Comment on egusphere-2022-833
Anonymous Referee #2

Dear Authors,

The paper presents a framework to estimate transmission losses from satellite imagery and predict their hourly time series using random forest regressors. The study is interesting and shows novelty, but it must be improved in many aspects in my point of view. Moreover, there are some critical points that should be carefully addressed by the Authors. Please, find below my comments and suggestions.

Major comments:

- Page 2, Line 42-44: “Furthermore, as noted by Cook (2015), it is unusual for two gauging stations to be located on the same river without intervening tributaries. Therefore, quantifying transmission losses from two existing gauging stations is rarely possible.” I disagree. In reality, for large dryland rivers, where the most studies on channel transmission losses were undertaken, upstream river discharge is much larger than runoff produced between the streamgauges. Considering allogeneic dryland rivers, this runoff is practically null. Therefore, quantifying transmission losses from two or more existing gauging stations is perfectly possible in drylands.
- The described perceptual model of the surface-groundwater interaction of the study site (2 Study site) must be much better spatially presented, showing profiles along and across the main river and aquifer units.
- As you described in the study site, is it correct to say that the water lost in the ephemeral losing reach is immediately available again downstream?
- The ephemeral reach is always a losing one? Even during high floods?
- Page 6, Line 137: Why did you use five inflection points for the rating curve?
- 3.2.1 Transmission losses derived from the river drying front locations: the main problem of this methodology is there are just five days of comparison between the GPS and satellite drying front positions, although daily satellite images were available. Therefore, more fieldwork should be done, in order to properly estimate the uncertainty
on the satellite wetted river length estimation.

- Page 7, Line 180: you wrote that “the higher uncertainties are typically associated with shallow and low flow in the smaller braids.” However, your fitted linear model showed a rather different result with higher uncertainties related to larger flows.
- Page 13, Line 245: “... and the effect of the peaks becomes an important control” How?
- 4.3 Reconstructed transmission loss time series: What did we learn about the transmission losses in the Selwyn River when the machine learning approach was applied? If there is nothing to add to our understanding of the process, I suggest either excluding it or to use another time series model.
- You should compare your study with previous studies conducted on other ephemeral streams, including those from other climates. It is fundamental to place your findings in the context of transmission loss research.

Minor comments:

- I suggest moving the Figures A1, B1 and C1 from Appendix to the main text.
- Page 7, Lines 190-192: “In the course of the model development, more predictors (e.g. river flow, water temperature, groundwater level, date) have been tested but they appeared to not improve significantly the predictions.” Have you tried any statistical criterion, such as AIC?
- Please, reconsider the terminology of “reconstructed” transmission losses, because reconstruction of time series is a quite different topic. You should use just “predicted” transmission losses.
- Page 12, Lines 235-236: “The estimated transmission losses range from 0.14 to 1.5m³/s/km. Most of the estimated losses (56%) are below 0.6m³/s/km and correspond mainly to baseflow periods and river drying phases.” Please, provide a box-plot of the transmission losses, and add more statistical details.
- Conclusions: It is not necessary to use citations in the conclusion.