Review of “AI-Driven TanDEM-X Penetration Bias Estimation in Antarctica Using ICESat-2 and ECMWF Data: Implications for the NASA Surface Topography and Vegetation Decadal Survey Incubation study” by Ankita Vashishtha et al.

Vashishtha et al. used a set of radar-derived parameters and a set of climate parameters as features in 14 non-linear regression models to predict the penetration bias of X-band microwave (defined as the difference between TanDEM-X heights and ICESat-2 heights). The radar-derived parameters and climate parameters are selected based on their data distributions and correlation coefficients between features. The most variable and independent parameters are adopted as features in the regression models. Based on the error metrics of the predicted penetration bias, the authors demonstrated the results of the RF and DNN models and their related sensitivity when the temporal difference between TanDEM-X and ICESat-2 acquisitions and the features in the regressors are altered. It was concluded from this analysis that incorporating the climate parameters into the regressors could improve the performance of the prediction, whereas the temporal difference between TanDEM-X and ICESat-2 acquisitions has a negligible impact. The authors also showed that InSAR coherence and amplitude are the most important features among the radar-derived parameters, while the 2 m air temperature, near-surface wind patterns and surface pressure are the most important features among the climate parameters that affect the regressors. Finally, implications for future missions and studies are proposed based on this study.

Overall, this study is insightful for the cryosphere community, as it provided a novel and interesting algorithm to derive high-resolution and potentially dynamic DEMs for Antarctica (perhaps also other snow-covered regions). The error metrics provided in the manuscript are also promising. The obtained X-band microwave penetration may also have the potential to assess subsurface snow/firn properties, which is interesting for cryosphere studies. However, my major concern is that the manuscript suffers from serious issues related to language, clarity, and structure. The current presentation lacks the level of professionalism expected in scientific writing, which makes it difficult to fully evaluate the contribution or impact of the work. A substantial rewrite is required to meet the standards expected for scientific publication. Please find the detailed comments below.

General comments

1. A lot of punctuation errors, subject-verb errors, errors in citation format, and informal languages are present throughout the manuscript (mainly up to Results section). The detailed examples will be provided in specific comments and technical corrections.

2. While I do not wish to discourage any innovative or fun ways of structuring a manuscript, I found the flow of the paper quite difficult to follow. In the Introduction section, the authors started by listing a wide range of satellite remote sensing techniques together with their limitations. Almost none of the listed techniques was used in this study. Suddenly, TanDEM-X was introduced, without any mention of the strengths and limitations. The regions of interest and data are proposed here, instead of conventionally in the Data section. An overview of the paper structure is proposed, followed by some literature reviews and (In)SAR principles as subsections...

Hereby I would like to suggest a preferred structure in the following order:
- Introduce the importance of accurately estimating (I would not say characterizing) Antarctic surface elevation and mass balance, as the authors already did.
- Introduce TanDEM-X, including what the mission is, what the mission does, and the strengths and limitations. The penetration bias and related previous studies can be introduced here as well.
- ICESat-2 data can be used to estimate penetration bias due to the theoretically negligible penetration into snow (Michel et al., 2014). However, the spatial gaps between the tracks, the cloud-cover impact, and the longer revisit time (compared to TanDEM-X) make it not ideal for constructing a high-resolution DEM solely based on ICESat-2 data. I understand that the authors already mentioned this limitation in the Discussion section. However, I would

recommend also introducing it here to clear the doubts such as “if the ICESat-2 data are already used as reference, why would the authors not use them to construct DEMs directly?”
- Introduce the opportunities brought by AI methods in estimating X-band penetration bias

In this way, Fig. 1 and Table 1 can also be moved to Study Area and Data section, in order to better match the titles of each section.

3. From Table 1, it is difficult to understand what the features and the target variable are. I would suggest adding a diagram or an explanation to clearly indicate what should be the input and output of the regression models.

