General Comments

I started by reading the abstract and introduction then jumping into section 5. Many of the suggestions I make are from the perspective of someone who hasn't carefully read the body of the paper. There are also some cases where multiple ideas were combined into a single sentence. I thought Section 5.2 in particular did a nice job of already incorporating many of the ideas we've covered in class. There was good flow from old ideas to new ideas within sentences, and it was easy to tell what was meant to be emphasized.

Specific Comments

Section 5.1 - Line 1

The initialism EB is used without elaboration twice in the first sentence. I know this is used throughout the paper, so someone reading through really should know what it means, but the discussion section might be a place where readers jump to. It might be worth considering writing it out again here.

Also, I'm guessing "EB as a benchmark" refers to a generic empirical Bayes model without any bells and whistles. It may be helpful to make this more explicit (particularly if my guess is wrong).

Section 5.1 - Line 8

I think this sentence could be rearranged to strengthen the emphasis and remove redundancy. Something like "Posterior Biasing outperformed the other two models when predicting peak height." This puts Posterior Biasing in the topic position and predicting peak height in the stress position.

Also, when I read this I wondered how the models were evaluated and found the description in Section 2.3. From the perspective of someone skimming for results and discussion this might be difficult to find. My first thought was to move this description later in the paper, but it looks like there are a couple other results sections interspersed through the paper. I'm not sure what the best solution is, but it may be helpful to create a consolidated results section. Another option might be to refer to the section during the discussion so that a less thorough reader can find what they skipped.

Section 5.1 - Line 13

I think this sentence could be rearranged to make the regional EB model the topic and stress the comparison with the benchmark model. Maybe you could do something like "The EB model with regional effects does worse or about the same as the benchmark EB model."

Section 5.1 - Line 14

This second sentence seems to contain two related ideas: first that the performance of regional EB should be better than regular EB due to model flexibility, and second that the actual implementation used approximations that hurt its performance. If this sentence were split in two then you could have two stress positions to emphasize both observations.

Section 5.1 - Line 17

This sentence also seems to contain two ideas: first that we have this other model, and second that it does well. It might be helpful to split this into a sentence mentioning the targeting model with a brief description of what is different about it plus a second sentence describing its performance in cross validation. This sentence also mentions "the height". In the interest of consistency it might be good to keep referring to "peak height".

Section 5.1 - Line 19

The sentence says "although [the model] does outperform [PB], there is no distinguishable

difference..." I find this confusing and in need of clarification. Is this meant to be saying the error is lower but within one standard deviation so it's maybe not meaningful? This interpretation doesn't seem to be reflected in the table.

Section 5.2 - Line 7

This may be an unreasonable pet peeve of mine, but I would prefer "half as big" to "twice as small".

Section 5.2 - Line 11

This sentence seems to contain a few ideas that could be separated. First, only one tuning parameter choice was tried and it may not be optimal. Second, an optimal parameter choice could be found with cross validation. Third, it would take time running code to do this. The third idea could possibly be removed, unless the amount of time picking a different parameter would require makes it infeasible.

Section 5.2 - Line 25

There seem to be two ideas in this sentence: first that not all metrics are equally important, and second that the domain scientists know which metrics actually matter. I would consider rephrasing this to something like "While there are many ways of assessing predictions, some provide more useful insights. It is important to rely upon the advice of epidemiologists [etc.] when designing assessments.

This part also seems to be pointing toward an area of future work. I think this could be made more explicit.