

Review of the manuscript ""Evolution of the Auckland Volcanic Field (New Zealand) boundary" by Craig Miller et al.

I have now read the manuscript (and I apologize once again for the delay). The study tries to provide a quantitative answer to the question "will the next vent in the AVF be outside the presently defined AVF boundary?". To this end, the authors adopt a set of analysis applied iteratively to the record of age-known eruptions in the AVF, as well as to a set of synthetic eruption catalogues generated on the basis of a (uniform? see below) spatial distribution over the AVF boundary defined by a previous paper Sporli and Eastwood (1997). Further, they quantify the effect of a precautionary buffer region around the boundary, analysing the effect of different buffer sizes.

In my opinion, the manuscripts needs major clarifications and substantial changes before being suitable for publication. In particular I try to list my major concerns (not in order of importance) in the following:

20/09/2026.

Dear Reviewer,

Thank you for your thorough and useful review of our preprint which has helped to strengthen our manuscript. We have addressed all your comments as detailed in reply below.

Regards

Craig Miller on behalf of the author team

Comment

1) some of the stated results and conclusions are obvious to me. For example, lines 221-224 state the obvious: if I generate N random points with a UNIFORM pdf within a predefined area (which is what the authors do with the synthetic catalogues), it is obvious that the area covered by the points increase with increasing N , so that it then becomes more likely for the $(N+1)$ -th point to fall within the previously-covered area. Same applies to the sentence in lines 207-208.

Reply

Thanks for the comment. We would like to emphasise that the generation of these synthetic data provide a critical comparison for us to examine the extent to which the growth of the convex hull of the Auckland volcanic field follows a random pattern, as well as the necessary buffer zones to apply for a field in its current growth stage (i.e., number of vents = 54). The statements we make in lines 207-208 and 221-224 are factually correct, and while we agree they may be intuitive to some (or many) readers they are still descriptions of our findings and approach and thus must be discussed in this paper for clarity of communication.

Comment

On the other hand, it would be interesting to quantify the "probability gain" by expanding the hull in time: in my opinion the authors should not only measure how often a new vent falls outside the previously-defined hull, but consider also how larger the new hull needs to be with respect to the previously-defined hull. Or, in other words, penalize a large hull that needs to be enlarged with respect to a small one. Would it be feasible? I think it would add information, more useful than just counting how often a vent falls outside.

Reply

Thank you for your interesting suggestion. We have tackled this in several ways and illustrated it in a new figure, Figure 9.

Firstly, we computed the cumulative probability of the next eruption being inside the convex hull which helps visualise the "probability gain" as the field develops. This is illustrated in Figure 9A.

We then implemented a penalty score based on the change in hull area at each eruption compared to the prior area. We compute a fractional growth (g_i) at each eruption as the change in area divided by the prior area.

$$g_i = \frac{A_{new} - A_{prior}}{A_{prior}}$$
$$gi = \Delta A / A_{prior}$$

From this we calculated a penalty score, S_i as

$$S_i = \frac{1}{1 + g_i}$$

This can also be thought of as a hull "efficiency", whereby an efficient hull is one that stays small and doesn't need to grow much, or if it does grow, it is only by a small amount.

Conversely an inefficient hull is one that needs to keep expanding to accommodate new eruption centre locations. The hull area change, and the efficiency score S_i , of each step is shown in Figure 9B

We use our hull efficiency metrics in two ways to modify the probability score. First, we compute a "spatial probability gain" score, PG_i shown as the purple curve in Figure 9C, plotted at each eruption.

$$PG_i = \frac{P_{inside}(i)}{A_{prior} / A_{final}}$$

Where $P_{inside}(i)$ is the running cumulative probability of the next eruption falling inside the convex hull divided by the ratio of the prior hull to the final hull. PG_i probabilities are larger than P_{inside} , reflecting the probability gain approach.

