This paper presents a tephra fallout hazard assessment for two different types of eruptions, strong short-lived (SSL) and weak long-lived (WLL) considering ground load thresholds for roof collapse and vegetation. For SSL two different One Eruption Scenario (OES) are considered, 1990 (category 1) and 122 BC (category 2). For WLL hazard assessment is done for the 2002-2003 OES and also for an Eruption Range Scenario (ERS). In any case, eruption source parameters have been constrained on the basis of tephra deposits, which supposes a step forward. The paper is methodologically correct and focus on quantification of a relevant hazard. However, there are some (relatively minor) aspects that, in my opinion, should be better addressed before publication.

**General comments**

1. The authors use wind profiles from soundings at Trapani, at more than 200 km away from Mt. Etna. An obvious question is how these wind fields differ from those downwind Etna, very especially for low columns (WLL cases) and winds blowing from W (i.e. from Trapani, which seems to be the predominant direction). A mountain exceeding 3 km height does certainly affect lower troposphere wind direction and speed downwind. Is there any particular reason for using IAF soundings at Trapani and not other meteorological datasets (e.g. more local profiles extracted from ERA-Interim reanalysis)? Some comparison would be worth since this could affect somehow the results, particularly for WLL.

2. Some model input data for the scenarios is missing. A Table summarizing the main inputs for all scenarios/eruptions would be useful. From the text, I have gathered the following:

<table>
<thead>
<tr>
<th>Year</th>
<th>Scenario</th>
<th>Fi mean</th>
<th>H (km a.s.l.)</th>
<th>Mass (kg)</th>
<th>volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>SSL1 (OES)</td>
<td>-0.5</td>
<td>?</td>
<td>1.5x10^10</td>
<td>?</td>
</tr>
<tr>
<td>122BC</td>
<td>SSL2 (OES)</td>
<td>?</td>
<td>24-26</td>
<td>?</td>
<td>0.28 km³</td>
</tr>
<tr>
<td>2002-03</td>
<td>WLL(OES)</td>
<td>+0.5</td>
<td>?</td>
<td>?</td>
<td>43x10^6 m³</td>
</tr>
<tr>
<td>-</td>
<td>WLL (ERS)</td>
<td>?</td>
<td>3.6-7.0</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td>1998</td>
<td>SSL calibration</td>
<td>+2</td>
<td>12</td>
<td>1.3x10^9</td>
<td>?</td>
</tr>
<tr>
<td>2001</td>
<td>WLL calibration</td>
<td>+2</td>
<td>?</td>
<td>2.3x10^9</td>
<td>?</td>
</tr>
</tbody>
</table>

3. I partly disagree with section 4. The authors use two eruptions (1998 and 2001 for SSL and WLL respectively) to “calibrate” some model inputs through best-fit, including the diffusion coefficient. This can be very misleading because the diffusion coefficient depends on the particular meteorological conditions during the
days or the eruption. It is true that the values obtained (200 m\(^2\)/s for SSL and 1800 m\(^2\)/s for WLL) are consistent with turbulent diffusion being higher in the lower atmosphere, but the uncertainty is large. This should be at least mentioned. Also, TGSD from the “calibration” cases seems to be very different from those of the scenarios (see Table above). Given that particle size strongly affects the residence time, can this also affect the “effective diffusion”?

4. The authors consider thresholds for roof collapse (100, 200, 300 kg/m2) and damage to vegetation (10 kg/m2). However, all results except one are shown for roof collapse only (Figures 10, 11 and 14). Why? It would be more interesting to show same results (e.g. 300, 100 and 10) for the 4 cases.

5. In section 3 nothing is said about topography. Is this also an input for TEPHRA? If not, the impact on the results of assuming a flat topography for Etna should be commented, particularly for WLL.

6. Finally, aggregation processes are not even mentioned. I know the complexity of accounting for aggregation in the hazard assessment, but this limitation should be at least mentioned. How could affect the results?

