Interactive comment on “New strategies for vertical transport in chemistry-transport models: application to the case of the Mount Etna eruption on March 18, 2012 with CHIMERE v2017r4” by Mathieu Lachatre et al.

1 Answer to Anonymous referee #1, received 16 may 2020

We wish to thank the referee for his/her helpful comments. The comments of the referee are in bold, our answers in normal black, the new elements added to the text are in blue.

1.1 Major comments

Firstly, this paper seeks to address two major concerns regarding vertical transport: 1. Vertical transport is poorly represented in most modern chemistry transport modeling efforts, resulting in excessive numerical (and eventually horizontal) diffusion; and 2. The naïve, or brute-force, solution to this increasing the number of levels in the simulation is expensive. This paper has done an excellent job of exploring answers to the first question, but does not provide any insight into the second.

The two “smart” solutions which the authors propose have their own downsides; the Després and Lagoutière (hereafter DL) advection scheme, while antidiffusive, is also only first-order accurate, while the “directly interpolated winds” (hereafter WRFW) approach violates mass conservation. The utility of the paper would be significantly increased if the authors gave a quantitative assessment of the computational overhead associated with each method and compared it to that associated with the naïve approach. Timing alone, in terms of the number of CPU-hours spent on each simulation, would help with this.

Parameters / Resolution	20	50	99
NODIV-VL	943	1177	1376
WRFW-VL	938	1193	1380
NODIV-DL	957	1193	1389
WRFW-DL	936	1187	1302

Table 1: Number of CPU hours for each simulation setup

The number of CPU hours spent on each simulation is provided in Table 1 above. They do not fit any theoretical scaling. The scaling of the computational load relative to the number of vertical levels n is known to be at least proportional to n (and proportional to n^2 if the CFL in the vertical direction constrains the timestep). Here our observed the scaling is sub-linear which is unexpected.

Here the configuration was 384 CPUs for CHIMERE and 128 CPUs for WRF. The configuration of CHIMERE is extremely light, with only 1 advected species and no chemistry, so that most likely the meteorological simulation, an extremely complex process with several prognostic variables, was using most of the CPU time, with the CHIMERE CPUs likely spending part of the time waiting for the input meteorological fields, at least in the lightest configuration with 20 model levels. It would have been more efficient in terms of computational time to use fewer CPUs for CHIMERE at least in the simulations with 20 levels to balance the load between meteorology and chemistry, but since the point here was to compare the results of the various simulations we preferred to choose an "all other things being equal approach" where the only change in configutaion between a simulation with 20 levels and its 99-levels counterpart is the number of levels.

This underloading of CHIMERE CPUs is very specific to the present configuration since we advect only one species (typically hundreds of species in a CTM simulation). We have observed that in full-fledge CHIMERE simulations with realistic chemistry and using pre-calculated meteorological fields the scaling of computational time according to the number of vertical levels is linear or superlinear.

Due to these limitations, we are unfortunately not able to use our results to provide a more precise information on computational cost.

Similarly, the lack of mass conservation in the WRFW approach causes serious concern. I applaud the authors for their frankness in discussing this limitation. However I believe that a full understanding of the advantages and drawbacks of each approach demands a fuller discussion of this issue than is currently given in Section 3.2.

In Figure 3, it is not clear to the reader why the total domain mass differs so much between each simulation, and it is critically important to the core question of the paper to know why the mass is changing. Specifically, it would help greatly if the authors could quantify on or with Figure 3: 1. How much mass has been (erroneously) lost through the domain upper boundary, based on integrated vertical mass fluxes at the upper boundary;

In Figure 3 (reproduced below), considering the "NODIV" simulations which are mass conservative, SO_2 mass loss is only due to fluxes through the model upper boundary. For this wind strategy, the differences between 20, 50 and 99 vertical levels are explained by the plume proximity to model upper boundary, which can be observed in Figure 7.

