This manuscript presents thermodynamic simulations to constrain the extent of submarine permafrost along the U.S. Beaufort Shelf near Oliktok Point. The study explores 18 model configurations that vary lithology, pore-water salinity, geothermal heat flux, and simplified borehole-based heterogeneity, and then compares the resulting ice-saturation patterns with previous seismic, resistivity, and DTS-based interpretations. The topic is potentially valuable, especially given the growing interest in combining process-based modeling with emerging fiber-optic observations in Arctic shelf environments.

However, in its current form, I do not find that the manuscript supports several of its strongest claims. My main concern is that the study is framed in places as if it had robustly constrained the actual present-day conditions along the transect. In contrast, the evidence presented is more consistent with a scenario-bounding or sensitivity-analysis study. The manuscript shows that some parameter combinations appear more consistent with previous geophysical interpretations than others, but it does not yet provide a sufficiently direct observational evaluation to justify the stronger inferences currently made in the Abstract, Discussion, and Conclusions. In my view, the study has clear potential, but substantial revision is needed before it is ready for publication.

Major comments:

1. In its current form, the study is best understood as a set of thermodynamic sensitivity experiments exploring how simulated submarine permafrost extent changes under different assumptions about lithology, pore-water salinity, geothermal heat flux, and simplified subsurface heterogeneity. That is a useful contribution. However, the manuscript often frames the results more strongly, as if the study had robustly constrained the actual site conditions. I suggest that the authors define the contribution more clearly and consistently from the outset. This is primarily a scenario-bounding or sensitivity-analysis study, and the framing of the abstract, discussion, and conclusions should match that scope.

2. A central weakness of the manuscript is that the simulations are not evaluated against observations in a way that clearly demonstrates model capability for this specific site. The discussion refers to seismic, resistivity, and DTS-related information, but these are used mainly for indirect or qualitative comparison rather than for clear calibration or validation. As a result, the study does not yet show how well the model reproduces the observed system; rather, it shows that some tested scenarios appear more plausible than others. This distinction should be made much more clearly throughout the manuscript.

3. The borehole-based heterogeneous simulations are presented as more realistic, but they remain highly idealized. In particular, the extension of lithologic units laterally across the shelf from a small number of boreholes is a major assumption, not a minor simplification. Likewise, the treatment of bottom-water temperature, the use of only two geothermal heat-flux values, and the adoption of a fixed 40% ice-saturation threshold all strongly influence the simulated distribution and inferred extent of submarine permafrost. These assumptions are acceptable as part of a conceptual sensitivity study, but they currently carry too much weight in the interpretation. The manuscript should either test their influence more explicitly or discuss their consequences more critically.

4. I encourage the authors to moderate several of the stronger inferences. In particular, the suggestion that the preferred scenarios imply pore-water salinity near seawater values and higher geothermal heat flux, the interpretation of deeper high-resistivity signals as hydrocarbons rather than deeper permafrost, and the suggestion that localized DTS

anomalies may indicate methane seepage all seem more speculative than the manuscript currently acknowledges. These are interesting hypotheses, but they should be presented as such. At present, the wording in several places makes them sound more firmly established than the results justify.

Minor comments

1. Please reconsider the use of "eighteen models." This is likely to be interpreted by many readers as eighteen distinct models, whereas this is one simulator/framework explored across eighteen configurations.

2. There are several language issues and small editorial problems that should be corrected in revision. Examples include:

"ResFrac Coporation" in the author affiliation.

Line 199: "These were chosed".

Line 392: "∼450 m.." contains a double period.

Line 413: "? - a companion study...".

3. The data availability statement still includes a "doi:TBD," and several cited outputs are listed as "in prep." This may be acceptable at the discussion stage, but the authors should ensure these are updated before final publication.