Reviewer 1

We are grateful to Reviewer 1 for his constructive and detailed reviews, his enthusiasm for the method and for his perspicacity in picking some important flaws of our manuscript (e.g., the formalization of Equation 1). Critical comments included:

1. A better definition of climatic pre-eruption variables in the model;
2. A clarification of multicollinearity between features;
3. Limitations of model inference.

We have answered below each comment in a detailed way, which we believe addresses all of Reviewer 1's comments.

Minor comments:

- **Line 71**: Removed « five ».
- **Figure 1a**: We added labels where possible. However, labeling the innermost isopach is obscuring key features of the proximal deposit.
- **Line 172-174**: The sentence was rephrased.
- **Figure 2**: The figure was reworked.
- **Line 196-197:** Yes, we modified the text following this suggestion.
- **Lines 223-224**: We modified the text following this suggestion.
- **Equation 1**: Thank you for pointing that. In fact, expressing the CDI in a formal mathematical way is not trivial. We propose a new version of Eq. 1 and proper referenced its indices throughout the text.
- **Equation 1** and discussion: We have made this point clearer when introducing the CDI (Section 3.1.2).
- **Line 244**: This was corrected.
- **Line 253-54**: This is an excellent point – and not acknowledging this limitation is clearly an omission from our part. This is indeed a critical aspect that will be investigated in future iterations of the model once a satisfactory modeling method has been identified to account for multi-target predictions. We added clarifications just before and within the *Caveats and future research* section.
- **Line 285**: The use of steady-state analytical models such as Tephra2 is unsuitable for modeling a month-long eruption. We do not feel a justification is required.
- **Line 301-306** : We added clarifications in Section 3.2.2.
- **Section 3.2.4** – land cover: The year is indeed specified (i.e., 2015). We have added a comment addressing the limitation raised by the reviewer.
- **Line 333**: We removed any reference to *One-Hot Encoding*.
- **Line 381**: This is indeed very important and has been rephrased.
- **Line 452**: Done
- **Line 497**: The statement has been modified.
- **Lines 510-511**: It is difficult to explain the fundamentals of gradient-boosted trees in the manuscript. Therefore, we have removed this sentence from the text, but we added a short description of each parameter in the caption of Table 4.
- **Line 550-551**: This was removed as it is better explained further in the text.
- We source the use of *abiotic* vs *biotic* factors from existing literature (e.g., Arnalds, 2013). Note that although *abiotic* factors are commonly restricted to environmental parameters, we choose to keep *abiotic* factor to also englobe socio-economic components of agriculture (and especially crops) vulnerability. In addition, as described

in the discussion, we restrict here any causal inference "*to effects that rely on phenomena that have been either witnessed in the field or experiments".* In this sense, we agree that temperature and precipitation most likely have different roles in explaining the impact, but we are currently reluctant to *overinterpret* the results of our model.

- **L578**: About 2'500 different combinations of SHAP dependence plots can be produced. Although we focus here on 19 of them, the results highlighted in the text are based on an exhaustive review of a majority of them. It is however impractical to include and describe all of them in the text. Although we agree that their interpretation requires caution, we have adopted the most conservative interpretation when suggesting noticeable patterns in our data and only highlight them when they agree with other sources of information (e.g., post-EIA). In this sense, we have modified the text to only use conservative phrasing (e.g., "suggest" rather than "show") to stress the care required in inferring an unrealistic degree of causation.

- **Line 615**: Rephrased.

- **Line 636-638**: Thank you for pointing the role of elevation in our model. Firstly we have added more details in the discussion section that addressed this spatial dependency (Caveats and future research section). Secondly, we are facing here a problem of multicollinearity rather than simple correlations. As explained in Section 3.2, we have reduced an initial dataset of ~300 variables to about ~40 based on standard exploratory data procedures aiming at i) reducing the amount of collinearity between features and ii) improving the model's prediction by identifying and removing uninformative variables. This procedure is highly iterative and somehow standard, which is why we have decided not to include it into details in the manuscript. We would however like to point that elevation and landcover were not overly correlated. We have added clarifications in Section 3.2. Thirdly, one benefit of the *XGBoost* library and gradient-boosted trees compared to other decision trees is their ability to handle multicollinearity (we have added clarifications and associated references in Section 3.4.1). Finally – and linked to the previous point and multiple previous answer, variables in Earth and environmental sciences are rarely purely orthogonal. Again, this is a problem for model inference, which we keep to a minimum. Conversely, elevation proves to be an important variable even when the model is trained on individual landcover classes. This outlines how i) multicollinearity occurs over a variety of variables, which makes the choice of removing specific variables more difficult and ii) despite this, the model remains informative.

