Responses to the comments of anonymous reviewer #2

1) Overall quality and general comments

Rock glaciers are key indicators of permafrost in alpine regions, formed by a seasonally frozen detrital layer overlying supersaturated debris of ice or pure ice, and characterized by gravity flow. Their distribution is influenced by topographic and climatic factors at different scales, and they play a crucial role in high-altitude hydrology by storing ice and water. Traditionally, rock glaciers are classified as active, inactive, or relict based on ice content and movement. However, rising permafrost temperatures have led to an accelerating trend, encouraging an updated classification that considers sediment transport efficiency. In the regional territory of South Tyrol, two rock glaciers activity classifications coexist (Autonomous Province of Bolzano/Bozen and Bertone et al., 2019). By combining geomorphological characteristics, climatic driving factors, and InSAR products, the authors develop a statistical model to refine the classification of rock glaciers.

This study represents an innovative contribution since it integrates multiple variables into a multiclass generalised additive mixing (GAM) model to predict rock glacier activity. Using remote sensing, ground-based data, and digital terrain models, the workflow involves extracting velocity and environmental attributes at a regional scale, calibrating and validating a multiclass predictive GAM, and applying it to classify landforms based on their activity status.

The integration of remote sensing data and statistical modelling significantly advances current methods for assessing rock glacier dynamics. The study is well-structured, with a clear research objective and methodology. The statistical approach, particularly the use of a multiclass GAM model, is effective for the research aims. The discussion is robust, highlighting both its contributions and its limitations. The figures and tables are clear, informative, and support the understanding of the concepts. Finally, this work advances the understanding of rock glacier dynamics by refining their classification system and linking their activity states to a range of predictor variables.

2) Individual scientific questions

1.1.) 3.4.1. Statistical modelling

How did the authors ensure the robustness of the GAM model in terms of the selection and evaluation of predictor variables?

The selection of predictor variables for the GAMs model was conducted through exploratory data analysis, which enabled the screening of a broad set of morphometric and climatic descriptors (Table 2). From this analysis, eight variables were chosen based on the interquartile ranges that exhibited the greatest divergence among classes.

1.2.) 4.1. Exploratory Data Analysis

How did you decide which variables to retain for further analysis, and why were some variables, such as elevation, excluded to avoid redundancy despite their high discriminatory power? Could you clarify the rationale behind this choice?

From the exploratory data analysis, we selected variables that exhibited the greatest interquartile variations in boxplot distributions, as these differences enhance class discrimination. Some variables, such as elevation, were excluded despite their discriminatory power because they are strongly correlated with retained predictors, like land surface temperature (LST). Similarly, aspect and total insolation influence LST and were excluded to avoid redundancy, as their contribution is already captured through LST. This helps minimize redundancy of information, ensuring a more efficient and interpretable mode.

3) Specific comments on the manuscript

2.1.) Line 119: How many rock glaciers are present in the analyzed dataset?

The dataset used includes 1779 features. This information is reported at line 127 in the manuscript. We modified the sentence to better clarify this point.

2.2.) Line 127: The classification 'n.d.' is unclear. Could you please clarify its meaning and usage in this context?

We added the definition of "n.d." that stands for "not defined".

2.3.) Line 148: Could you explain in more detail how the variables were extracted and assigned to each individual rock glaciers?

Morphometric and terrain attribute analyses were conducted using ArcGIS 10.8 and SAGA GIS, based on a 10m DEM resolution. All derived products (e.g., slope, and aspect) were generated for the entire South Tyrol region and successively clipped over the boundary of each rock glacier presents in the dataset (1779 in total). For each feature, we calculated the mean values of environmental and climatic variables. Additionally, for the DInSAR-derived variables, we computed further statistical descriptors, including variance and the 25th, 75th, and 90th percentiles, to better capture their internal variability. The details of these analyses are provided in Section 3.4.1.

2.4.) Lines 207-210: "Using this SCD parameter, a potential correlation between the rock glaciers' activity at a regional level was made[...]" Could you explain this statement more clearly? How was the correlation assessed, and what were the main findings regarding the SCD in relation to the rock glaciers' activity?

We thank the reviewer for this comment and agree on the scarce clarity of the sentence. To enhance understanding, we have revised it in the text accordingly (Lines 206-210): "we do not consider SCD as a predisposing factor for the development of rock glaciers due to its implications for the thermal state of permafrost. Instead, we consider the temporal duration of snow cover in relation to the observed activity of rock glaciers, viewing SCD primarily as a factor influencing the modulation of activity states rather than as a prerequisite for their onset."