4. I understand that it would be extra work and could be impossible to implement snow grain size in the regression models. However, I wonder what the authors think about the effect of snow grain size (in addition to density) on the estimation of penetration bias. As far as I understood, not only does snow grain size affect the volume scattering of radar (Tsang et al., 2022), but it can also introduce a scattering bias in ICESat-2 heights (Smith et al., 2018). I see that the effect of grain size is briefly mentioned in the Discussion section, but without a general discussion about what may affect the penetration bias, such a brief touch on the topic seems abrupt and unexplained.

5. I am curious about how the updated DEMs in the regions of interest compare with the existing CryoSat-2 DEM (Slater et al., 2018) and the Reference Elevation Model of Antarctica (REMA; Howat et al., 2019) in the corresponding regions. Would the difference between the DEMs indicate limitations in the proposed methodology, the CryoSat-2 methodology, or the stereoscopic-imagery-based methodology?

6. Line 183. “The typical winter season” is defined as the period between April and June. However, the Kuipers Munneke et al. (2018) study defined the winter season as the period between April and October. Could the authors motivate why they chose this period, and would it affect the result if the investigated time period is extended?

7. The authors put the assessment metrics of different regression models in the Supplement and only presented RF and DNN results in the main content, probably based on the best performance. Maybe I missed something, but the selection criteria and locations of the figures are not introduced very clearly. Table 5 is introduced in Line 370, but Fig. S1 only appeared in Line 374. This writing flow makes the readers wonder a bit why Table 5 only shows RF and DNN instead of the previously mentioned set of machine learning and deep learning models. I would suggest the authors add some comments about the result selection at the beginning of Section 4.1.

This is personal preference, but it might be clearer and simpler to move the figures from Supplement to Appendix.

Specific comments

1. Lines 19–28. This part of the Abstract seems somewhat repetitive, which dilutes the application of TanDEM-X and the importance of mitigating its penetration bias. I would suggest the authors change the flow into:

TanDEM-X is an InSAR mission operating in X-band (8–12 GHz frequency) -> it can be used to characterize snow and ice surfaces, analyze snow layers, and construct DEMs -> however, due to the absorption, reflection and transmission of X-band microwave in the snow layer, penetration biases are introduced and consequently cause the InSAR-derived DEMs to deviate from the true physical surface -> therefore, correcting X-band penetration biases is important...

I also wonder whether it is necessary at all to mention the characterization of snow and ice surfaces and the analysis of snow layers, because these applications are not mentioned afterwards anymore.

2. Lines 69–71. The listed features seem to cover both radar-derived parameters and topographic parameters. I would like to see the motivation for using them. Why not use more (or less) parameters than these?

Similarly, it would be more helpful to provide motivations for including the "environmental features" mentioned in Line 72.

3. Line 75. *"We then implemented five machine learning and nine deep learning algorithms, leveraging radar and environmental features to predict penetration bias values."*

I would suggest removing "values" for consistency in the manuscript.

4. Figure 1. The figure seems unnaturally flattened. Moreover, the map of Antarctica should typically be in polar stereographic projection visualized with an equal aspect ratio, so I would not recommend flattening it.

I wonder what the point of showing the IMBIE drainage basins is...the concept of drainage basins is not used in the rest of the manuscript.

Caption: "...for which ICESat-2 data, TanDEM-X data and ECMWF features is extracted" -> "...features are..."

5. Lines 99–104. Salinity is mentioned here multiple times, but is it useful for the application over the Antarctic Ice Sheet?

6. Line 106. *"InSAR coherence also effects the penetration depth..."*

This sentence is grammatically and logically wrong. I would recommend writing "a deeper penetration depth results in a lower InSAR coherence due to an increase in volume scattering (Deng et al., 2024)" or something similar.

7. Line 111. I am not sure that the cited work says *"overall backscatter intensity decreases due to surface scattering"*. What I found in Park et al. (2021) is: *"backscatter over wet snow can also be increased due to the complex wet snow metamorphism, including an increase in snow surface roughness and an increase in the snow grain size during overnight refreezing"*.

8. Lines 114–117. It is nice to distinguish the different concepts between "penetration depth" and "penetration bias", but it would be nicer to clarify again (apart from the title) that this study focuses on the bias.