In the PGi the denominator is the probability of an eruption centre landing inside the prior hull that could be expected under a null model of spatial randomness. If new vent locations were drawn at random from a uniform distribution over the field's eventual full extent, the chance of falling inside a hull covering a given fraction of that extent equals the denominator. Therefore, if $PGi = 1$ the observed P_{inside} matches what would be expected from the hulls areal coverage and the field shows no more or less clustering than its footprint size implies. If $PGi > 1$ eruptions occur inside the hull more often than their areal coverage would predict. This suggests a degree of clustering beyond what the hull size suggests. If $PGi < 1$ then eruptions occur outside the hull more often than expected from the hull's areal coverage.

As such the PGi metric evaluates a hull's predictive success against its size, rather than just "in" or "out". PGi penalises hull that are large yet contain few events and rewards hulls that stay compact with successive eruptions.

For the AVF PGi doesn't exceed 0.86 and is typically 0.5 to 0.85 meaning the fields hull coverage has consistently grown at least as fast as its eruptions. It indicates the AVF spatial envelope has not yet shown clustering strong enough to "outperform" its areal growth, consistent with a still expanding, imperfectly bounded volcanic field, rather than one with a well-defined, highly predictable core region where eruptions frequently occur.

The second way the probability score is modified is using the hull step efficiency metric S_i as a multiplier on the P_{inside} , as shown by the yellow line in Figure 9C. S_i is a per eruption measure of how much the hull expanded to accommodate that eruption. $S_i = 1$ if the eruption landed inside the hull and $S_i \rightarrow 0$ depending on the size of expansion required. If the current eruption requires a large hull expansion, S_i drops and P_{inside} is lowered for that step. An eruption just outside the previous hull, only slightly penalised P_{inside} . Conversely if an eruption is inside the prior hull, no penalty is applied. The metric is designed to flag large expansions of the hull and minimise "just outside" eruptions, compared to the binary in or out counting of the P_{inside} . S_i penalised P_{inside} scores are lower than P_{inside} , reflecting the penalisation approach.

For example, the eruption of Wiri Mtn (eruption 31) and Green Mountain (eruption 47) only change the hull area by 0.018 and 0.043 km² respectively, meaning the penalty score $S_i = 0.9999$. Compared to Rangitoto eruption which added 21.05 km² resulting in a penalty score of 0.937.

The final probably P_{inside} taken after the last eruption of Rangitoto is modified to 0.61 for the Penalised Probability or 0.69 for the probability gain, compared to 0.65 for unpenalized P_{inside}

Comment

2) why do the authors generate catalogues of 250 events if they have 51 in the real catalogue? In figure 7a (as well in figure A1-a-d-g) they should stop the cumulative plot at x=51 on the x-axis, and maybe run a statistical test to check if the real curve is compatible with the ones based on the simulated catalogues (it does not seem to me). To me, it is quite striking that the synthetic catalogues seems unable to capture the real area covered by the 51 AVF centres, as the synthetic catalogues all seem to cover smaller areas than the real AVF. This is very important and casts doubts on the "uniformity" of vents in the real catalogue. In other words, if I have understood correctly the plot 7A, it shows that the 51-data synthetic catalogues do not capture the larger area covered by the real 51 centres. This is maybe due to the spatiotemporal clustering (see below point 3) found on the real data? Would it make sense to generate another set of synthetic catalogues of 51 events each, not uniformly random but considering the

clustering angle and distance shown in Figure 8b, and see if then the new synthetic catalogues capture the larger area?

Reply

We have also addressed this question from the first reviewer. Generating 250 events allows us to use the synthetic data as a predictive tool to look at the future behaviour of the field. We have updated the text to reflect our reasoning:

To extend the record beyond the range of the current observations and examine the probability of future eruptions occurring within the present field boundaries, we developed a synthetic dataset from 1,000 simulations of AVF development. Each realisation consists of a sequence of 250 volcanic centres randomly generated from a uniform distribution within the bounding ellipse defined by the parameters of Spörli and Eastwood (1997). Generating 250 centres provides a convenient yet realistic number that extends the simulated field beyond the range of the observed record, permitting it to be used as a predictive tool and ensuring that it is long enough to show the asymptotes in non-linear trends (see section 3.2). This allows us to determine how well the simulations align with the observed record where they overlap and then explore how those trends evolve in the future. Furthermore, many volcanic fields often have hundreds of eruptive centres with Valentine et al. (2021), reporting some fields such as Pinacate in Mexico having more than 400 eruptive centres in the Quaternary. Allen and Smith (1994) considered the AVF to be in its infancy and as we are interested in the future evolution of the field it is important to extend our analysis beyond the 54 reported eruptions to a number that is both mathematically stable and geologically realistic.

