp 1–11
Mardia KV (2013) Statistical approaches to three key challenges in protein structural bioinformatics. J R Stat Soc C 62(3):487–514
Mardia KV, Jupp PE (1999) Directional statistics. Wiley, Chichester
Mardia KV, Edwards R, Puri ML (1977) Analysis of central place theory (with discussion). Bull Internat Stat Inst 47(93–110):139–146
Mardia KV, Kent JT, Laha AK (2016) Score matching estimators for directional distributions. arXiv:1604.08470
Mardia KV, Sriram K, Deane CM (2018) A statistical model for helices with applications. Biometrics 74(3):845–854
McCullagh P (1996) Möbius transformation and Cauchy parameter estimation. Ann Stat 24(2):787–808
Mukhopadhyay N (2002) A conversation with Kanti Mardia. Stat Sci 17(1):113–148
Mukhopadhyay N (2015) A conversation with Kanti Mardia: Part II. In: Dryden IL, Kent JT (eds) Geometry driven Statistics. Wiley, Chichester, pp 59–85
von Mises R (1918) Über die "Ganzzahligkeit" der Atomgewichte und verwandte Fsragen. Phys Z 19:490–500
Wang F, Gelfand AE (2013) Directional data analysis under the general projected normal distribution. Stat Methodol 10(1):113–127