Responses to Reviewer 1 (Dr. N. Eckert)

Thank you very much for the positive comments and constructive suggestions. We agreed to all the comments, and made changes in the manuscript accordingly. We would note that the line numbers correspond to the manuscript without comments/changes.

(1) Workflow: Because of the several modelling steps and datasets, following the workflow is not that easy. An additional figure would really help.
[Answer] We agreed and included the figure of the workflow in Appendix B.

(2) Spatial modelling: Maybe it is because I cannot fully figure out the workflow, but the way the explicit spatial modelling is done is not crystal clear to me. As far as I understand, three random fields are fitted on the data (presumably using max likelihood techniques), and the corresponding variograms are used as known covariance matrixes in the Bayesian modelling step. If this is true, the claim of a Bayesian geostatistical estimation method is not justified, or at least not fully. I feel I am right because the conditioning on tau appears nowhere in the detailed conditional distributions. Also, it is clear that inferring the variograms adds substantial difficulty to the Gibbs sampling. Yet, it remains feasible, see Lavigne et al. (2016) as an example. Anyhow, things should be written more clearly and the choice made justified.

[A] We found that Lavigne et al. (2016) was very interesting. Thank you very much for pointing it out. In this study, we assumed that the geostatistical model and parameters were well constrained by a large number of point measurements (snow depth probe data). The goal of our Bayesian estimation was used to include indirect information that could be used to estimate snow depths (such LiDAR DEM and GPR). Although we could infer the geostatistical parameters during the Bayesian estimation, such indirect datasets were known to be not powerful enough to improve the estimation of geostatistical parameters [Day-Lewis and Lane, 2004; Murakami et al., 2010]. In addition, we included the uncertainty and prior information for the correlation parameters between snow depths and GPR as well as the ones between snow depths and LiDAR DEM.

To clarify, we added this sentence (Line 390), "Although we may include the uncertainty of those geostatistical parameters in the Bayesian estimation (Diggle and Ribeiro, 2007; Lavigne et al., 2016), we assume that those parameters are fixed in this study. This is because we have a large amount of point measurements (snow depth probe data), and also it is known that indirect information (such as geophysics) does not significantly improve the estimation of geostatistical parameters (Day-Lewis, 2004; Murakami et al., 2010)."

Additional Reference:
Day-Lewis, F. D., and J. W. Lane Jr. (2004), Assessing the resolution-dependent utility of tomograms for geostatistics, Geophys. Res. Lett., 31, L07503, doi:10.1029/2004GL019617.

Murakami, H., Chen, X., Hahn, M. S., Liu, Y., Rockhold, M. L., Vermeul, V. R., Zachara, J. M., and Rubin, Y.: Bayesian approach for three-dimensional aquifer characterization at the Hanford 300 Area, Hydrol. Earth Syst. Sci., 14, 1989-2001, doi:10.5194/hess-14-1989-2010, 2010.

(3) Modelling assumptions: The model proposed is rather straightforward, and, in fine, seems to be sufficient in cross-validation. However as the paper is written, the modelling choices made seem to be a bit too much convenience choices for the Gibbs sampling rather than data-driven, letting the reader the feeling that alternative modelling structures could do even better. We indeed all love handling only Gaussian fields, but sometimes things are a bit more complicated. For instance, I would expect some skewness in the variables analysed by the authors. Also, there is no justification for the three chosen variogram models, etc. Even if no sensitivity study is attempted (the paper is already long), a few words should at least be added in discussion to justify the probabilistic choices made.

[A] We agree that there could be another modeling choice. However, we have checked whether the distribution of the residual snow depth (after removing the linear correlation to microtopography) was Gaussian. We included this sentence (Line 573): “We used the Shapiro-Wilk normality test to confirm that the residual of the linear correlation, defined by tau in Equation (1), follows a normal distribution (the p-value of rejecting this hypothesis was 0.21).”

In addition, to justify the choice of variogram models, we included (Line 354): “The covariance models (equivalent to variogram models) can be selected to minimize the weighted sum of squares during variogram fitting.”