**Minor comments**

- Pg 2, line 22. “distance from” \(\rightarrow\) “distance to”
- Pg 3, line 11. Actually, none of the references cited here, except Barberi el tal. (1990) use numerical models but analytical/semi-analytical. There are several other examples in literature for fully numeric tephra hazard assessment.
- Pg 3, line 12. Reference Costa et al. (2006) does not deal with hazard assessment.
- Pg 3, line 15. “numerical models of…” \(\rightarrow\) “models of…”
- Pg 3, line 20 onwards. Last paragraph describes violent Strombolian activity, but nothing is said regarding subplinian, although both types have been introduced in line 20-22.
- Pg 4, line 1. “airports in Catania and Reggio…” \(\rightarrow\) “the airports of Catania and Reggio…”
- Pg 4, line 28. “hazard assessment has been evaluated…” \(\rightarrow\) ”hazard has been evaluated…”
- Pg 5, line 1. “The model was implemented to include…” \(\rightarrow\) “The mode was configured to include…”
- Pg 6, line 2. “These kinds of” \(\rightarrow\) ”This kind of…” ??
- Pg 6, line 17. Computing techniques?
- Pg 6, line 18. MPI is a library, not a series on commands.
- Pg 7, line 15. What do you mean by “small fall time”? Unclear. I would better say “For particles having a fall time lower than FTT”.
- Pg 7, line 16. \(\sigma_{ij}\) is not defined.
- Pg 8, line 2. The parameter C cannot have units of m\(^2\)/s given the power 2.5 in eq. (2).
- Pg 8, line 13. “at the time zero” \(\rightarrow\) ”at time zero”
• Pg 8, line 20. “implemented” \(\rightarrow\) “modified” or “configured”
• Pg 9, line 20. “events to occur” \(\rightarrow\) “events occurred”
• Pg 10, line 3. “where it began to blow” \(\rightarrow\) “where it rotated”
• Pg 10, lines 13-14. “equals to” \(\rightarrow\) “equal to”
• Pg 10, line 24. “of altitude from a vent” \(\rightarrow\) “above a vent”
• Pg 11, line 3. “changed between” \(\rightarrow\) “varied between”
• Pg 11, line 5. “10800s, PR” \(\rightarrow\) “10800s, and PR…”
• Pg 11, line 10. “data taken from” \(\rightarrow\) “data from”
• Pg 12, line 21. The deposit in 1990 covered the WNW flank, implying predominant wind direction of 315° (according to the author’s wind direction criterion). This seems to be a very anomalous situation according to Fig 7b and 7c…
• Pg 13, line 10. It is a bit surprising that granulometry for 1990 is much coarser than for other weak eruptions…
• Pg 13, section 6.2. Granulometry for SSL2 is not given. Is assumed equal to that of 1990?
• Pg 14, line 2. I would not give this volume with 3 decimal digits…
• Pg 14, line 15. “Tall eruption columns…” \(\rightarrow\) “Eruption columns…”
• Pg 14, line 17. Falls?
• Pg 15, line 21. “makes a comprehensive…” \(\rightarrow\) “compiles a comprehensive…”
• Pg 15, line 22. “reliability” \(\rightarrow\) “accuracy”?
• Pg 16, line 2. I do not understand why these references here, all from other hazards…
• Pg 16, line 4-5. No. Parallel programming has nothing to do with “improving of physical models”. At most, it allows to run more complex models efficiently…
• Pg 16, line 7-8. Unclear sentence.
• Pg 16, lines 11 to 15. This paragraph is true only in part. Even if the quasi-steady approach is used, semi-analytical models still have the limitation of homogeneous wind field, and therefore can not be used for long-range transport (fine ash).
• Pg 16, line 16. “It should be noted, for SSL…” \(\rightarrow\) “It should be noted that, for SSL…”
• Pg 16, lines 25-26. Disagree. See general comments on this.
• Pg 17, line 12. “have shown” \(\rightarrow\) “show that”
• Pg 17, line 15. “us” \(\rightarrow\) “as”
• Pg 17, line 22. Please explain better…Why OES-SSL1 affects only agriculture? How is the big/small is the mass in this scenario compared with WLL (values missing in the table above)?
• Pg 18, line 19. I do not think that Eulerian/Lagrangian is relevant here…

• Figure 6. Do you have 4 soundings per day at Trapani?
• Figure 6. The wind direction criterion used throughout this paper is not the standard one used in meteorology (0° wind coming from N, rotation clockwise).
• Figures 9,10,11 and 14. Difficult to see. Should be zoomed around the area of interest (where contours exist). All contours should be labeled.
• Figures 12 and 13. For the ERS, the sampling strategy seems to be purely random, which is not the best choice in this case, with too little sampling. For example, why the peak in Fig 13?