The paper uses the one-dimensional, averaged governing equations of atmospheric boundary-layer flows under the assumption of horizontal homogeneity to analyse wind veer. The theory is used to derive mathematical expressions relating wind veer to wind shear and turbulent stress, and it is also used to analyse the physical mechanisms giving rise to wind veer. The theoretical results are then evaluated based on canonical RANS simulations and real observations (i.e., including for example effects of stability and heterogeneity) from 4 different sites. The authors also propose a practical approach to estimate the wind veer based on typical measurement data. The paper presents a detailed analysis of wind veer in both idealised and realistic conditions, and it provides a lot of insights and explanations for why a certain behaviour is observed. As wind veer is likely to become more and more important for wind energy purposes, I believe this work provides valuable insight into the theory and the statistics of wind veer. My main concern is that, as the paper is quite lengthy and heavy on math, more structure is needed to guide the reader. Please find below more detailed comments and suggestions as to how to further improve the manuscript.

Main comments

1. It is not clear from the introduction what the objectives of the study are and how it is organized. After summarizing the literature on the importance of veer, the authors quite abruptly dive into the mathematical derivations and formulations in section 2, and it took me about 20 pages until halfway section 3 to realize what the paper was actually about. The paper is quite lengthy, so more indications of how the paper is structured are needed in the introduction to guide the reader. For instance, it is by no means clear from the introduction that the authors will test the theoretical formulations of veer against RANS and observational data at 4 different sites.

2. Section 2 is quite heavy on math with a lot of different formulations for the same quantity, and it is not always clear what you are going to do with all these formulations. For example, there about 4 different formulations for the wind direction φ (or $\cos\gamma$) (Eq. 6, 13, 25, 29) and no less than 10 for the veer $\partial\varphi/\partial z$ (or $\partial\cos\gamma/\partial z$) (Eq. 9, 10, 12, 14, 15, 16, 21, 24, 38, 39)! More structure and guidance is needed in section 2 to make sense of all the different formulations. It could for example help to indicate upfront what expression you will derive in the various subsections and for what purpose (or alternatively, e.g. for 2.3.1 and 2.3.2, what aspects you will analyse and why), and it could also be useful to provide a summary or a table with the final expression which you will (mainly) use to analyse the data in section 3.

3. Eq. 9 and 10 show that, in the absence of baroclinicity, the veer only depends on the curvature of the stress profile. Later, it is shown that veer also depends on the shear (see Eq. 12, and 14-16). How is this possible? Is it perhaps so that the shear affects the veer by its impact on the stress

profile, so that Eq.9-10 are implicitly dependent on the shear? Please comment on the physics explaining these different forms in the paper.

4. The authors propose two practical equations to find veer based on shear, equation 38 and 39. Eq. 38 only accounts for the shear, while Eq. 39 includes the contribution from the cross-wind stress. In Fig. 15 it is shown that both equations give comparable results, and for simplicity equation 38 is recommended. However, this seems to contradict with the earlier finding from fig 5, showing that the contributions of shear and crosswind stress are an order of magnitude larger than the veer and mostly balance each other. How is it then possible to get good results with Eq. 18 in which the crosswind stress is neglected? Please comment on thing apparent contradiction in the paper.

Other scientific comments

1. Line 98: by definition $B = -z/L$. I believe this is incorrect and should be $B/\varepsilon_0 = -z/L$.

2. Eq. 8: It seems to me that you only use the first form of $\mathrm{d}\varphi$ for later derivation and interpretation. The derivation of the second and third form is not so straightforward, and, given that these forms are not being used (as far as I can tell), I think they make the math in this section unnecessary complicated. Moreover, I believe the second form (with S and S^*) is missing a minus sign (or equivalently, the factor i should be in the denominator instead of in the numerator). It is not clear to me how you derive the third form from the second one. Unless these complex forms are essential for later derivation or interpretation, consider removing them.