9. Lines 126–131. Is there any reference for the concepts mentioned here? It would also be more concise and clear if the relationships could be expressed in the format of equations. For example, it seems that Eq. 9 of Dall (2007) is helpful.

10. Line 134. It would be nicer to briefly explain to the readers of The Cryosphere what "interferogram flattening" is.

11. Line 146. The penetration depths between 10 m and 12 m seem contradictory to what is mentioned below. These values also seem to deviate from Table 6 of Rizzoli et al. (2017).

12. Line 152. The Michel et al. (2014) work is not about InSAR penetration bias, but radar altimeter. Suggested removing.

13. Line 167. Please specify which "previous models" were applied for which purposes, and please also provide references.

14. Line 182. “These data have been collected between...” could be misunderstood, possibly indicating that ICESat-2 collects data only in winter seasons. I would recommend the authors be specific that they are the ones who selected these winter data.

15. Lines 189–194. I understand the authors would like to have a detailed motivation and background introduction about the time difference between TanDEM-X and ICESat-2. However, adding too much information before introducing the time difference adopted by the authors dilutes the real message. I would recommend that here it is sufficient to only keep Lines 195–199.

16. Line 209. *“To retrieve elevation profiles, the collected dataset of point elevation measurements is sampled by utilizing different beam spacings and arbitrary repeat-track geometries, along with the estimation of Root Mean Square (RMS) errors result into the elevation change measurements.”*

The sentence does not seem optimal, both grammatically and logically. From what I understood, I would suggest rewriting into *“The elevation profiles are derived by sampling the collected dataset of point elevations using different beam spacing and arbitrary repeat-track geometries. The elevation change measurements are derived by computing the root mean square errors (RMSE)”*, or something similar. I am also not sure if I understood the differences (or similarities) between “elevation profiles” and “elevation change measurements”. Hopefully the authors could clarify this.

In addition, throughout Section 2.1, I found it difficult to distinguish which part is the characteristics of ICESat-2 ATL06 product, and which part is post-processing of the authors. I would suggest the authors split the two processes in two separate paragraphs.

17. Line 227. *“TanDEM-X is also designed to maintain a sharp balance between height of ambiguity with phase disambiguation (Zink et al., 2021).”*

I wonder how important this balance is, as it does not seem to be mentioned in the result analysis. If it is really so important, I also recommend the authors introduce what phase disambiguation refers to to the readers of The Cryosphere.

18. Line 229–231. The data acquisition period is repetitive, as it was already mentioned in Line 182. Meanwhile, most of the acquisition periods in Table 2 are 2021 to 2023 instead of 2024. Please be consistent.

19. Lines 232–240. Is it possible to add a flowchart to show how ITP and TAXI work?

20. Line 249. *“Fourteen atmospheric features complement our dataset (Table 1).”*

Why would the authors use these fourteen features to “complement” the dataset? This introductory sentence is quite abrupt and does not really show the importance of using these features. Meanwhile, these features were called “environmental features” and now “atmospheric features”. Please be consistent. Strictly speaking, snow density should also not be part of “atmospheric features”.

As an additional comment on the Data section, I recommend clearly stating the purposes of each dataset used by the authors.

21. Line 259. I wonder what the added value of repeatedly mentioning (also the modifications made to) the CSV files throughout the Methodology section is. While I appreciate that the authors provided the code and introduced the correct data format to the potential users, I believe such specifications would be more appropriate in the code documentation. Given the already extensive length of the manuscript, I suggest omitting these minor details.

22. Figure 3. I would appreciate it if the input and output of the models could also be specified.

23. Lines 270–272. My concerns are the same as comment 21. The projection is not critical compared to the core logic and methodology of this manuscript.

