

Hello Joseph,

You explained the issue of omitted variable bias in your rain intensity model very clearly. It is a great example because weather models, just like the fraud models I have worked on before, often have hidden factors that influence the outcome. I liked how you logically connected how solar radiation could serve as a proxy for cloud cover, which would, in turn, affect both air temperature and humidity.

It made me reflect on how, during one of my projects, we initially left out some important factors like account age and transaction history, which caused our model to give us misleading results. Only after including them did we realize that the relationships between the variables became more realistic and accurate. This is very similar to how you described the potential changes in your model after including solar radiation (Marginal Revolution University, n.d.).

I also liked how you mentioned that the omission might have caused an overestimation of the effects of air temperature and humidity. It is very interesting to see how a single variable like solar radiation could adjust the model and provide a more accurate prediction. Have you considered how you might collect reliable solar radiation data for Chicago? I would like to know if historical solar radiation data is available publicly for urban areas like Chicago, which could be helpful if you plan to improve the model further (Nitka & Burnecki, 2019).

Thank you for sharing such a well-thought-out post.

All the best,

Avinash

References

- Marginal Revolution University. (n.d.). *Omitted variable bias*. MRU. Retrieved April 3, 2025, from <https://mru.org/dictionary-economics/omitted-variable-economics>
- Nitka, W., & Burnecki, K. (2019). Impact of solar activity on precipitation in the United States. *Physica A: Statistical Mechanics and its Applications*, 527, 121387. <https://doi.org/10.1016/j.physa.2019.121387>