Response to Mark Stoelinga

Referee comments appear in black and author responses appear in blue.

We thank the reviewer for his thoughtful comments.

181-182: I think centered RMSE (cRMSE) is essentially the same as what I've heard and referred to as bias-corrected RMSE (or BCRMSE), in which you first calculate the mean model bias error, subtract it from all the model values, then calculate RMSE. And, I believe both are essentially equivalent to the standard deviation of the errors as well. All that is neither here nor there. However, I do think the sentence in lines 181-182 should be clarified, to say that "a value of 0 for cRMSE indicates that all values, ***after removal of the respective model or measured means***, lie on the 1:1 regression line".

Thank you for the suggestion. We have added your proposed clarification to this sentence as "A value of 0 for cRMSE indicates that all values, with model bias removed, lie on the 1:1 regression line".

189 (paragraph): Might be good to show model versus measured mean shear exponent, a metric that the wind industry uses extensively and is highly familiar with its typical range of values.

Thank you for the suggestion. We have added a new Figure 8 and a discussion of model versus measured wind shear exponent, finding that lidar-derived exponents are in good agreement with past evaluations in the mid-Atlantic and that WRF-derived exponents are underestimated.

325-326: There is an interesting result in Fig. 8 that you do not comment on, which is similar to behavior other have seen and commented on (including, I believe, one or more of you in previous work, and myself). What I'm referring to is the opposite effect of TKE amount in the near-project versus distant wake environment. Within and near the project, behavior is intuitive: higher TKE dissipates wakes and leads to smaller waked wind deficits. However, farther away, as evidenced by the distance northeastward of the first (0.5 m/s) contour, as well as the area of this contour reported in the text, it is actually slightly farther (and covers more area) with TKE than without it. In other words, at distance, higher TKE actually helps wakes, whereas near or within the project it hurts wakes. I saw the same behavior, and I'm certain you and others have commented on it previously. Do you have any new insights into this behavior?

This comment was clarified and retracted by the reviewer in a later comment posted in the online discussion.

Appendix E. The authors and I have had discussions in the past about the nature of the noise seen in difference fields (turbines minus no turbines wind speeds). I'm not opposed to the idea that they are purely numerical; I agree that is the most likely explanation. However, I still consider it possible that even the distant differences are perhaps partly physical rather than numerical. They tend to occur in an unstable boundary layer or in convective scenarios. These scenarios are characterized by small-scale, high-amplitude, chaotic structures (convective cells) whose initiation locations are

random and probably sensitive to even the smallest perturbations, which may include very subtle and fast-moving gravity waves or other disturbance triggered by the presence of the turbines. For the purpose of energy production, though, they are probably inconsequential because they tend to cancel each other out when averaged either spatially or temporally.

Apart from noise adjacent to the farms, we have observed noise also appearing far upwind of the turbines where the introduction of wind plants should make no discernible difference to the atmospheric state (tens of kilometers upwind of the induction zone, with little or no noise in the induction zone). Even if gravity waves were involved here, gravity wave deflection should maximize close to the wind plants before dissipating, making it more likely that these features are numerical, but we agree that numerical noise is worth looking into in future studies.