General comments

Lake ice phenology is a sensitive indictor to climate change.

The paper proposed a new double-threshold moving t test (DMTT) algorithm applied to CETB dataset, and reconstructed the ice phenology of Lake Ulansu from 1941 to 2023 using Random Forest based on the ERA5 data, which have been proved relatively good performance. The method is novel in identifying lake ice phenology and the influencing factors, but there still need some improvement.

(1) The introduction is not well organized, and need to improve.

(2) The logic of methodology is not clear. The overview of Data makes the reader confusing.

(3) The caption of figures should be concise, and the author need to check and improve them.

(4) The correlation between solar radiation and lake ice phenology is lower than we expected. The authors analyzed this in terms of month. Freeze-up and break-up date are one value in a winter, and other parameters varies each month, how to connect them together?

In conclusion, the subject of the study is suitable for "The Cryosphere" and can potentially be of interest for the journal's wide audience, and we suggest a major revision.

Specific comments

The Introduction is not clear and attractive. The meaning of long-term tracking of climate records such as lake ice phenology, is not well expressed. The research meanings in the first paragraph of Introduction are not closely related with the content, and the author should rewrite. The scientific problem and study goals are note clear.

Line 38 the best quality is not appropriate.

Line 41 MODIS just mentioned 8-day product, and the author should update recent work.

Line 46 The active microwave is not suitable for the large lakes. This sentence should be improved, and supplement more contents.

P47 One advantage of passive microwave is frequent revisit, which is not mentioned.

The second paragraph is too long, and should be divided. The core of this work belongs to passive microwave, and could be discussed alone.

The fitting methods of lake ice phenology based on climate records (such as air temperature) are note well discussed, just listed some previous works.

The author explained how the study organized, and the goals of this study is not clear. The last paragraph should be rewritten.

The general description of ice regime should be added in Study are, like ice thickness, regular winter recreation.

Line 75 300 need to check, or provide the citation.

Line 80 how to exclude the influence of vegetation?

Line 83-84 remove the date.

Line 85 The definition of hydrological year is not accurate. For example, HY 2022 lasted from August in 2021 to July in 2022.

Line 90 The title of figure 1need to revise.

Line 95-105 The logic of flowchart is not clear, and Figure 2 should be revised. The input and output is not clear. The different colors have certain meanings? The method is too short.

Line 119 high spatial resolution Tb data are essential, gramma error
Line 124 These data were sourced from the SMMR on the Nimbus 7 satellite; the SSM/I on the F08, F10, F11, F13, F14 and F16 satellites; and the SSMIS on the F16 satellite. Move this to the above content.
Figure 3 have two choices: (1) make the content short and concise, and add the figure in the supplementary materials; (2) move this part to Results.
Figure 4 “between 1 August and 31 July from 1979 to 2023 for Lake Ulansu”: check this.
“The solid lines represent the interannual average brightness temperature and air temperature time series, respectively. “ delete this sentence. Lack the description of Figure 1 (a), just (b) appears.
“
Line 174 “proportion lake water greater than 0.70” how the thresholds are determined? Why do not use 0.8/0.2?
Line 236 “Seventy percent” and “30 %”, keep the same expression.

Line 258 ERA5 have various types of climate data, which one you used? It is not clear. The input and output of RF is confusing.

Figure 6 Please add the math equation and basic index of linear regression. Please add the lengend of different lines, it would be better to remover the description of dashed lines. (a)-(f) explained separately.
Line 280 add the yearly changing rates.
Line 282 The title of 4.2 is not proper.
Table 3 Linear tread is yearly changing rates? Please add the unit. The linear tread of wind speed is 0. Check this?

Figure 8 The correlation between solar radiation and lake ice phenology is lower than we expected.

As for random forest, we have following questions:
(1) The abbreviation of correlation coefficients is r, not R. Need to check the whole manuscript.
(2) The evaluate the performance of RF, the determination coefficients R^2 is more usually used, rather than R.
(3) How the author avoid overfitting in the work? Need to explained more clearly.
(4) Why the author chose random forest rather than other methods?

The two paragraphs in Author contribution are repeatedly expressed.