

Interactive comment on “Developing a novel hybrid model for the estimation of surface 8-h ozone (O₃) across the remote Tibetan Plateau during 2005–2018” by Rui Li et al.

Anonymous Referee #1

Received and published: 12 February 2020

The subject is appropriate to Atmospheric Chemistry and Physics. This paper developed a novel hybrid model named random forest-generalized additive model (RF-GAM) to estimate the surface 8-h O₃ levels across the remote Tibetan Plateau during 2005–2018. This model displayed excellent prediction performances of O₃ spatiotemporal variations when compared with other seven machine learning models, and can be applied in the remote regions with sparse monitoring sites. This study is a valuable work. Therefore, I recommend clearly the acceptance for publication of this manuscript after revisions. Several editorial comments for improving the information content and presentation of the paper are listed as follows:

C1

1. When you analyzed spatiotemporal variations of simulated 8-h O₃ concentrations and nonattainment days at the monitoring sites, why do you not use observed values for comparison? It may verify the results and rich your discussions.
2. Since Community Multiscale Air Quality (CMAQ) model and other three-dimensional models have been widely applied to the simulation of surface O₃ concentrations, you'd better add a paragraph to summarize these results and compare your results with others to reveal the superiority of your novel model by adding more references such as Eder, B., and S. Yu, 2006. A performance evaluation of the 2004 release of Models-3 CMAQ. *Atmospheric Environment*, 40: 4811–4824.
3. As for the temporal variations of the simulated 8-h O₃ concentrations and nonattainment days as depicted in Fig. 7 and Fig. 10, the change pattern from 2005 to 2018 is not very obvious. Would you mind showing interannual variations in a clearer way?
4. L225–227 Regarding the other five models (i.e., RF, generalized regression neural network (GRNN), backward propagation neural network (BPNN), Elman neural network (ElmanNN), and extreme learning machine (ELM)), you should put all related results from them into SI because you already prove that RF-GAM model is the best and it is unnecessary to present the results of other worse ones in the main text part.
5. There are many English grammar errors in the manuscript. I only list some of them here for your reference below (please make all necessary corrections before publication).
6. L16: It should be “to predict the surface 8-h O₃ concentrations. . . .”
7. L21: It should be “to predict the surface 8-h O₃ concentrations. . . .”
8. L25 It should be “the estimated O₃ mean concentrations. . . .”
9. L44: It should be “The chemical reactions between NO_x and VOCs in the presence of sunlight were. . .”

C2

10. L85: It should be “the daily satellite data enable the fine-scale estimations of O₃ level. . .”
11. L90-91: It should be “the global distributions of O₃ concentrations based on simple linear . . .”
12. L103-104: It should be “can obtain the contributions of each predictor to air pollutants. . .”
13. L146-147: It should be “simulate the gridded 8-h O₃ concentrations over Tibetan Plateau for the first time. . .”.
14. L162: It should be “lack of anthropogenic activities and most of the residents focus on southeast and south parts of. . .”
15. L171: It should be “The data quality of all the monitoring sites was assured on..”
16. L175: It should be “The O₃ column amounts (DU) during 2005-2018 were downloaded from the Ozone Monitoring. . .”.
17. L294-295: It should be “we also investigated the spatial variabilities of the predictive accuracy for RF-GAM model. . .”.
18. L337-337 about the effects of RH on the O₃ formation: Please add the following reference to have more discussions in terms of chemical mechanism: S. Yu, 2019. Fog geoengineering to abate local ozone pollution at ground-level by enhancing air moisture. *Environ Chem Lett*, 17(1), 565–580, doi: 10.1007/s10311-018-0809-5.
19. L353-354: It should be “It was well known that photochemical reactions of BVOCs and NO_x in the presence of sunlight were beneficial to the O₃ formation..”. Also please add more references about this such as: S. Yu, et al., 2006. Performance and diagnostic evaluations of a real-time ozone forecast by the Eta-CMAQ model suite during the 2002 New England Air Quality Study (NEAQS). *Journal of the Air & Waste Management Association*, 56:1459-1471.

C3

20. L397-398: It should be “Besides, the 8-h O₃ concentrations in Tibetan Plateau displayed significantly seasonal discrepancy. . .”.
21. P44: It should be “Table” in stead of “Tab.”.
22. P2, L34: “shared with” => “shared”
23. P3, L62: “surfacce” => “surface”
24. P3, L63: It should be “. . .concentration displayed the gradual decrease in the recent ten years”.
25. P4, L69: “lacks of” => “lacks”
26. P8, L163: “is consisted of” => “consists of”
27. P20, L422, 434: “Tab. 2” should be “Table 4”

Interactive comment on *Atmos. Chem. Phys. Discuss.*, <https://doi.org/10.5194/acp-2019-972>, 2020.