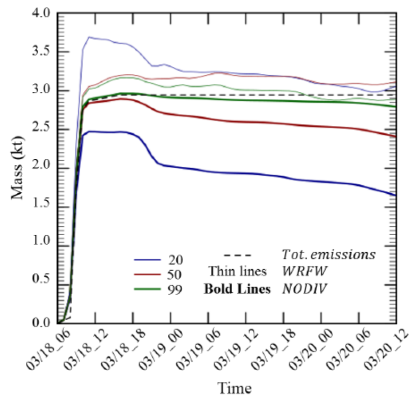

Figure 3. SO_2 mass evolution in model domain (kilotons). Line color indicates the vertical levels configuration, thickness indicates the vertical wind strategy considered. Dotted line represents the cumulated SO_2 mass emitted during Etna volcanic Eruption.

and 2. How much mass has been lost through the domain side boundaries, based on integrated horizontal mass fluxes at the domain boundary. These quantities should enable the authors (and reader) to determine how much of the mass at a given time is spurious, and the degree to which loss through the boundaries is offsetting artificial mass production.

On this note, on lines 2-3 of page 14, the authors mention that the "spurious evolutions in tracer mass become weaker, less than 5 %" once the plume is more diffuse. Does this really mean "the total domain mass is $< 5\,\%$ of the total emitted mass", or is it saying that the amount of mass created spuriously in each time step is $< 5\,\%$ of the current domain total? I assume the former, but if so, does this really mean that the error is $< 5\,\%$, or just that the additional spurious mass is now offset by some loss of mass through the domain boundaries?

The negative trend due to leakage through top of domain is observed mostly in the simulations with 20 and 50 levels. For the simulation with reconstructed wind, this leakage is the only term of mass loss: therefore, we can identify easily the magnitude of this term without additional calculation.

The idea of error compensation is interesting. However, a close look at the curves shows that the decreasing trend due to mass leak at top of model is present in the simulations with interpolated wind as well (thin lines in Fig. 3 reproduced above), and with a comparable magnitude. We think that the effect of mass balance inconsistency due to the divergence of wind field are visible in the small deviations of the curves corresponding to the WRFW simulations around this long-term trend, giving them a more wiggly aspect than those with NODIV which display only a slow and steady decrease. These small movements

$$\frac{\partial \widetilde{C}_{i,j,k}}{\partial t} + \left(\widetilde{F}_{i,j,k+\frac{1}{2}} - \widetilde{F}_{i,j,k-\frac{1}{2}}\right) + \left(\widetilde{F}_{i+\frac{1}{2},j,k} - \widetilde{F}_{i-\frac{1}{2},j,k}\right) + \left(\widetilde{F}_{i,j+\frac{1}{2},k} - \widetilde{F}_{i,j-\frac{1}{2},k}\right) = -\varepsilon_{i,j,k}, \qquad (6)$$

are indifferently positive or negative. We think that the effect of this term is not necessarily the effect of "additional spurious mass", but can be indifferently positive and negative as shown by Eq. 6. Actually, errors in discretized calculation of divergence will tend to compensate each other between neighbouring cells, so that we think that the relatively weak effect of the mass inconsistency term as soon as the plume is spread over many cells is due to this error compensation, between neighbouring cells:

If, for example, $\widetilde{F}_{i+\frac{1}{2},j,k}$ is overestimated, this will introduce a negative contribution in $\varepsilon_{i,j,k}$ but a positive and opposite contribution on $\varepsilon_{i+1,j,k}$.

This also explains why the $\varepsilon_{i,j,k}$ term has a much more drastic impact in the first hours of the eruption, because in these hours a substantial part of the total tracer mass is concentrated in one single cell above the vent: then the sign and magnitude of the error term $\varepsilon_{i,j,k}$ in this precise cell becomes critically important and no error compensation occur since the opposite errors on neighbouring cells will act on much smaller tracer concentrations.