The apparent misfit of the observed and modelled data reflects an observational limitation of our dataset. The absolute spatial extent of the Auckland volcanic field is unknown and, for our calculations, is based on a best-fit ellipse (Spörli et al., 1997) based on the distribution of current vents. Therefore, the estimated area of the AVF reflects the minimum area, while the observed area ratios shown in Figure 7A reflect the maximum area-ratio.

The fact that the most recent eruptions lie above the predicted line (Figure 7A) suggest that the true area of the AVF is likely larger than the best-fit ellipse. As an example, if we increase the AVF ellipse area by 15%, the area ratios fall match the synthetic values (Figure 7A dashed brown line). If we assume that the AVF should follow a similar trend to the synthetic dataset (which is likely a fair assumption), this suggests that the true extent of the AVF likely lies somewhere between the current best-fit ellipse and a 15% larger area.

Comment

3) section 3.3: it needs to be clarified that you are looking to spatio-temporal clustering (clustering of subsequent events) rather than clustering by itself. The latter could mean that there are preferred clusters of location, regardless the succession of events.

Reply

We have clarified section 3.3 to reflect a temporal aspect. Note that our synthetic dataset has no temporal aspect, so spatio-temporal clustering is limited to it being the “next” eruption within the threshold distance. We have also improved figure 8, to show both spatial and temporal aspects of clustering on the observed data. The section is reworded as follows.

The locations of volcanic centres in the synthetic model simulations are randomly and successively drawn from a uniform distribution bounded by the AVF ellipse. As such, the distribution of distances between successive eruptions defines the null hypothesis to test whether vents within the AVF are also randomly distributed in both time and space. Across the synthetic dataset, the 5th percentile of distances between successive eruption centres is 2584 m, and we use this as a distance threshold for spatial clustering. By contrast, in the observed AVF, 18% of vents (9 out of 50) occur within 2584 m of the previous eruption and are therefore considered to be spatio-temporally clustered (Figure 8A). Seven of these nine spatially clustered eruptions occur within a 900-year time gap indicating, which is much shorter than the mean time between successive eruptions of 3800 years, further confirming the temporal component of the clustering. The azimuthal trend between the nine clustered centres indicates a NNE/NE structural alignment (Figure 8B)

Comment

4) discussion on completeness of the data: as figure 2 shows, there is a possible incompleteness in the AVF data (which is a very common problem), since the apparent increase in the eruptive rate in recent times compared to old data could be simply explained with the missing old events or old eruptive centres obliterated by younger ones. I know this is an unavoidable problem, but all of the analysis assumes that the record is complete, otherwise it makes no sense. So I think it would be honest to state this in the introduction and in the conclusions.

Reply

This is a good point and indeed in the last decade the number of known eruption centres has increased through new discoveries so we can't categorically rule out missing eruptions, although based on geophysical evidence we think it is unlikely. We have added sentences acknowledging this in the Chronology of Eruptive Centres section.

Comment

5) Section 3.3 and figure 8a: I needed to read this iteratively several times to figure out what the number 2584 is. I figured out, then, it is the 5th percentile of the distance between successive centres in the synthetic uniformly-generated catalogues. If this is the case, my suggestion is to rewrite all of this section starting from this, or in any case rewrite this part to explain better because it is really cumbersome in the present form.

Reply

We have rewritten this section for improved clarity: see reply to previous comment #3.

