CORRESPONDENCE

Reply to “What Is the Maximum Entropy Principle? Comments on ‘Statistical Theory on the Functional Form of Cloud Particle Size Distributions’”

WEI WU

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, Norman, Oklahoma

GREG M. MCFARQUHAR

Cooperative Institute for Mesoscale Meteorological Studies, and School of Meteorology, University of Oklahoma, Norman, Oklahoma

(Manuscript received 21 December 2018, in final form 21 August 2019)

ABSTRACT

We welcome the opportunity to correct the misunderstandings and misinterpretations contained in Yano's comment that led him to incorrectly state that Wu and McFarquhar misunderstood the maximum entropy (MaxEnt) principle. As correctly stated by Yano, the principle itself does not suffer from the problem of a lack of invariance. But, as restated in this reply and in Wu and McFarquhar, the commonly used Shannon–Gibbs entropy does suffer from a lack of invariance for coordinate transform when applied in continuous cases, and this problem is resolved by the use of the relative entropy. Further, it is restated that the Wu and McFarquhar derivation of the PSD form using MaxEnt is more general than the formulation by Yano and allows more constraints with any functional relations to be applied. The derivation of Yano is nothing new but the representation of PSDs in other variables.

1. Introduction

Wu and McFarquhar (2018, hereafter WM18) examined the application of the principle of maximum entropy (MaxEnt) in the field of cloud physics for determining the theoretical form of cloud particle size distributions (PSDs). MaxEnt was first proposed by Jaynes (1957a,b) and then widely used in various scientific communities [e.g., see Banavar et al. (2010) for a recent general review]. The lack of invariance of Shannon–Gibbs entropy under coordinate transform in continuous cases was realized later by Jaynes himself 10 years after its use was proposed, with the concept of relative entropy and the approach of transformation group proposed by Jaynes (1968) himself for solving this issue. The same problem with the lack of invariance also exists in previous studies applying MaxEnt for deriving cloud PSDs (Zhang and Zheng 1994; Liu et al. 1995; Yano et al. 2016). Based on the approach proposed by Jaynes (1968), WM18 proposed a general framework for the use of the MaxEnt to derive cloud PSDs and a general form

of PSDs. More specifically, WM18 showed that cloud PSDs can be represented by a four-parameter generalized gamma distribution if only one constraint exists, with the constraint variable being in the form of a power-law relation with the particle size.

Yano (2019, hereafter Y19) questions the need for the use of relative entropy by WM18. Y19 ignores the later development of MaxEnt as discussed in WM18, and wrongly claims that the original MaxEnt approach is sufficient and that only a unique variable needs to be specified, which “is a very basic nature of the distribution function discussed in any basic textbooks on statistics and probability.” This claim ignores the later development of MaxEnt. The details are discussed in section 2 and summarized in section 3.

2. MaxEnt and later development

In this section, we correct some of the misconceptions about the development and evolution of the application of MaxEnt contained in Y19.

First, the development of MaxEnt originates from statistical mechanics, and its successful application to

Corresponding author: Wei Wu, weiwu@ou.edu

DOI: 10.1175/JAS-D-18-0374.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the [AMS Copyright Policy](#) (www.ametsoc.org/PUBSReuseLicenses).

an ideal gas inspired Jaynes's initial papers (Jaynes 1957a,b). Therefore, the first conclusion of Y19 that the "applicability of this principle to a particular system cannot be deduced by physical arguments in any obvious manner" is incorrect. However, we agree with Y19 that the application of MaxEnt is not limited to a system at equilibrium, even though that the principle was inspired by an equilibrium system. WM18 never mentioned that the application of WM18 was limited to systems at equilibrium.

