Referee comment on "Estimating spatiotemporally continuous snow water equivalent from intermittent satellite observations: an evaluation using synthetic data" by Xiaoyu Ma et al., EGUsphere, https://doi.org/10.5194/egusphere-2022-470-RC2, 2022

The manuscript ‘Estimating spatiotemporally continuous snow water equivalent from intermittent satellite track observations using machine learning methods’ by Ma et al. provides an interesting synthetic study regarding SWE observations from potential satellite tracks for track-to-area transformations. In general, I think such studies are important to be prepared for novel satellite missions and their ‘real world’ applications. The study is carried out in the Upper Toulomne River Basin in California in the Western United States. The authors apply statistic and machine learning techniques to answer the following four research questions:

(1) How does the spatially distributed April 1st SWE inferred from TTA compare with the synthetic truth, and how do their differences vary in dry, normal, and wet years?

(2) What are the dominant variables for the April 1st SWE estimation in statistical and machine learning TTA methods, and which method has the highest accuracy?

(3) How does the accuracy of the domain-wide SWE estimates from TTA approaches evolve within a season at different temporal observation resolution?

(4) How does the performance of TTA change as a function of the spatial sampling density (number of hypothetical ground tracks), and what is the preferred number of tracks?

The manuscript is well written and organized. I agree with all points raised by reviewer 1. In addition, I have some further points, which need clarification and/or additional information before publication.
Specific comment:

Please change the unit of SWE in mm (instead of m or cm) throughout the paper (text, figures, tables), as mm is the unit for SWE.

In general, please discuss if the pixel size and satellite track width of approx. 1 km is suitable to accurately describe the quantity of snow and SWE accumulation and ablation within the catchment (and, also for which applications is this resolution sufficient?). Could there be any limitations and need for higher spatial resolutions?

Title: The title is a bit misleading. The study is fully synthetic, but the title ‘promises’ somehow ‘real’ satellite track observations. I would recommend to at least include the word ‘synthetic’ (or ‘hypothetic’) in the title. Moreover, as you use statistic and machine learning methods it would make sense to add also statistic in the title. Suggestion: ‘Estimating spatiotemporally continuous snow water equivalent from intermittent synthetic satellite track observations using statistic and machine learning methods’.

Abstract: The abstract should be revised: Please add information on the potential future satellite mission (P-SoOP), on which you relate your work on TTA, and give information on the spatial resolution of your synthetic satellite tracks.

38+: I am missing a solid literature review
a) on satellite based SWE and snow height derivation. This should at least include the following references:


b) on other TTA or point-based methods.

57: Please add ‘wet snow’ as further limitation.

82: Please give some more information on P-SoOP in the manuscript, including when it is planned to be launched and if the track width matches your synthetic assumption. What is the expected accuracy of P-band based SWE estimates?

103: Why not also areas in high latitudes?

122: The F2022 snow reanalysis dataset should be described in more detail. What is the meteorological input to generate SWE? Are the applied meteorological variables for the
F2022 snow reanalysis dataset the same than those you used for your machine learning approaches (is independency given)?

195: Please define how you classified the years in dry, normal and extremely wet. What are the thresholds?

338 and Figure 3: As you clearly mention, DNN performs best. However, it is worth to mention that RF shows the lowest PEBAS values for the normal and wet years. Also, more clear statements why DNN performs best would help.

565: Please add some more discussion on the fact that topography plays an important role regarding the ‘choice’ of the satellite tracks. Does this play a role in Figure 11 as an increase in the number of ground tracks shows outliers in the course of MAE for WY2017 – 3 ground tracks and WY2008 – 5 ground tracks.