Two new paragraphs have been added in Section 3.2 to discuss these points :

In the simulations with the reconstituted non-divergent wind field, substantial mass leak through the top of model can be observed as soon as the injection starts in the 20-level simulation (in which injection is done in the highest model level): the mass of tracer present in the domain never exceeds 85% of the expected mass. For the simulation with 50 vertical levels, this phenomenon is also visible. Another strong episode of mass leak through model top occurs in the simulations with 20 and 50 vertical levels and with reconstructed wind fields from March 18, 18UTC to March 19, 00UTC. This episodes causes an additional drop in tracer mass of 20% in the simulation with 20 levels, 5% in the simulation with 50 vertical levels. This episode of leak also affects the simulation with 20 vertical levels and with interpolated wind fields, reducing tracer mass concentration by about 10% from March 18, 18UTC to March 19, 00UTC. In these three simulations (20 and 50 levels with non-divergent winds, 20 levels with interpolated winds), a continuous decreasing trend in tracer mass is observed throughout the simulation. This drop is directly attributable to leak through model top since the tracer plume is far away from the horizontal boundaries of the domain.

And:

No physical process can explain this overshoot, and it is directly attributable to the choice of lifting the mass conservation constraint in the formulation of transport in order to permit the use of a realistic wind field. If we take March 19, 00UTC as a reference time at which the eruption is terminated, the first strong event of leak through model top is terminated as well, we can observe that the mass evolution in all three WRFW simulations undergoes small variations from one hour to the next but stay confined in very narrow ranges : 3.3 to 3 kt for the simulation with 20 vertical levels, with a decreasing trend attributable to leakage through model top, 3.1 to 3.25 for the simulation with 50 levels and 2.9 to 3.1 kt for the simulation with 99 vertical levels. The fact that these variations in total mass become marginal in this latter part of plume advection, when the plume is spread over a large geographic areas reflect the fact that numerical errors in the evaluation of divergence mechanically tend to compensate each other between neighbouring cells so that their global impact on a plume that is dispersed over many cells is small.

A broader concern which does not appear to be discussed in detail is the fact that the simulation is driven by fields which are sometimes at a lower vertical resolution. CHIMERE is driven by WRF, running with 33 models, but CHIMERE interpolates this data to its target vertical resolution (Briant et al 2017). Is this interpolation done in a divergence-conserving fashion? If not, does this constitute an uncontrolled-for additional term, in the sense that different vertical grids could introduce different amounts of artificial divergence?

The interpolation of the wind fields is done in a linear fashion which in principle is divergence conserving, but CHIMERE interpolation works directly on winds and not mass fluxes which actually may bring some additional errors in divergence. Our concern was to have all simulations forced with the exact same meteorological simulation, and we decided to retain the typical number of levels that is used in CHIMERE (Briant et al., 2017). The statement that "different vertical grids could introduce different amounts of artificial divergence" is therefore correct. We explicitly draw the reader's attention towards this point in section 2.2.1 of the revised version:

$\varepsilon_{i,j,k}$ depends on the resolution of the meteorological model (which is identical for all our simulations), and on the resolution of the chemistry-transport model, so that this error term that essentially traduces divergence errors due to interpolation depends on the vertical resolution of the model. It is identical between simulations that have the exact same number of domains. Choosing interpolation strategies that reduce this error term is a promising path to mitigating excessive vertical diffusion, as discussed in Emery et al. 2011, but is not investigated here.

Finally, the authors rely heavily on the trajectory of the plume as a metric of the simulation's fidelity. While the equation to determine error (equation 16) is an interesting formulation,

This is true because the plume's horizontal location is the only reliable observation that we had, due to the large uncertainty and error bars in the satellite retrievals of its altitude. Therefore, indicators like the one in Eq. 16 were, unfortunately, our only way to provide a comparison of model simulations with real-world data. We agree that this measure is only an indirect way to observe potential improvements in the vertical direction and reduction in plume diffusion.