Second, Y19 falsely claims that MaxEnt must be applied only to a certain restriction variable and identifying the restriction variable is "the first main challenge." Unfortunately, there are no methods to identify this "restriction variable," at least not in Y19. This is actually the lack of invariance issue for the Shannon–Gibbs entropy that Y19 failed to acknowledge. Y19 clearly misunderstood the lack of invariance issue for Shannon–Gibbs entropy when applying MaxEnt as explained in detail in section 3 of WM18, and discussed in section 4a of Y19 about the probability density function (PDF), which is not supposed to be invariant. The invariant requirement applies to entropy, not to the PDF of any state variable. Obviously, the PDF is not invariant: the PDF represented by one state variable can be easily converted to a PDF represented by another state variable using the relation

$$f_1(x)dx = f_2(y)dy \quad \text{or} \quad f_1(x) = f_2(y)\frac{dy}{dx}. \quad (1)$$

WM18 never stated that functions f_1 and f_2 are the same, and again this is not the lack of invariance issue of Shannon–Gibbs entropy stated in WM18. All the "derivations" in sections 4b and 4c of Y19 are just the transformations of a PDF using Eq. (1) with the additional term dy/dx that is falsely called $I(D)$ in Y19. This is nothing close to the invariant measure in WM18. The additional term dy/dx needs to be applied for any PDF transformation, including all the PSDs derived in WM18. Furthermore, the lack of invariance of Shannon–Gibbs entropy for continuous cases was not first noted by WM18 as implied by Y19. Interestingly, relative entropy was used in three references cited by Y19 (Verkley and Lynch 2009; Verkley 2011; Verkley et al. 2016). This problem has been recognized first by Jaynes (1968) himself and in several subsequent papers, as discussed in Banavar et al. (2010) and other papers as summarized in the introduction of WM18.

It should also be noted that the derived entropy in Eq. (15) of Y19 in section 4d is in the same form as the equation for relative entropy, but only with a very specific invariant measure, $I(D) = dx/dD$, with x being the chosen state (restriction) variable. This shows that

Shannon–Gibbs entropy for continuous cases is ill-defined, and instead should be in the form of relative entropy to be invariant, in contrast to the arguments of Y19. The main shortcoming of the derivation of section 4 in Y19 is that a particular state variable x as a function of particle size D needs to be chosen [and therefore a fixed parameter μ in Eq. (7) of Y19 instead of a variable μ in Eq. (37) of WM18], and there are no methods to decide which variable to choose. Actually, this is just the term dy/dx in Eq. (1), which is universal for the representation of PSDs in another state variable. The derivation of WM18 is more general than that in Y19, and can include more complex situations such as having more than one constraint variables and/or having the constraint variables in any functional relation with the state variable as described in Eq. (27) of WM18.

Third, the constraints do not need to be "strictly conserved" as suggested in the third conclusion by Y19. For example, an ideal gas with slow changing temperature will still have molecule speeds distributed in the form of Maxwell–Boltzmann distribution, with the parameters of the PDF changing accordingly. The system does not have to be at equilibrium. However, identifying the constraints is indeed a challenge. It is important to know what restriction variables and how many are needed, and this depends on the particular physical system. This is explicitly noted in WM18 with the statement that "the value of n_c is determined by the knowledge of the system that is being considered, and can vary according to the behavior of the particular system that is being modeled," with n_c being the number of constraint variables. In the case of cloud PSDs, potential restriction variables can be any variable describing the particle property, such as particle size, projected area, surface area, mass and fall speed.

Fourth, contrary to the statement of Y19, introducing relative entropy does not need one to identify the "restriction variable" as stated in his second point. There is no so-called first main challenge. But identifying the constraints will be a challenge, as stated in his third point and in WM18, and WM18 suggested an approach in the conclusions that "the development of idealized models to simulate the evolution of cloud particles can also provide another perspective, from which the application of MaxEnt may provide more theoretical basis on the appropriate constraint for the system that should be used."

Finally, unlike the false statement in Y19, it should be emphasized that WM18 never stated that the principle of maximum entropy suffers from a lack of invariance. Rather, the lack of invariance occurs only when certain definitions of entropy (e.g., Shannon–Gibbs entropy)

are used in its application. This has been explained in detail in [section 3](#) of [WM18](#).

