This paper investigated the limitations of existing power outage models, including bounded prediction, out-of-distribution prediction, and physics-aware uncertainties. The authors found some of the existing state-of-the-art models may generate unrealistic predictions, and cannot generalize well to extreme events that are not sufficiently represented in the training datasets. The authors discuss some potential ways to address the shortcomings of these models. I have some major comments that authors need to address before publication:

1. The problems mentioned by the authors, including limited generalization ability, unbounded predictions, and unreasonable uncertainty variations, are common problem for general machine learning models. Many machine learning community researchers proposed different methods to address these problems. How unique and critical are they for power outage predictions?

2. Now there is a variety of more complex power outage prediction models [1], are there any specific reasons for the authors to choose to evaluate traditional machine learning models? These traditional models are known to be less representative.

3. It is unclear to me why beta regression should perform well in general cases. I think it also has its own problems such as strict distribution assumption and does not address the representativeness issues which eventually cause the poor generalization problem. Could you provide any justifications and performance comparison regarding why Beta regression should be used?