(...)it would be helpful to provide a more quantitative assessment of the amount of numerical diffusion. Variation in the maximum

volumetric mixing ratio, (...)

Figure 7 (reproduced below) displays the highest column vertical profile evolution for each simulation. It can be observed that SO_2 mixing ratio is highly impacted by diffusion parameters chosen (please note that the scale use is irregular), and that simulations with the WRFW-DL configurations preserve a much higher maximal VMR than their conterparts with NODIV-VL.

Also, we believe that Figure 8 as well as Figures S5 and S7 in the supplements that have been added in the revised version bring additional elements in this line. Generally speaking, we have chosen to look at a more synthetic parameter like the minimal volume containing 50% of mass plume rather than a value of maximal VMR, which is more dependant on the details of all simulations. Figures 8a and S5 can be directly interpreted in terms of VMR, since the typical VMR in the plume is inversely proportional to plume volume.

(...) the total area of the plume above some minimum VMR, or the total entropy would be useful for quantifying how much numerical diffusion is being introduced.

A calculation very similar to the one suggested by the Referee on area is already present in the manuscript (Section 3.6, Fig 8 of the submitted manuscript and Fig. S5 of the revised manuscript). Here we propose to use the minimum volume containing at least half of the SO_2 mass as a synthetic indicator of how much the plume has been diffused. This is very similar to the proposal of calculating the area above some minimum VMR except that we chose to did it in 3d with volumes instead of areas, and we thought that calculating the volume containing at least half of the plume was a useful method to avoid introducing an arbitry threshold on VMR.

On Figure 8b), we calculate the volume ratios for each parameters (i.e. WRFW vs NODIV; DL vs VL) to provide a quantitative assessment of diffusion reduction on 3 dimensions. To illustrate the differences implied on for plume's surfaces, Figure S6 and Figure S7 (in suppl.) have been added to show the horizontal dispersion of plume on various simulations after 2 days.

We believe that entropy is delicate to interpret for many people including ourselves particularly when, as it is suggested here, we do not speak of a thermodynamic entropy of air but on the artificial construction of a mathematical entropy value for a tracer distribution. We agree that entropy of tracer concentration fields is a useful way of measuring numerical diffusion but we feel that discussing issues in terms of plume volume as we have done is much easier to interpret for the particular case we treat here, as we deal with a physical quantity whose absolute value has a meaning.

This would also allow the authors to account for the effect that spurious vertical diffusion can have in accelerating spurious horizontal diffusion (relevant papers discussing this issue and metrics of numerical diffusion are e.g. Rastigejev et al 2010, Lauritzen and Thuburn 2012, Eastham et al 2017, Zhuang et al 2018).

We agree with the Reviewer that more discussion on this point was useful. The results we obtain are in line with Eastham 2017 and Zhuang 2018: reduction of vertical diffusion has a direct impact on horizontal diffusion as well. Here in