24. Lines 279–282. *"Without any correction applied the differences between ICESat-2 elevation data and TanDEM-X DEM height are characterized by an RMSE of 4 m and median and mean value of -1.6 and -1.48 m along with minimum, maximum values as -18 m, 20 m respectively in line with previous studies (Rizzoli, et al., 2017)."*

This sentence here is unclear with a grammatical mistake. Suggested change: *"Without any correction applied, the differences between ICESat-2 heights and TanDEM-X heights are characterized by an RMSE of 4 m, a median of -1.6 m, a mean of -1.48 m, a minimum of -18 m, and a maximum of 20 m. The statistics are consistent with Rizzoli et al. (2017)."*

Please feel free to improve the idea based on this version. In general, I suggest the authors add commas between clauses and consider splitting long and detailed sentences to improve clarity.

25. Line 283. *"Firstly, we have removed the data points lying in the floating ice region by removing the data points having TanDEM-X DEM height less than 100m."*

I am curious why the authors are not interested in the floating-ice region. Moreover, this sentence could be more concise. I would suggest revising into *"Points with TanDEM-X DEM heights below 100 m were removed to exclude the floating-ice region."*

26. Line 286. What does "finite accuracy" mean? Why is it important for this study?

27. Line 287. *"For heights of ambiguity ranging between 30 and 80 meters expected tilts in a scene range between 7 to 19 cm (Gonzalez et al., 2024)."*

If I understood correctly, did the authors try to say *"For heights of ambiguity ranging between 30 and 80 meters, the expected tilts in a scene range between 7 cm to 19 cm (Gonzalez et al., 2024)"*? I am not sure if I understood the purpose of mentioning this range "7 cm to 19 cm". Is it somewhat connected to the previous sentence "low-varying offsets and tilts"? Does it work as a quality indicator? Is it useful for the setup of the following experiments? Please specify.

28. Line 289. If the points with a height of ambiguity less than 30 m only take up 1% of the dataset, how do they introduce imbalance? Should we consider them as outliers? What about the points with a height of ambiguity higher than 80 m? Why would the authors not apply the same outlier removal criteria shown in Eqs. 1 and 2 also for the height of ambiguity?

29. Lines 319–323. Many of the abbreviations do not seem to match the ones in Figs. S1–S4. I suggest double-checking.

30. Table 3. Could the authors explain how the hyperparameters listed in the table are chosen?

31. Line 335. I would avoid using words like "obviously", as they can sound somewhat casual, and the observation may not be as obvious to the reader as it is to the authors. This sentence has another problem. Please refer to item 11 of technical corrections.

32. Line 339. *"Based on this analysis we decide to remove some of the highly correlated variables and leave only seven features including temperature of the air at 2m above the land surface, snowfall, surface pressure, snow density, snow evaporation, wind speed and wind direction."* -> *"Based on this analysis, we removed several highly correlated variables and*

kept the following seven features: 2 m air temperature, snowfall, surface pressure, snow density, snow evaporation, wind speed, and wind direction."

But it seems like this selection is only applied to the so-called "environmental features". The following text gives the impression that the so-called "radar features" are all used. If this is the case, I suggest the authors state clearly what exactly the retained features are.

33. Figure 6 caption. *"ICESat-2 elevation data is not used as input; only TanDEM-X DEM values are considered for the AI models."*

Please check the consistency of the manuscript, as "data" should generally be plural and seems to be also used that way in most parts of this manuscript. Regarding this sentence itself, I find it appearing rather randomly. It does not really describe the figure, and the authors filtered out a lot of other parameters apart from the ICESat-2 heights, so why would the ICESat-2 elevation data be mentioned here only?

34. Line 361. I find it odd to suddenly mention passive sensors. They are not used in this study as far as I understood.

35. Line 366. This sentence is unclear due to grammatical issues and dense structure. If I understood correctly, it would be better to rewrite it into *"Initially, all models are trained using 6 radar features as input: TanDEM-X DEM height, coherence, amplitude, height of ambiguity, local incidence angle, and slope. Penetration bias is used as the target variable."* Please also decide whether the authors prefer to use TanDEM-X or TDX throughout the manuscript.

36. Line 368. This is another sentence that I am not sure how to interpret due to grammatical issues and dense structure. According to my understanding, I would suggest rewriting into *"For a 30-day time difference, the dataset contains 276,549 data points as input. This number is reduced to 215,620 when considering a 15-day time difference."* Please also check for consistency when you mention "time difference" or "time interval".

