Methane is one of the most important greenhouse gases and its concentration increase in the atmosphere is of large concern to the society due to methane's strong warming potential. Previous reports showed that methane emissions from reservoirs are an important natural methane source. However, few models can capture the dynamics of methane emissions from reservoirs. In this study, Lomov et al. describe a process-based lake model that includes representation of processes that are important for reservoir methane dynamics and present a new set of valuable data in Mozhaysk reservoir, Russia to validate the model. I find that this study is interesting. However, its scientific and presentation quality have not reached the standard of this journal.

First, I agree with the first reviewer that the manuscript looks like a rough draft rather than one ready for publication. There are many grammar errors and typos. Sometimes, the writing is difficult to follow. Thus, a thorough proofreading is needed.

Second, the model description is in a poor format. For example, the manuscript lacks an overall description of each term in the right-hand side of Eq. 2 for their physical meanings. In addition, the description of each term should be better organized. The whole section 2.4 in the current form is very difficult to follow. Also, as the model development puts methane dynamics in the central focus, a detailed description of methane biogeochemical processes is necessary. For example, the authors need to provide details on the parameterization of methane production, oxidation and transport. Importantly, the authors should only introduce the processes that actually have been turned on in the numerical experiments. For example, internal wave induced diffusivity and detritus sedimentation are described in section 2.4 but were not used in the simulations.

Third, the upscaling of methane flux observations from station IV to the whole reservoir surface is very arbitrary. As shown in Fig. 2b, the statistical relationship between riverbed
methane emission at station IV and that from floodplain is not significant. As such, using this probably erroneous total methane emission estimates to validate the model does not make sense to me. Instead, the authors should only use the observational data for model validation. For example, the authors can compare the modeled and measured methane emissions for each station. In my view, although the model is presented in an longitudinal integrated form, the methane emissions over the riverbed portion of each station can be easily isolated. In the end, once the model is validated, the authors can use it to estimate the total methane emission for the reservoir.

Forth, although it is fine that the model fails to reproduce observations in several cases, it is important to dive deeply into these model errors. In my view, the clear explanation of the successes and failures of the model validation is the main value of a model development study. It will inspire other model developers to improve the representation of related processes and help interested users choose appropriate models for their applications. However, most of the analyses in section 3.2 are only qualitative and much speculative. For example, how does the error of simulated bottom water temperature shown in Fig. 6 affect the simulated oxygen and methane dynamics? Are methane production and oxidation rates correctly simulated? From the large bias in simulated bottom methane concentrations, it seems to be not the case. Also, from Fig. 12, I cannot tell what the model does right and what the model does wrong. I am also not confident that the model get the right result for the right reason.

**Specific comments**

L44-45: This sentence and Figure 1 give me an impression that the introduced model will represent methane flux through macrophytes. But I cannot find it in the manuscript.

Table 1: What do mean and maximum width refer to? From Figure 2, I cannot see the width of the reservoir is so small.

L143: replace "was" by "can be"

Table 2: Could you use a more standard date format, such as MM/DD

L149: were evaluated

L219: from all the descriptions, I still do not know how the horizontal heterogeneity of some variables, such as oxygen and methane, are represented in the model but the authors emphasized that the longitudinal difference is substantial.
L239: replace "inequality" by "imbalance"

L244-247: please specify this background diffusion. How was it used in the simulations?

L251-253: This is too vague. What does this momentum flux represent?

Figure 4: Could you also show observed diffusion and ebullition?

L321: thermal stratification

L322-323: Do you have simulation results to support this claim?

Figure 5: The use of R value is misleading for statistical significance. Please use R2 instead.

Tables 5-7: Please use bias (difference in mean), root mean squared error (RMSE) and NSE for model performance metrics. Pearson R is not a good metric for model validation.

L423-424: But the error of simulated oxygen in the middle layers, such as metalimnion, would be important for methane dynamics. Could you show the comparison?

Figure 9: Where are the surface oxygen comparison for 2019 and bottom oxygen comparison for 2016 and 2017?

L481: Do you represent the emission of ice-trapped bubbles in the model? If not, I cannot understand why high ebullition occurred in early spring.

Figure 11. As you speculated that the drop of hydrostatic pressure is responsible for many of these high ebullition, could you present its dynamics in the figure?

Figure 12. Could you also compare the model with observations for diffusion and ebullition?