3. The wording used to introduce and derive equations in section 2.3 is not always clear. For example, line 143, what do you mean with "taking alternately the time derivative of (8)"? Line 152: what do you mean with the "dimensionless deviation of the wind from streamwise"? Line 168: "We note that (14) and (15) are more direct alternatives to dealing with functions of φ_G". Not clear to me what this is about. It is also not clear how Eq. 16 is obtained.

4. I'm confused about Eq. 17: The text says that β_{ma} is the angle between $\partial S/\partial z$ and $\langle sw\rangle$, but Eq. 17 uses $\arg(S)$. Are you implying that $\arg(S) = \arg(\partial S/\partial z)$? I don't see how this holds mathematically. Please explain.

5. Line 192: Please add a reference for the Rotta parametrization.

6. Eq. 18 is introduced to show the root of the misalignment between shear and stress, but it is not entirely clear to me from the surrounding text how this equation explains the root of the stress-shear misalignment. Line 197-199 only talks about the absence of misalignment when the flux-gradient relation is used, but a clear interpretation of Eq. 18 and how it explains

misalignment is missing. Are you saying that misalignment is directly related to the turbulent transport of turbulent stress (the last term in Eq. 18)?

7. Section 2.3.2: I found the section title misleading. The section does not talk about alignment of shear and stress, but instead discusses two canonical solutions for the veer and the relation with shear.

8. Eq. 21: How is this equation obtained? Do you get it from Eq. 10? Moreover, on line 222 you mention it "defies analytical solution", but I'm not sure what you mean with that. The differential equation has two unknowns $\varphi(z)$ and $U(z)$ so you need an additional equation anyway to solve it.

9. Line 243: It would be more clear to indicate in the subsection title that this is the Ellison regime you referred to earlier. The current subsection title "Linear diffusivity profile: ..." is not clear and I had to search for where this Ellison regime is actually discussed.

10. Fig. 1 left plot is not entirely clear to me. Which part of the curve is for low/high values of z? Line 260 says that you can see that "the Ekman solution produces less mixing away from the surface ... and consequently a higher shear exponent than Ellison's." Is this statement referring to the left plot? Difficult to appreciate without knowing where low/high values of z are in the plot.

11. Where is Eq. 35 coming from? Please provide a reference.

12. Line 375: How did you find the exponent -1.4? Can this be derived mathematically from Eq. 33 or is this an empirical estimate based on visual observation?

13. Line 429-430: "... follow a curve which resembles the nondimensional M-O shear function ...". Not all readers will know the shape of the Φ_M curve by hard. Please show the curve or provide a reference.

14. Line 450: Not clear what the angular brackets in $\langle\alpha\rangle$ mean.

15. Line 487: "The lowest [veer scale] corresponds to the offshore Høvsøre case, while the highest ...". It is not particularly clear how this can be seen. The veer scale defines the slope of the PDF in log scale, correct? As in, the higher the veer scale, the lower the slope?

Minor/technical comments

1. Variable Ro_0 is first used on line 264 but only defined on line 290. Please define upon first use.

2. Line 274: Should $|S/G|$ be $|S|/|G|$?

3. Line 307: "along with ~~with~~ (30) or (32)..."

4. Line 374: "... that is not inconsistent with" ⇒ Avoid double negation and just use "consistent with"

5. Probability density function values are sometimes difficult to interpret. For example, in figure 7 (right), a value of $P(\alpha, L^{-1}) = 1000$ is not very meaningful unless you know the bin size $\mathrm{d}\alpha$ and $\mathrm{d}L^{-1}$. It would be much clearer if you could show results in terms of their relative frequency of occurrence.

6. Caption of Figure 15: what are the two empty brackets "()"? Should these contain a reference to an equation?

7. Line 604: "condtions" ⇒ "conditions"

8. Line 667: "convering" ⇒ "covering"?

9. Line 703: "underpreictions" ⇒ "underpredictions"

10. Line 716: "stabiilty" ⇒ "stability"