This manuscript presents an 11-year field experiment in an organic olive grove combined with RothC simulations to evaluate soil organic carbon (SOC) dynamics under different organic amendments (compost, biochar, and their mixture). This manuscript involves a considerable amount of work on field experiments and model simulations. The availability of long-term experimental data is a notable strength and provides valuable data for model validation, yet these uncertainties in model validation and model prediction accuracy are not adequately discussed. I have the following suggestions for improving the research.

Major

1. This manuscript compares the model-simulated SOC with observed values, and a key result of the study is the poor agreement between modelled and measured SOC, particularly for the biochar treatment (e.g., R^2=0.10, high RMSE). However, the manuscript does not adequately explain the reasons for this discrepancy. More discussion is needed to explore potential sources of inconsistency, such as model structural limitations, the absence of potential mechanism processes (e.g., priming effects), and parameterization choices.

2. Has the model been sufficiently validated with consideration of soil heterogeneity? What is the scalability of the single-point validation? That is, can the model be reasonably applied to other regions? It is recommended that the authors further validate the RothC model using literature data or by testing alternative validation data from other soil carbon models (e.g., MIMICS-BC, EPIC, APSIM). At the very least, this issue should be discussed.

   References:

   Han M, Zhao Q, Wang X, et al. Modeling biochar effects on soil organic carbon on croplands in a microbial decomposition model (MIMICS-BC_v1.0)[J]. Geoscientific Model Development, 2024, 17(12): 4871-4890.

   Lychuk T E, Izaurralde R C, Hill R L, et al. Biochar as a global change adaptation: predicting biochar impacts on crop productivity and soil quality for a tropical soil with the Environmental Policy Integrated Climate (EPIC) model[J]. Mitigation and Adaptation Strategies for Global Change, 2015, 20(8): 1437-1458.

   Archontoulis S V, Huber I, Miguez F E, et al. A model for mechanistic and system assessments of biochar effects on soils and crops and trade‐offs[J]. Gcb Bioenergy, 2016, 8(6): 1028-1045.

3. Given that the compost contains sheep manure, it is likely a source of nutrients. How were these nutrients accounted for in the model? Was the nutrient content of the added compost measured experimentally? Because the nutrient may stimulate soil microbial

activity, thereby influencing soil carbon decomposition. This is a critical issue that the manuscript does not address. For example, on L175, the mixture of compost and biochar is treated as independent components without considering interactions. However, nutrients from compost could stimulate microbial activity and potentially affect biochar degradation. This will introduce uncertainty that should at least be discussed.

4. The irrigation was conducted in the field experiment. However, how was this considered in the model? Irrigation practice can affect soil moisture, which further influences soil carbon decomposition (e.g., soil moisture process in Millennial, MEND model).

References:

Abramoff R Z, Guenet B, Zhang H, et al. Improved global-scale predictions of soil carbon stocks with Millennial Version 2[J]. Soil Biology and Biochemistry, 2022, 164: 108466.

Wang G, Huang W, Zhou G, et al. Modeling the processes of soil moisture in regulating microbial and carbon-nitrogen cycling[J]. Journal of Hydrology, 2020, 585: 124777.

Minor

5. Line 50: In addition to these first-order kinetic soil carbon models, there are also microbial models, such as MIMICS and MEND, which might offer greater utility for understanding the impact of biochar addition on soil carbon dynamics (e.g., microbial process, priming effect). The authors should provide a more comprehensive overview of existing modeling approaches and discuss how their model choice compares with these alternatives.

6. Line 128-129: The determination method for SOC concentration should refer to the corresponding references.

7. Line 134: Why was the MAD threshold set to 4, and what percentage of the data was filtered out as a result?

8. Line 153: Is there evidence to support the assumption that complete mixing occurs within 5 months? Several methodological choices are not sufficiently justified and may bias the results.

9. Line 155-159: The current description of the RothC model is too general. At least a structural diagram of the RothC model should be provided, illustrating the carbon pools and carbon flows, to facilitate a more intuitive understanding of the simulation processes for different amendment additions. In addition, the model time step of the RothC simulations is not explicitly specified.

10. Line 188-200: Compost parameters are derived via model inversion, but biochar parameters are adopted from the literature without calibration. This inconsistency may limit the comparability of treatments and should be explicitly acknowledged and discussed.

11. Line 189: More details or a summary of the inversion procedure should be provided in the main text. Currently, the description is too dependent on supplementary material.

12. Section 2.3.3: The model assumes constant plant carbon inputs derived from equilibrium runs. However, the field system is a perennial olive grove, where biomass accumulation is likely to increase over time. The authors should justify this assumption, discuss its implications, and consider uncertainty tests or sensitivity tests on the main model parameters to strengthen the credibility of the findings.

13. Section 2.3.3: The input variables information (e.g., source, spatial and temporal resolution) of the model, including those used for long-term equilibrium and forward simulations, should be described in detail. A summary table would greatly improve clarity and reproducibility.

14. Line 219: Performance metrics (R^2, RMSE, Bias) should be explicitly defined at first mention, including their calculation, unit, and interpretation.

15. Line 241: The note “Bars represent standard deviation” is suggested to be added to the end of the Figure 2 caption.

16. Fig. 4: Statistical significance should be added.

17. Line 377-380: The authors should explain why compost treatments are more predictable than biochar treatments. One reason is likely due to the lack of mechanisms to capture biochar-induced changes in native soil carbon dynamics. These uncertainties should be discussed.

18. Line 454-455: The authors mention that a two-pool decay model was tested, which seems that the test was conducted by the authors, yet no supporting results or references are provided. Please clarify this point.