This paper deals with the application of machine learning modeling to ecological data and its comparison with more classical linear modeling. The topic is relevant and the article is generally clear with interesting results. I have however some points that should be addressed or amended to my opinion.

Major points:

- The authors do not justify the use of gradient boosted trees method. There are other techniques that could be applied to such a dataset and the reader should be shortly informed about the advantages and drawbacks of gradient boosted trees with respect to other major methods.
- It is not clear why the linear model could not take into account tick density data, which seems to constitute an asset of machine learning model. Even though this is not the subject of the paper, there are examples of zero inflated generalized linear models (with Poisson distribution : see Bah et al., 2022, DOI: 10.1111/tbed.14578, but negative binomial could also be considered) which consider both occurrence and abundance. Data with no occurrence could also be used by applying a log(y+1) transformation. Therefore, these kinds of linear model may lower the advantage of using machine learning techniques and this should be acknowledged
- My main concern lies in the use of machine learning methods to help interpreting ecological interactions. It appears straightforward that machine learning significantly improves the predictive capacity of models and this is shown by the present paper. I am not convinced by the use of GBM to investigate the influence of environmental features with the present study. An increase in tick density with deer density, with a kind of linear-plateau relationship, seems relevant. But the variations of this influence for intermediate deer harvest does not appear to be based on biological grounds. The same applies to the influence of the temperature in June of the year before which have not been accounted for in other ecological studies on ticks, to my knowledge. These variables also arise from the linear model, which gathers many variables to my knowledge. In any cases I do not see the advantage of using GBM in this context. The authors acknowledge that GBM could point out particular issues to be addressed in detail, in the discussion (as linear models could also do) but they should be more cautious throughout the text.
- How does the quality of fitting vary with the maximal number of environmental features involved ? Why have you fixed this number at a value of 30 ? If you lower this number, does it have an influence on the fitting ?

Other points:

- P6 L21-22: the difference between RMSE and R2 for both models is quite low. I do not think that we could say that the density model is outperforming the abundance linear model.
- In the discussion, I don’t think it is worth getting into the detail of model results (P8 L21-27).