37. Line 383. "...RMSE values nearer..." I cannot see the point of using a comparative form of an adjective. Please rephrase.

It is also important to mention the units of the errors throughout the Results section.

38. Line 384. "The reason may be cited..." is quite unconventional. Based on my understanding, here is a suggested revision: *"The probable reason for the compromised performance of the Decision Tree regression model is that it uses a single tree, which makes decisions by splitting features at threshold values."* Please feel free to improve based on this, if I understood it wrong.

39. Line 393. The introduction of the features is repetitive. There should not be a comma after "including" in this sentence.

40. Line 395. This sentence is difficult to follow due to missing verbs, dense structure, and inconsistent pluralization. Please consider rewriting into *"Table 5 shows that DNN is the best-performing model among all the data models with an RMSE of approximately 1.04 m, an MAE of 0.74 m, an R2 score of 0.84, and an explained variance of 0.84",* or something similar. In general, please specify the reported metrics precisely and avoid vague phrases such as "almost equal to" or "along with high corresponding values".

41. Figs. 7–10. For a more straightforward comparison, I would suggest combining them into one figure. There could be four subplots showing four scatter-plots, and one subplot showing all histograms in different colors. In addition, these figures, together with Fig. 4, seem unnaturally stretched.

42. Line 401. "Similarly, Table 5..." Please refer to comment 40 to improve clarity.

43. Line 405. “Figure 9...” The sentence has the same problem as the comments mentioned above. Suggested rewriting: *“Figure 9 shows the DNN model results using data with a 15-day time difference. Compared to the case with a 30-day time difference (Figure 7), a slight improvement in performance is observed.”* Please also check the consistency of the preferred word for “time difference between TanDEM-X and ICESat-2” throughout the manuscript.

44. Line 407. “...reduced between” is incorrect. Please specify: is it “reduced to” or “reduced by”?

45. Lines 410–419. This is a rather extensive description of the Supplementary material. As already mentioned in general comments, would it be clearer to put everything in Appendix?

46. Line 435. *“As the DNN model with model architecture given in Figure 11 is most efficient...a SHAP analysis...”* This sentence seems to imply that SHAP analysis was performed only because the DNN model showed the best performance. However, Line 459 mentions that SHAP analysis was also conducted for the RF model and the corresponding figures are not shown. This gives the impression that SHAP was applied to both models regardless of performance. To avoid confusion and to improve the flow, I recommend briefly introducing the SHAP analysis and its purpose in the Methods section, so that its use in the Results section does not seem unexplained.

47. Line 452. Would the result change if the authors use more or less than 1000 samples in the SHAP analysis?

48. Line 457. *“...wind speed, wind direction and surface pressure which also shows broad data distribution...”* Did the authors mean surface pressure shows a broad data distribution? But from what I understood, all input features should show a broad distribution. What is the purpose of mentioning it here?

49. Line 475. When the authors mentioned the corrected TanDEM-X DEM, is it possible to provide a map indicating the difference before and after the correction? I assume it can be interesting to see the spatial patterns of the improvement, apart from pure statistics.

50. Line 481. This was already mentioned in general comments. When the authors mentioned that the corrected TanDEM-X DEM achieves an accuracy similar to that of stereophotogrammetric DEMs, it would be interesting to see a comparison with REMA.

51. Line 482. It would be clearer to specify that the “block-wise calibration” is a procedure used in the Wessel et al. (2021) study. I am also not sure if I understood “residual elevation scatter”. According to the Wessel et al. (2021) study, *“For the HPB the mean penetration bias and standard deviations vary from −1.68 to −5.66m and 0.92 to 1.20m, respectively. They are located in the interior of Antarctica and are well distributed over the continent to serve as ground control for the adjacent blocks.”* These values (0.9 m to 1.2 m) seem to refer to the standard deviation of penetration bias.