- **Line 660-665**: We added clarifications in Section 3.2.2.

- **Line 703/720-721**: We feel this goes back to concerns about multicollinearity and model inference, which we have already address. Please refer to comments for Line 578 and 636-638.

- **Line 712-715**: Rephrased

- **Line 781-784**: The paragraph was remodeled.

- **Conclusions:** A lot could indeed be added in the discussion and the conclusion. However, since the paper is already long and mostly represents an introduction to the method, we decided to focus on pragmatic conclusions.

Reviewer 2

Point 1

Earth Observation data provides a global observation network in space and time, which can be used as an indirect proxy to infer surface processes. EO sensors always balance three types of resolutions (i.e., spatial, temporal and spectral) to achieve specific purposes. For instance, Landsat provides 30 m pixels at a poor temporal resolution whereas MODIS provides 250 m pixels at a higher frequency. With this in mind, the EO community is entirely aware that satellite imagery is not able to resolve the complexity suggested by Reviewer 2, especially when, in our case, the methodology considers large areas and require dense time series. However, EO data enables the monitoring of vegetation over widespread areas, for which MODIS is often the preferred sensor. Therefore, we feel that i) the manuscript already extensively covers literature focusing on vegetation monitoring from EO data and ii) expressing and detailing this uncertainty is beyond the scope of the paper. Please note that all points stated in our discussion and future objectives target this exact purpose through validations and comparison with field mapping.

Point 2

We partly agree with this statement. We agree that impact mechanisms have been discussed for a long time in the literature, both from an ecological perspective (i.e., references proposed by Reviewer 2) and from an "impact" perspective (i.e., the dominance of references proposed in our manuscript), with the conclusion that no consensus is yet possible. One limitation to this is the opportunistic nature of studies in the field, which are too limited (both in number and in spatial coverage) to provide a sufficiently large number of observations required to capture the full variability of the involved processes (e.g., eruption types, climates, crop and vegetation types etc). We would fully agree with Reviewer's 2 comments should our method ambition to replace these field-based studies. However, the motivation for our method is the realization that generalizable models of volcanic impacts – at least for disaster risk reduction perspectives – probably will never be developed using only field-based studies, and we therefore explore here an alternative way to generalize these *in situ* observations rather than replace them. In addition, we fully acknowledge the limitations of our methodology, and limit causal inference to specific case-studies where impact mechanisms suggested by various sources point to supporting our interpretation. We therefore feel that most issues raised by Reviewer 2 in this comment are comprehensively addressed in our manuscript and supported by more recent literature, although using an impact rather than an ecological perspective.

Point 3

This statement has been deleted. Please note that i) the timing of the eruption relative to the phenological cycle of the plant is mentioned in the same paragraph, for which we have added relevant references and ii) further investigations of this relationship is identified as the first point for future iterations of the method in the discussion section.

Point 4

The method (and the CDI) are indeed based on vegetation indices, which provide a proxy for *biomass production*. Here, we attempt to provide a proxy for *impact*. Following comments of Reviewer 1, the purpose and limitations is now more detailed and discussed in the perspective

of existing techniques. In a nutshell, the CDI is designed to not only capture *negative impacts* but, on the longer term, to also capture the recovery. It was developed with the idea of quantifying impact as a *budget* (i.e., comparing short-term negative losses with potential long-term gains in fertility), and presents many advantages to estimates *rates* of impact and recovery compared to existing anomaly quantification methods. We believe that changes made to address Reviewer 1's comments also address this issue.

Minor comments

- Line 32 : Done
- Line 125: The map was reworked according to both reviewer's comments
- Line 175: Done
- Line 190: Done (and good to know!)
- Line 267: Done
- Line 269: Done
- Line 270: This sentence was rephrased following reviewer 1's comments
- Line 310: Done
- Line 800: Done