

Manuscript Summary

This manuscript details the changes in SO₂ and BrO emissions at Masaya using an unusually long degassing dataset, which in combination with re-analysis meteorology the authors use to investigate the trends and associations of data, whilst being rigorous in their retrieval methodologies.

General Comments

This is a well written manuscript with a generally high level of presentation throughout. It is logically structured and is easy to read. My one major comment is on the treatment of statistics in this piece, which reads as a little bit muddled (and confusing in places), with several elements lacking. These elements need to be improved before the manuscript can be published, particularly as some of the conclusions and discussion rely on some of the statements made about trends through time or difference between values, yet differences between times (BrO/SO₂ ratios for example at the different phases of the lava lake arrival and the activity) are only stated in a qualitative manner. More statistical rigour is needed. I outline this below in other general comments and in the specific comments.

The abstract is very long. I found it difficult to follow exactly what the key purpose but importantly the major conclusions and discoveries were. Needs shortening. Further comments below.

A very thorough discussion of methodology and sources of error throughout. Substantial and rigorous. Excellent.

A minor thing, I found it difficult that some Figures and Tables were presented before aspects of the main text, which explain some of the formation of the Figures, some rejigging to make sure that this doesn't happen would be great.

Treatment of statistics:

Part 1: Correlation vs. Regression. It is difficult to see what form of analysis has been performed, as frequent reference to 'correlation coefficients' are made, and yet the resultant number is provided as a percentage (with negatives occasionally). Correlation coefficients are presented in the -1 to 1 format (as you do in one point in the manuscript). The use of the percentage here throughout is confusing as we could commonly use the regression coefficient in this manner, i.e., an R² of 47% (47% of the variation in one variable can be counted for by another). The correlation would be reported as 0.47. This is where the confusion arises. Have you conducted regression and are providing an R² and calling it a correlation coefficient? Or have you conducted correlation and are formatting it incorrectly. If you were using Pearson's correlation, then the actual correlation value for an R² of 47% would be ~0.69. This is an important distinction, and it is important that the reader has confidence in the actual statistical technique used – correlation or regression.

Further, regarding regression. Any trend identified can then come with a p-value, is there a significant trend through time? So, where you identify a trend in the manuscript we also need to see the associated p-value to see whether this is the case.

Part 2: In parts of the manuscript where you are comparing significant differences between variables (between different phases for example), you need to statistically test this, i.e., avoiding the qualitative terminology currently used. After determining normality of the data, we can then use a variety of techniques for two-variables (t-test variations / Mann Whitney etc.) and others for three+ variables (Anova / Kruskal-Wallis) dependent on circumstances. This would allow rigorous interpretation of differences and back up the points you make in the manuscript.

Overall, the only major comment being the treatment of statistics I consider that this manuscript would be acceptable after minor revisions. The authors will need to be careful that the results of the additional statistical analysis match with the framing of the discussion.

Reviewed by Dr Tom D. Pering

Specific Comments

- Line 19: What is an 'extremely significant' annual cyclicity? Do you mean statistically significant? There are no degrees of 'extremeness' beyond this.
- Line 21: Correlation is not measured as a percentage; it is a standardized set of values between -1 and 1. So what is the -47% signifying? Is this regression?
- Line 57: I would say in situ methods are not able to retrieve bulk gas emissions, suggest removing 'may, however'.
- Line 65: Why are chlorine and fluorine compounds 'obvious candidates'? Needs more detail here.
- Line 77-79: Needs evidence, why is it the best accessible proxy for volcanic processes? References? Examples? The next paragraph (lines 80-95) then goes on to say that interpretation is difficult, so those two sections don't tie in together. Based on the subsequent paragraph a combined DOAS + MultiGas approach would seem far simpler. If by accessible in line 79, do you just mean that you can just use one instrument? If so, tailor that sentence in that manner.
- Line 170-175: Where was this data acquired from? Needs a link or detail.
- Line 216: What does 'hardly affected' mean? Is that the 10% figure at the end of the sentence? If so rephrase to use this value. 'Hardly' isn't quantifiable.
- Line 226: Retrieval of the background SO₂ slant column. Content fine, but it might be helpful to the reader to summarise the 'four approaches' into a Table.
- Line 302-304. Needs more context, why the 'actually'?
- Line 351: Coefficient – this isn't a correlation coefficient, do you mean 0.89? Latter fit suggests regression? Confusing statistical phrasing. See general comment.
- Figure 8: (d), you highlight a relationship between wind speed and plume height. What is the regression coefficient (the percentage model fit)? What is the p-value? Is it a significant fit? It is also unclear whether the fir is on the grey dots or the black dots.
- Figure 8: (e and f), this is unclear, did I miss in the text why you have split this up into 0-5, 5-10, and 10+ ms blocks [I note that I see this stipulated in text following the Figure but question remains for 5-10]? There appears (not tested) to be a broad relationship between flux and wind speed? So why separated? Needs justification. Same comment regarding statistical terminology.
- Line 375: OK, what does weak mean? What is the regression coefficient (R² value)? What is the associated p-value? Is this a statistically significant relationship? The scatter plot looks like a smudge of points.
- Line 380: Its only a best guess if you present the model with some statistical rigour, which it is not currently.
- Lines 384-400: Explanation here makes broad sense, but I wonder why you did not use a low flux threshold instead of wind speed? For example, omitting below 0.1 x 1000 t/d?
- Table 4 – Part 1: How is daily variation measured? Is this a standard deviation? Or range? Or iqr? And in each case how is the error determined? Is this 1 standard deviation? Particularly important to clarify.

- Table 4 – Part 2: annual trend and amplitude of cyclicity. How did you measured the trend? Can you have a significant trend of -0.1 with an error the size of the trend itself? P-values? Significance? How did you calculate the amplitude of the cyclicity?
- Line 461 and 462: Correlation coefficients listed here. Confirm that this is indeed what you have.
- Line 470-473: OK, significant variability and different averages. You need to test this statistically, see general comments details.
- Figure 9c: Orange label is for linear trend. But none is indicated? One would expect obvious linear trend in d (your residual plot) therefore. But it isn't obvious. Is this correct? Also how was your annual cycle determined?
- Line 479-481: See general comment on significant difference.
- Line 482: Remove the word 'extremely'.
- Line 486: What does 'basically the same' mean? The values afterwards look different to me.
- Line 490: Significance of the trend?
- Line 493: How did you determine outliers?
- Line 556: Rephrase to remove 'basically vanished'.
- Line 553-558: See general comments on correlation.
- Line 580: see previous comment, anti-correlation needs proof.
- Line 718: What does 'basically not correlated' mean in quantifiable terms?
- Line 715 onwards: Correlation terminology, see general comments.
- Line 787: Reference should probably be made to the Aiuppa paper here (already cited in this manuscript), which talks about this subject exactly.
- Lines 787-806: Interesting analysis, but framing of this will depend on the reassessment of statistics in the manuscript.

Technical Comments

- Line 10: Correct 'We make plausible'. Sentence needs shortening and clarification.
- Line 15-18: Shorten, too much distance between the mention of 'former periods' and what those periods are.
- Line 253 and 293 and 531 and 744: Don't use 'w.r.t' use with regards to.
- Line 295: 'There is also a significant number of scans'
- Line 299. 'Gaussian distribution' singular.
- Line 571-572: Rephrase needed.
- Line 577 – There *are* a number of possibilities
- Line 718: Remove 'basically'.
- Line 741: Alter phrasing away from 'basically vanishing'
- All graphs, check the 2 is subscripted in SO₂.