52. Line 486. *“For example...DEM products can be integrated with other datasets...”* This sentence is extremely vague. Please specify how to integrated the DEM with which datasets for which purposes.

53. Line 490. Change detection is a similar idea to the “elevation change” concept mentioned in the following section. Would the authors consider merging the two sections and making the story more concise?

54. Line 498. *“This is critical because even modest systematic errors can translate into...”* It would be better to specify that these errors are in heights.

55. Line 509. *"Notably, in radar percolation zones where seasonal melt occurs, uncorrected InSAR measurements can differ by up to 10–15 m from the actual surface (Wessel et al., 2021)."*

I am not sure if I follow this logic. In principle, percolation zones with seasonal melt should be characterized by high-density rough layers on the surface and near the surface. I would not expect such a high penetration bias. I looked up how Wessel et al. (2021) commented on the difference between dry and percolation zones: *"Here, the temperatures are coldest (Macelloni et al., 2019; Scambos et al., 2018), and the SAR signal penetrates the most in dry, cold firn (Ulaby et al., 1986), whereas the coastal areas show lower penetration which clearly corresponds to the brighter reflecting percolation areas in the amplitude mosaic (Fig. 11). This variation in the SAR penetration over the whole AIS raises the absolute linear error (LE90) to 6.25m, which is calculated by sorting the absolute differences thresholded by 90% of the values."* I strongly recommend the authors specifying what they meant.

56. Line 514. "bias-corrected bistatic InSAR data": I assume the authors meant that they corrected for the DEM rather than the data themselves?

57. Lines 531–535. I am afraid I got lost in the logic of this paragraph. The paragraph seems to imply that over a long (30-day) time difference (again, please make it consistent and choose between difference, period and interval) between TanDEM-X and ICESat-2, both surface changes and volume changes can occur. But the authors seemed to try to explain why coherence outperforms amplitude, which means that the volume change should be more influential than surface changes?

58. Line 546. *"...underscoring that encapsulates a different aspect of the scattering physics..."* I tried to help revise this, but this part of the sentence is hardly understandable.

59. Line 587–589. I am not sure if I understood it. Did the authors mean that the ECMWF data are in a coarse spatial resolution, whereas the real-world climate effects are rather local hence require a high-resolution model?

60. Line 610. *"...confirms that SAR can serve as a powerful tool for measuring ice topography..."* Would it be more precise to say it is bistatic InSAR? Furthermore, on Line 615, the concept of "swath-based InSAR" is suddenly introduced. I would strongly recommend the authors be more consistent with the language they use to refer to TanDEM-X.

61. Line 620. "deep drawdown at a glacier terminus": does it mean a steep topography or a pronounced ice loss?

62. Line 630. I wonder why optical sensors should be mentioned suddenly.

Technical corrections

1. Please check in Table 1 and throughout the manuscript the misspelling of ECMWF.

2. Line 133. "...and interferometric phase change..." -> "...an..."

3. Line 134 and throughout the manuscript. Please check that there should be a space between the text and the citation. Some commas and dots are missed here and there.

4. Line 135. "almost equals" -> "almost equal"

5. There are two Rizzoli et al. (2017) reference entries that should be distinguished with 2017a and 2017b.

6. Line 164 and throughout the manuscript. The use of "like" to introduce examples is informal in academic writing. It is strongly recommended replacing them with "such as".

7. Line 197 and throughout the manuscript. The abbreviation TDX is used in parallel with TanDEM-X. I would suggest the authors remain consistent with only one expression.

8. Line 205. “...with each pair is separated by distance of 3km...” -> “...with each pair separated by 3 km...”

9. Line 207 and many other places in the manuscript. (B. Smith et al., 2019) -> (Smith et al., 2019)

10. Line 214. “signal-to-noise ratio”

11. Line 215 and throughout the manuscript. The present continuous tense is generally not considered appropriate in academic writing. The accepted tenses should be past simple, present simple and present perfect. It is strongly recommended thoroughly checking it throughout the manuscript.

12. Line 273. “...the corresponding TanDEM-X elevation data and associated radar features is extracted from...” -> “...are extracted from...”