2.5. *) Figure 4: Does the term "look vector" refer to the Line of Sight (LOS) of the satellites? Could you also better explain if the shadowing and layover effects part is the C index analysis?*

Look vectors do not correspond directly to the LOS. They are the component of a 3D directional vector from the ground back to the sensors and they are described by two angles: the look vector elevation angle and the look vector orientation angle. The first measures the angle between the look vector and a horizontal plane at the ground pixel and indicates the sensor position above the surface. The latter is defined as the angle between the East direction and the projection of the look vector on the horizontal surface plane. These angles are considered in combination with the DEM to highlight the areas that, due to topographic and geometric conditions, are affected by layover, shadowing and foreshortening. C index is related to the evaluation of the visibility for each landform but can be better interpreted as representative of the percentage of movement detected from the satellite on the ground. So, after the exclusion of layover and shadowing areas, we used it as parallel information to quantify the robustness of the SAR measurement over each rock glacier. Rock glaciers expose N-S have a lower C value in comparison to those which have more favorably oriented towards east and west.

2.6. *) Figure 4: Is the vLOS referring to vertical velocity? Additionally, could you adjust the color scale bar to range between -8 and 8 cm/year to improve the clarity of the data representation?*

vLOS does not refer to the vertical velocity, but to the velocity component along the line of sight. We did not compute the vertical velocity, but kept the 1D LOS information because, as explained in lines 549-554, we prefer to mitigate the introduction of biases and assumptions that may arise from geometrical reprojections.

We modified the color scale.

2.7. *) Lines 244-248: "For each rock glacier polygon, mean values for environmental and climatic variables were assigned based on the values within the polygon boundary. Furthermore, for DInSAR-related variables (i.e., velocity and coherence), additional statistical descriptors [...]". Can you explain how the uncertainty was computed for each rock glacier, based on the SAR data coverage? How did you assess the spatial uncertainty within each polygon?*

The spatial uncertainty within each rock glacier polygon is not quantified by a single index, but it is assessed by evaluating the SAR data coverage and quality, also adding a filter on coherence (>0.25) and velocity (±2mm/yr). The C-map is also used to indicate the satellite's detection capability for each rock glacier, highlighting areas where signal coherence and data reliability might be reduced. Furthermore, we filtered the satellite data to exclude regions affected by layover and shadowing, ensuring that only valid pixels were included in the analysis.

2.8. *) Line 243: "Starting from the distribution map of the rock glaciers and considering their displacement range, we made two distinctions [...]". Could you clarify the rationale behind the choice of a 100-meter buffer around each mapped landform? How was this distance determined, and how does it affect the classification?*

The selection of a 100 m buffer was chosen since it provides a balance that ensures meaningful data extraction for analysis while avoiding excessive noise from unrelated features. In cases where adjacent or coalescing rock glaciers occur, the rims were cut to avoid any overlap between features, ensuring that the boundaries of one rock glacier do not encroach upon another. Additionally, due to differences in the orientation and spatial distribution of the landforms, it is highly unlikely for an entire rock glacier to fall entirely within the 100 m buffer zone of another. Regarding the suggestion for further analysis of the increment values: we concur that differentiating between intrinsic movement and externally driven movement is an interesting prospect. However, this aspect was not explored further in the current study, as our primary focus was on classifying rock glaciers activity. The incremental differences derived from buffer-based analysis could indeed serves as a basis for future investigations into the dynamics of rock glacier systems, particularly in distinguishing between intrinsic and external movement drivers.

2.9. *) Lines 264-266: “To discern the key factors influencing the distinction between A, R, and T rock glacier classes, we performed an initial Exploratory Data Analysis. This exploration served […]”. Could you provide more details on how this exploratory analysis was performed, and how it helped with the model?*

We refer to our response to comment 2.2. The Exploratory data Analysis (EDA) was conducting analyzing the distribution of statistical descriptors of morphometric, climatic and DInSAR derived parameters in all the mapped features grouped in the three main activity classes. This step is fundamental to extract the most representative variables controlling the distinction between A, R and T rock glaciers. In GAM, using less significative parameters as predictor variable would provide less sharp classification with associate a lower prediction capability. We thus considered a group of variables, eight in total, that have a physical control on the activity (e.g. LST, SCD) or are direct consequences of it (VRM, velocity etc.) and that, at the same time, provide a statistical distinction between activity classes.

3.10) *Lines 267-272: “GAM was employed to investigate the associations between the chosen predictor variables derived from both environmental and DInSAR datasets and the response variables. GAM provides […]”. Could you provide more explanation on the use of GAM in this context? A brief discussion of the relevant literature and how GAM has been applied in other studies would strengthen this section.*

We thank the reviewer for this suggestion. We have integrated the text with additional references on the application of GAM in similar studies. We selected GAM over a linear model because it can effectively capture complex, non-linear relationships between response variables and multiple independent environmental predictors. This flexibility is particularly important in our study, where the relationships between geospatial and DInSAR-derived variables may not follow a simple linear trend. By employing GAM, we ensure a more accurate representation of the underlying associations in our dataset.