Object: Response to reviewers' comments on "STEFLUX, a tool for investigating stratospheric intrusions: application to two WMO/GAW global stations" by Davide Putero et al.

Dear Editor,

Please find below our point-to-point replies (bold italic) to the specific comments raised by the two reviewers. We believe all comments have been addressed and we followed all suggested changes. Modifications as respect to the original manuscript are included in the new version uploaded, and marked with red and blue colors in the manuscript version below.
We thank the reviewers for their useful and valuable comments and we hope the manuscript now meets the journal's specific standards for publication.

Yours sincerely,

Davide Putero (on behalf of all the co-authors)

Reviewer #1:

General Comments: This paper presents STEFLUX, a new tool that detects stratospheric intrusions affecting a specific location during a specific time period. STEFLUX is well described and the results are thoroughly discussed revealing the benefits and restrictions of the tool. As the transport of stratospheric air masses into the troposphere is of great importance, STEFLUX can be used in conjunction with observations for several scientific purposes. Therefore, I consider the paper to be an interesting study and recommend its publication in ACP, but only after addressing the following comments.
We thank the Reviewer for his/her valuable suggestions and his/her encouraging evaluation. In the following, we report our point-to-point replies to each of the raised points. Modifications to the text are performed in the revised version of the manuscript and are marked in red and blue colors in the manuscript version below.

Main comments: It seems that there are inconsistencies between the skill scores values (False Alarm Rate, ORSS) presented in the manuscript (Page 7 line 27 – Page 8 line 5, Table 1) and the contingency tables presented in Table 1. Moreover, the presentation of the results in Table 1 needs to be more reader-friendly. I suggest the following:
1. Include (in Section 4.1.2) the formulas used to calculate all skill scores along with the corresponding references, i.e. ORSS=(AD-BC)/(AD+BC) (Thornes and Stephenson, 2001), explaining what A, B, C and D stands for in your case.
2. Assign A, B, C and D to the respective values in Table 1.
3. Check calculations for the skill scores. It is likely that your results are better (higher ORSS values and lower False Alarm Rate values).

4. Add a label for each table in Table 1, in order to be clear which approach is the "reference" and which is the "predictor". i.e for Table 1a,c "SIO vs STEFLUX" and for Table 1b,d "STEFLUX vs SIO".
5. Revise the discussion (for skill scores) in the manuscript if needed.
We thank the Reviewer for this important comment, and we apologize for the inconsistencies between the wrong values given in the text and those reported in Table 1. Our new results are indeed better than those presented in the first version of the manuscript, with a lower false alarm rate (hereinafter called Probability of False Detection, POFD, equal to 0.45), and higher ORSS (see the updated Table 1) for all comparisons. These values have been updated in the text (Page 8, Lines 24-25) and in Table 1.
Moreover, as suggested by the Reviewer, we made the description of the skill scores and the layout of Table 1 clearer. Formulas for the accuracy, false alarm ratio, probability of false detection and ORSS have been inserted (following Thornes and Stephenson, 2001), explaining also what A, B, C and D stand for (see Sect. 4.1.2). Table 1 has been updated by specifying what these capital letters refer to and by adding labels for the "STEFLUX vs SIO" or "SIO vs STEFLUX" comparisons. The caption has also been modified accordingly.

Minor comments: Please add degree symbols for lon and lat values in the manuscript. i.e. Page 3, line 20.
Done.

Section 4.1.1: Please include a definition for STEFLUX SI day. i.e. threshold of at least 1 box crossing per day?
The definition of SI day was erroneously given in Sect. 4.1.2 (Page 7, Lines 10-11): "…SI days (with a threshold of at least 2 box crossings per day for STEFLUX, in order to retain robust information only and to discharge "erratic" events).". Thus, this sentence has been moved above (Page 7, Lines 17-18).

Page 7, lines 6-7: "(see the Supplementary Material)". Please specify exactly where in the Supplementary Material.
Done. Since old Table S1 has been moved into the main body of the text (see our answer to Reviewer #2 comments), the sentence has been modified to: "(see Table S1 in the Supplementary Material)".

Table 2: Add "(b)" in the second table.
Done.

Figure 1a: Please replace "STEFLUX [#]" with "STEFLUX [number of crossings]".
Done.

Figure 4: The map continents are not so clear. Please change map continents color if possible (maybe grey).
Done.