13. Line 284. “We consider only points where the TanDEM-X coherence is larger than 0.3 as high coherence values lead to a more accurate estimate of the interferometric phase.” -> “We consider only points where the TanDEM-X coherence is larger than 0.3, as high coherence values lead to a more accurate estimate of the interferometric phase.”

14. Line 285 and throughout the manuscript. When a reference entry is mentioned within the texts, it should be cited as “Gonzalez et al. (2024)”. I recommend the authors double check the citation rules.

Also this line, it should be “Similarly to Gonzalez et al. (2024), we...” For similar problems, please refer to item 24 of specific comments.

15. Line 311. “Before data feeding in the machine learning and deep learning algorithms, data is divided into training, validation and testing data.” -> “Before being fed into the machine learning and deep learning algorithms, the data is divided into training, validation, and testing sets,” or “Before feeding data into the machine learning and deep learning algorithms, we divide the data into training, validation, and testing sets.”

16. Line 323. “The accuracy of all these machine learning and neural network models are compared...” -> “...is compared...”

17. Line 334. “To understand potential interrelations between radar and atmospheric features we focus to plot histograms of our variables and calculate their respective correlation (Figure 5, Figure 6).” -> “To understand potential interrelations between radar and atmospheric features, we focus on plotting histograms of our variables and calculating their respective correlation (Figs. 5–6).”

18. Line 336. “Various environmental features like Ultraviolet Albedo and Near Infrared Albedo along with snow depth show high values of correlation.” -> “Various environmental features, such as ultraviolet albedo, near-infrared albedo, and snow depth, show high values of correlation.”

19. Line 337 and multiple places in the manuscript. This is an incorrect use of “whereas”; it is not typically used to start a sentence.

20. Line 344. “...ICESat-2 elevation data is excluded” -> “...data are...” Please also check throughout the manuscript for consistency.

21. Line 376. “best performed” -> “best-performing”

22. Line 386. “However, it’s not much effective when there is a complex interaction between the features of the input data. Whereas Random Forest which is an ensemble of multiple decision trees is very effective in complex non-linear regression tasks.”

This paragraph contains serious issues in both grammar and clarity. The first sentence is grammatically incorrect (“it’s not much effective” is not standard English and should be avoided in academic writing). The second sentence should not start with “whereas”, and the contrast introduced by “whereas” is misused (it is difficult to understand what kind of comparison is achieved here). I strongly recommend rewriting this part to improve both clarity and readability.

23. Line 399. “...a gaussian distribution of errors suggesting that...” -> “a Gaussian distribution of errors, suggesting that...”

24. Line 404. A comma should be added after “Interestingly”.

25. Line 454. Suggested revision: “The top five features which influence the DNN model prediction ability include coherence, amplitude, height of ambiguity, TanDEM-X DEM height values, and 2 m temperature. These features are ranked by importance.”

26. Line 456. “effecting” -> “affecting” or “influential”

27. Line 456. “Other significant atmospheric features which are affecting the DNN model includes wind speed...” -> “Other significant atmospheric features affecting the DNN model include wind speed...”

In this case, one can use a present participle (“affecting”), as long as it serves as part of a reduced relative clause.

28. Line 457. Suggested revision: “For 30-day-difference data, the top five features...”

29. Line 458. “The top performing features is coherence followed by amplitude. ” -> “The best-performing features are coherence and amplitude.”

30. Line 459. “...both types of datasets”

31. Line 459. Suggested revision: “We also conducted a SHAP analysis for the Random Forest model, the second-best-performing regression model after DNN. The SHAP analysis of the Random Forest model gave the same result as the DNN model for both the 15- and 30-day cases.”

32. Line 475. “...height achieves...”

33. Line 477. “about” is rather informal. I would change it into “approximately”.

34. Line 547. “...coherence excels when capturing volume scattering, while amplitude captures persistent surface characteristics, they are both essential.” -> “...coherence excels when capturing volume scattering, while amplitude captures persistent surface characteristics; they are both essential.”