3. Summary

In conclusion, the arguments made by [Y19](#) do not take into account developments made in the application of MaxEnt after its original development, even though he asserts that he “summarizes an essence of important developments made in the application of the maximum entropy theory” in his comment paper. The use of relative entropy in the application of MaxEnt solves the issue of lack of invariance that exists for the Shannon–Gibbs entropy in the continuous cases, and the approach of transformation group proposed in [Jaynes \(1968\)](#) can be used for a particular physical system. The four-parameter generalized gamma distribution is the form of a cloud PSD for the simplest case with just one constraint variable being in the power-law relation as a function of particle size. The MaxEnt PSD in the form of Eq. (27) in [WM18](#) derived using relative entropy is general and can accommodate more complex situations. The derivation of [Y19](#) is nothing new but the representation of PSDs in other variables.

Acknowledgments. The authors are supported by the office of Biological and Environmental Research (BER) of the U.S. Department of Energy Atmospheric Systems Research Program through Grants DE-SC0014065 and DE-SC0016476 (through UCAR Subcontract SUBAWD000397) and by the National Science Foundation (NSF) under Grants AGS-1213311 and AGS-1762096.

REFERENCES

- Banavar, J. R., A. Maritan, and I. Volkov, 2010: Applications of the principle of maximum entropy: From physics to ecology. *J. Phys. Condens. Matter*, **22**, 063101, <https://doi.org/10.1088/0953-8984/22/6/063101>.
- Jaynes, E. T., 1957a: Information theory and statistical mechanics. *Phys. Rev.*, **106**, 620–630, <https://doi.org/10.1103/PhysRev.106.620>.
- , 1957b: Information theory and statistical mechanics. II. *Phys. Rev.*, **108**, 171–190, <https://doi.org/10.1103/PhysRev.108.171>.
- , 1968: Prior probabilities. *IEEE Trans. Syst. Sci. Cybern.*, **4**, 227–241, <https://doi.org/10.1109/TSSC.1968.300117>.
- Liu, Y., Y. Laiguang, Y. Weinong, and L. Feng, 1995: On the size distribution of cloud droplets. *Atmos. Res.*, **35**, 201–216, [https://doi.org/10.1016/0169-8095\(94\)00019-A](https://doi.org/10.1016/0169-8095(94)00019-A).
- Verkley, W., 2011: A maximum entropy approach to the problem of parametrization. *Quart. J. Roy. Meteor. Soc.*, **137**, 1872–1886, <https://doi.org/10.1002/qj.860>.
- , and P. Lynch, 2009: Energy and enstrophy spectra of geostrophic turbulent flows derived from a maximum entropy principle. *J. Atmos. Sci.*, **66**, 2216–2236, <https://doi.org/10.1175/2009JAS2889.1>.
- , P. Kalverla, and C. Severijns, 2016: A maximum entropy approach to the parametrization of subgrid processes in two-dimensional flow. *Quart. J. Roy. Meteor. Soc.*, **142**, 2273–2283, <https://doi.org/10.1002/qj.2817>.
- Wu, W., and G. M. McFarquhar, 2018: Statistical theory on the functional form of cloud particle size distributions. *J. Atmos. Sci.*, **75**, 2801–2814, <https://doi.org/10.1175/JAS-D-17-0164.1>.
- Yano, J.-I., 2019: What is the maximum entropy principle? Comments on “Statistical theory on the functional form of cloud particle size distributions.” *J. Atmos. Sci.*, **76**, 3955–3960, <https://doi.org/10.1175/JAS-D-18-0223.1>.
- , A. J. Heymsfield, and V. T. Phillips, 2016: Size distributions of hydrometeors: Analysis with the maximum entropy principle. *J. Atmos. Sci.*, **73**, 95–108, <https://doi.org/10.1175/JAS-D-15-0097.1>.
- Zhang, X., and G. Zheng, 1994: A simple droplet spectrum derived from entropy theory. *Atmos. Res.*, **32**, 189–193, [https://doi.org/10.1016/0169-8095\(94\)90059-0](https://doi.org/10.1016/0169-8095(94)90059-0).