Comment

Also, I suggest to generate synthetic catalogues of 51 events only, recompute the 5th percentile of the distance between successive events and run a statistical test to check whether the observed number of real vents closer than that percentile is significant at 5% significance level (another option avoiding the test is to count how many catalogues have such a number of

successive events closer than that distance, and see if they are less than 1% or 5% of the 1000 synthetic catalogues).

Reply

We believe our current approach is similar to that suggested by the reviewer, except that it is more statistically robust.

Including only the distance between successive eruptions from the first 51 synthetic events produces a comparable lower 5th percentile value of 2,575m from 50,000 eruptions (50 distance values per simulation x 1,000 simulations). It, therefore, would not change our finding that 9 out of 50 events have occurred within that lower 5th percentile. However, including all 249,000 eruptions (249 distance values x 1,000 simulations) increases the statistical significance of the values and produces greater confidence in our estimated distance of <2584 m expected for 5% of the data.

Comment

5) Considerations on figure 6: I think the authors missed to consider the (apparent?) increasing rate of eruptions in time. When they compute that on average one event in 10000 years falls outside the hull with a standard deviation of 8000 years (lines 191-192), they assume that the rate of eruptions in time is constant, which is not (either a real feature or due to incompleteness as discussed in point 4 above, figure 2). I would have expected that the NUMBER of events outside the previously defined hull decreases in time (as the hull becomes larger and larger), and I think it does. The observed RATE of events outside the hull is a combined effect of the rate of eruptions increasing in time and the number of events outside the hull decreasing in time due to the hull area increasing. I hope I made this point clear, and I think it should be discussed. So lines 191-192 should be completely rewritten, and probably figure 6 should be combined with figure 2 in some way.

Reply

This is a fair point and we have modified lines 191-192 as follows: "Since 60 ka the rate of eruptions outside the convex hull is not uniform. Eruptions outside the hull are often "paired", with two events occurring close in time followed by long time gaps to the when the next eruption occurs outside the hull."

As Figure 2 presents the raw data so we prefer to leave it standalone rather than combine it with a figure that includes analysis of the data.

Comment

6) I am not sure this is the right word, but I think there is a bit of "misuse" of the word "random". "Random" is a general term and can have a different meaning, it can for example mean random according to poissonian model, or a uniform spatial model, or gaussian and so on, and still be random in the sense "stochastic". To improve the clarity, when the authors say "random", they should also specify according to what stochastic model, in my opinion, at least in some instances. For example, in lines 245 and 255 the authors state two very different things and both cannot apply simultaneously, I think: they say the spatial occurrence is Poissonian (I am not sure exactly what they mean with this) and then, a few lines below, that it is clustered. I think this

is very unclear, and to me confusing. Would it be possible to revise the draft and be more accurate on this?

Reply

We've clarified the use of the word random in places where it could be ambiguous and attached it to the type of model the random sample is drawn from. For example, in sections 2.2, 3.2 and 3.3 we refer to 250 volcano centres randomly generated from a uniform distribution.

In section 1 we now state the definition of the Poisson model – i.e. the location of an eruption centre is independent of the previous eruption centre. We note that a Poisson model can still show spatial clustering but not temporal clustering as it assumes events occur independently and with no memory of a prior event. However, two events could randomly occur close to each other at different times.

In the discussion we introduce the term Poisson cluster to describe the observed process. In a Poisson cluster a parent eruption centre generated by a Poisson process produces one or more child eruptions.

So even though AVF spatial distribution shows strong Poisson (non-clustered) attributes (Le Corvec 2013 figure 2) we show there is evidence for a Poisson cluster model operating at smaller scales spatially and temporally. The Poisson clustering is shown by volcanoes where the next eruption centre occurs within 2584m of the previous and with most of these events occurring within 900 years of the previous event. (Figure 8a).

The Poisson cluster process appears to have switched on and off throughout the last 200ka as shown by Figure 8a where eruption centres indicated by diamonds occur across the history of the field from eruption number 3 to 50. Importantly the 900-year temporal clustering is a direct observation and not based on a model that defines the temporal clustering timescale, however many of these dates are within the uncertainty of the dating method, so they could be more tightly clustered in time.