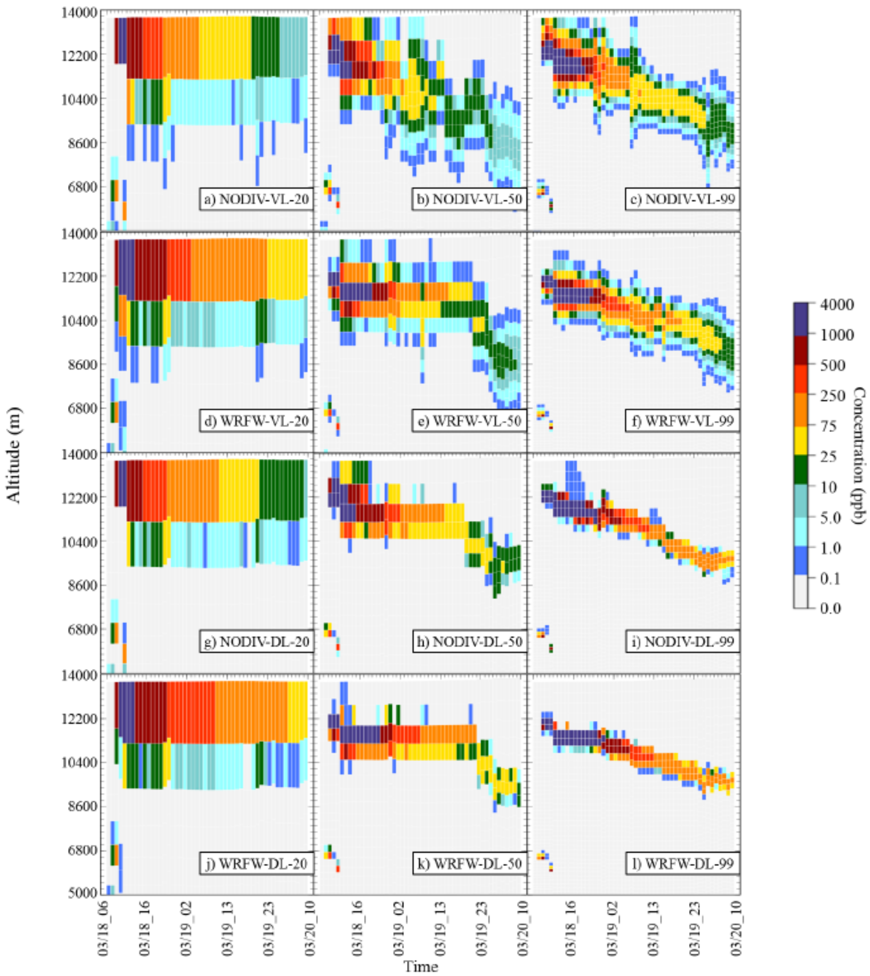

Figure 7. Evolution of SO_2 vertical profile (in ppb) corresponding to the maximum column for each step after the Etna eruption, for each tested model configurations. 1^{st} row: NODIV-VL; 2^{nd} row: NODIV-DL; 3^{rd} row: WRFW-VL; 4^{th} row: WRFW-DL. Left: 20 vertical levels; Center: 50 vertical levels; Right: 99 vertical levels. WRFW simulations values have been corrected to fit NODIV strategy masses.

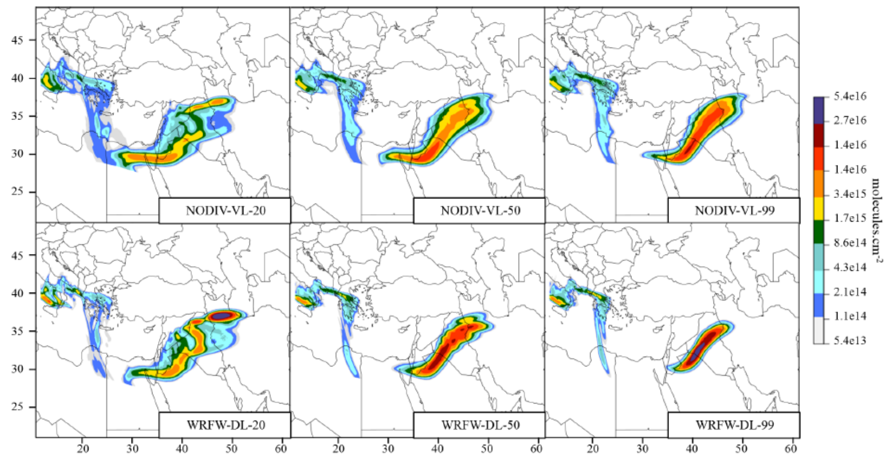

Figure S6. Volcanic plume integrated column dispersion on march 20^{th} at 11 A.M. UTC (2 days after the eruption).