35. Line 598. “Finally, the low importance of snow density underscores the need for improved snow/firn modeling in polar regions, current models may not represent the features most relevant to X-band radar, such as layering or grain size, reinforcing the value of using observable proxies like temperature and wind in the meantime.” -> “Finally, the low importance of snow density underscores the need for improved snow and firn modeling in polar regions. Current models may not represent the features most relevant to X-band radar,

such as layering or grain size, reinforcing the value of using observable proxies such as temperature and wind."

Reference

Dall, J.: InSAR elevation bias caused by penetration into uniform volumes, IEEE Transactions on Geoscience and Remote Sensing, 45(7), 2319–2324, https://doi.org/10.1109/TGRS.2007.896613, 2007.

Howat, I. M., Porter, C., Smith, B. E., Noh, M.-J., and Morin, P.: The Reference Elevation Model of Antarctica, The Cryosphere, 13, 665–674, https://doi.org/10.5194/tc-13-665-2019, 2019.

Kuipers Munneke, P., Luckman, A. J., Bevan, S. L., Smeets, C. J. P. P., Gilbert, E., van den Broeke, M. R., Wang, W., Zender, C., Hubbard, B., Ashmore, D., Orr, A., and King, J. C.: Intense Winter Surface Melt on an Antarctic Ice Shelf, Geophys. Res. Lett., 45, 7615–7623, https://doi.org/10.1029/2018gl077899, 2018.

Michel, A., Flament, T., and Rémy, F.: Study of the Penetration Bias of ENVISAT Altimeter Observations over Antarctica in Comparison to ICESat Observations, Remote Sensing, 6, 9412–9434, https://doi.org/10.3390/rs6109412, 2014.

Park, J., Forman, B. A., & Lievens, H.: Prediction of Active Microwave Backscatter over Snow-Covered Terrain across Western Colorado Using a Land Surface Model and Support Vector Machine Regression, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 14, 2403–2417, https://doi.org/10.1109/JSTARS.2021.3053945, 2021.

Rizzoli, P., Martone, M., Rott, H., & Moreira, A.: Characterization of snow facies on the Greenland ice sheet observed by TanDEM-X interferometric SAR data, Remote Sensing, 9(4),1-24, https://doi.org/10.3390/rs9040315, 2017.

Slater, T., Shepherd, A., McMillan, M., Muir, A., Gilbert, L., Hogg, A. E., Konrad, H., and Parrinello, T.: A new digital elevation model of Antarctica derived from CryoSat-2 altimetry, The Cryosphere, 12, 1551–1562, https://doi.org/10.5194/tc-12-1551-2018, 2018.

Smith, B. E., Gardner, A., Schneider, A., and Flanner, M.: Modeling biases in laser-altimetry measurements caused by scattering of green light in snow, Remote Sensing of Environment, 215, 398–410, https://doi.org/10.1016/j.rse.2018.06.012, 2018.

Tsang, L., Durand, M., Derksen, C., Barros, A. P., Kang, D.-H., Lievens, H., Marshall, H.-P., Zhu, J., Johnson, J., King, J., Lemmetyinen, J., Sandells, M., Rutter, N., Siqueira, P., Nolin, A., Osmanoglu, B., Vuyovich, C., Kim, E., Taylor, D., Merkouriadi, I., Brucker, L., Navari, M., Dumont, M., Kelly, R., Kim, R. S., Liao, T.-H., Borah, F., and Xu, X.: Review article: Global monitoring of snow water equivalent using high-frequency radar remote sensing, The Cryosphere, 16, 3531–3573, https://doi.org/10.5194/tc-16-3531-2022, 2022.

Wessel, B., Huber, M., Wohlfart, C., Bertram, A., Osterkamp, N., Marschalk, U., Gruber, A., Reuß, F., Abdullahi, S., Georg, I., and Roth, A.: TanDEM-X PolarDEM 90m of Antarctica: generation and error characterization, The Cryosphere, 15, 5241–5260, https://doi.org/10.5194/tc-15-5241-2021, 2021.