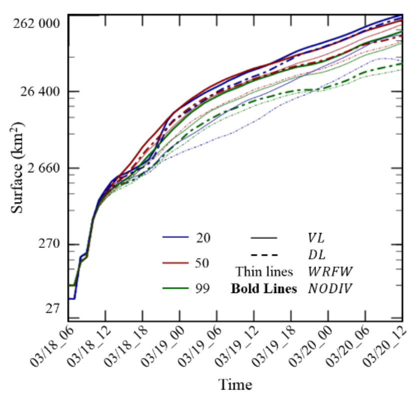

Figure S7. Minimum surface evolution calculated for 50 % of SO_2 total mass in the atmosphere.

the revised version we insist on the finding that this reduction can be obtained not only by improving resolutio but also, to some extent, by the approaches we advocate in the manuscript. A long discussion on this point has been introduced in section 3.6 based on new figures S6 and S7, and a corresponding statement is added in the conclusion as well.

1.2 Minors comments

I believe that there is an error in equation 15. Using the case of a local maximum (i.e. the first term of the Min operator is negative or zero), the estimated cell boundary VMR ends up being the cell mean VMR + 1, when it should presumably by the cell VMR only (specifically if this is meant to recreate the Godunov donor cell scheme for that condition). Although only a technical error, this is critically important to verification of the rest of the paper.

We are deeply grateful to the Referee for this in-depth investigation of our equation. This has permitted us to realize that there was actually a missing multiplicative factor in the equation and that this mistake would have made reproduction of our results in another model very difficult. The correct equation is as follows:

$$\bar{\alpha}_{s,k+\frac{1}{2}} = \alpha_{s,k} + \frac{1-\nu}{2}\text{Max}\left[0, \text{Min}\left(\frac{2}{\nu}\frac{\alpha_{s,k}-\alpha_{s,k-1}}{\alpha_{s,k+1}-\alpha_{s,k}}, \frac{2}{1-\nu}\right)\right] \times \left(\alpha_{s,k+1}-\alpha_{s,k}\right), \quad (1)$$

Even though the last factor was missing, the Referee's interpretation of the behaviour of Eq. 15 is correct and Eq. 15 would result into $\bar{\alpha}_{s,k+\frac{1}{2}} = \alpha_{s,k+1}$ in case of a local maximum, which would in our opinion lead to catastrophic instabilities since mass could never escape from a maximum whose downwind neighbour has zero VMR. As stated in the next sentence of the paper ("if $((\alpha_{s,k}-\alpha_{s,k-1})(\alpha_{s,k+1}-\alpha_{s,k}) \leq 0)$, no interpolation is performed and the scheme falls back to the simple Godunov donor-cell formulation"). This sentence may suggest that in case of a maximum the equation naturally falls back to the Godunov donor-cell formula. This is not the case. As we state more clearly in the revised version, Eq. 15 is applied if, and only if, the considered cell is not a local extremum, otherwise $\bar{\alpha}_{s,k+\frac{1}{2}} = \alpha_{s,k}$ is enforced:

As above, Eq. 15 is not applied in the case of a local extremum $((\alpha_{s,k}-\alpha_{s,k-1})(\alpha_{s,k+1}-\alpha_{s,k}) \leq 0)$. In this case, $\bar{\alpha}_{s,k+\frac{1}{2}} = \alpha_{s,k}$ is imposed and the scheme falls back to the simple Godunov donor-cell formulation

The same precision is broght for the Van Leer scheme (Eq. 14) since our initial formulation was suffering from the same ambiguity.

Section 2.1: it would be helpful to have details on how the vertical layers are placed (i.e. more detail on the different grid discretizations), and where the cell edges lie relative to the WRF vertical grid.

The various vertical resolutions can be compared on Figure S4 of the revised version.

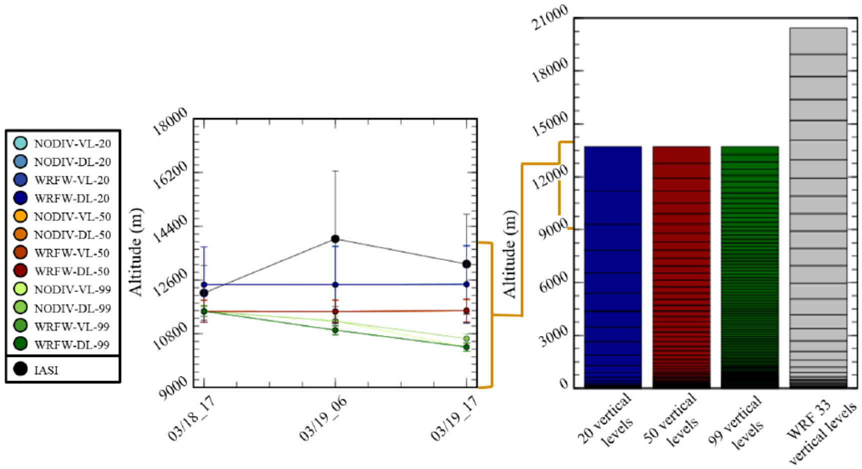

Figure S4. Center : Maximum concentration altitude evolution (IASI and CHIMERE), IASI brackets indicate values' uncertainties and CHIMERE brackets indicates cell's bottom and top. Right : Model vertical levels distribution for the 3 configurations from surface to top.

A sentence has been added to the manuscript :

The WRF model has been run with 33 vertical levels from surface to 55 hPa (28 levels are into 1013-150 hPa range), and with an identical horizontal grid.

P12 L6: 'independant' should be 'independent'

Modification has been done.

P18 L21: Currently this line appears to compare the Després and Lagoutière scheme to itself. Should the second instance actually be "van Leer (1977)"?

Indeed, this has been modified.

P20 L2: Why is increasing vertical resolution only meaningful in cases where plume injection altitude is known? I feel that this statement needs to be better qualified. A reduction in numerical diffusion should always correspond to an improvement in simulation fidelity, even if the initial conditions include error.

We agree with the reviewer that this statement needs to be better qualified. However, we still believe that when increasing accuracy, the probability that the model vertical distribution is totally separated from the real vertical distribution increases. It is true however that, most likely, the qualitative features of the plume including its concentration may be better reproduced in this case even though possible at the wrong location. Therefore, we replaced the question statement by the following which we believe is more precise:

In addition, increasing vertical resolution might give a false appearance of accuracy to the result when plume injection altitude is not known with a good precision.

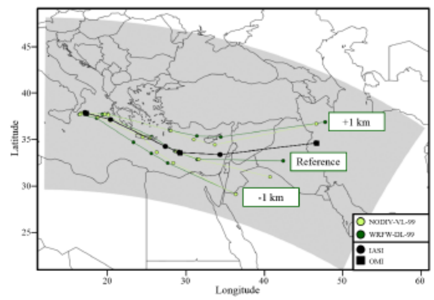

Figure 1. Satellite trajectory of the Etna volcanic plume (black line) built combining information from IASI and OMI instruments. CHIMERE simulated trajectory depending on SO_2 injection altitude of emissions (light and dark green lines - respectively NODIV-VL-99 and WRFW-DL-99). The grey area represents the CHIMERE simulation domain. White triangle indicates Mount Etna location.

1.3 Minor grammatical errors

page 1 line 15, "The CHIMERE CTM has previously been used to assess Eyjajallajökull eruption possible impact on air quality" should be "..to assess the possible impact of the eruption of Eyjajallajökull on air quality").

Phrase formulation has been modified.

I hesitate to bring these up as the errors are almost always very minor and do not impact the science of the paper, and it is usually possible to determine the authors' intended meaning. However, these issues do compromise the readability, and as such I would recommend the authors take another sweep through the paper to correct such issues.

We have performed a thorough checking of grammar and spelling in the manuscript and corrected these slips as best we could.