Reviewer #2:

In this study, the authors present a new tool, called STEFLUX, to select trajectories having crossed the tropopause downward at some time in the past days and arriving into a user-defined geographical box within a prescribed time-window. The trajectories are selected among a large set of pre-computed trajectories based on the ERA-Interim reanalysis from the ECMWF. Doing this, this is presumably a fast-computing tool since no trajectory computation is needed.
Output data allow for various applications, such as assessing the occurrence frequency of stratospheric intrusions (SI) in the lower troposphere at any place on Earth at regional scale, but also characterizing preferred entry regions in the UTLS, travel times until the target area, etc. The paper presents an illustrative case study, a skill assessment study with respect to SI detection based on (mainly ground-based) observations, and finally a climatology over 35 years of SI events over two focal areas.
STEFLUX is certainly a promising tool which may be helpful for a scientific community larger than the authors' research team. The paper itself is fairly well-built and written, and the presented scientific material and discussions are of good scientific quality.
Therefore, I recommend the publication of this study in ACP, but not before the author take in consideration the following comments and propose a revised version of their manuscript. I would appreciate if the authors could pay particular attention to my general comments 3 and 4.
We thank the Reviewer for his/her valuable suggestions and his/her encouraging evaluation. In the following, we report our point-to-point replies to each of the raised points. Modifications to the text are performed in the revised version of the manuscript and are marked in red and blue colors in the manuscript version below.

General comments
1. Method originality not fully clear
While reading Section 1, it is not straightforward to know what is new in the STEFLUX method compared to existing methods based on backtrajectories. For instance, one could wonder why don't the author simply initialize backtrajectories from the target regions and see if their cross the tropopause at some time in the past?
I guess one major advantage of the method is computation speed, and this is due to the fact that it works from pre-computed backtrajectories. But this is not clearly stated in the text.
More generally, I think the Introduction could be developed and depict more explicitely the state-of-the-art in the domain: what are the different types approaches? what are their drawbacks or limitations? etc. The originality of the STEFLUX method should thus be more emphasized.
As correctly pointed out by the Reviewer, one characteristic of STEFLUX is the computational speed, since no calculation of back-trajectories is required. This would be a particularly long and time-consuming task, especially when evaluating STE over long periods. The calculation of backward trajectories from the target point would have been possible for NCO-P and Mt. Cimone, but the aim of STEFLUX is to be a tool that can be easily applied at any point on the globe. Therefore, it is valuable to have a set of pre-computed forward trajectories (please refer to Škerlak et al., 2014, to see how the trajectories are generated), because a simple calculation of backward trajectories is not feasible. Also, as demonstrated in the paper, it is sufficient specifying only a few parameters (spatial coordinates and top lid of the box) in STEFLUX, for obtaining a quick and reliable estimate of the STE occurred over the desired time window. In addition to this,

being based on the constantly updated ERA-Interim reanalysis, STEFLUX allows for a STE estimate back to 1979, thus it is especially important from a climatological perspective. In order to highlight these motivations in the text, we modified as follows (Page 3, Lines 7-8): "The tool, called STEFLUX (Stratosphere-to-Troposphere Exchange Flux), is a relatively fast-computing algorithm which makes use of the pre-computed trajectories composing the STE climatology by…" and (Page 3, Lines 12-14): "Its computational speed and user-friendly approach (it is sufficient to specify only a few parameters to work) make it suitable for obtaining a quick and reliable estimate of the SI occurred at a specific place over the desired time window (including long periods which would otherwise require a lot of time-consuming calculations). A potential…".

Moreover, the Introduction part concerning the different "observations-based" methodologies to detect STE has been developed including several works, together with one review paper (Pages 2-3): "Usually, stratospheric influence is detected at a measurement site by analyzing the variability of in situ "stratospheric" observations (e.g., relative humidity, ^{7}Be, ^{10}Be, O_3, atmospheric pressure variability) and profiling datasets (radio/ozone-sondes), coupled with the analysis of satellite (e.g., total column of ozone) and various kinds of numerical weather prediction (NWP) model products fields. Many different methods, as thoroughly reviewed in Stohl et al. (2003), are based on this combined approach. Stohl et al. (2000) deployed a detection algorithm based on the in situ variation of experimental data and simulations with a passive stratospheric tracer. Similarly, other studies analyzed STE by coupling experimental data and back-trajectories (e.g., Cristofanelli et al., 2006, 2010; Trickl et al., 2010). Usually, specific threshold values are applied to in situ tracers' variability to detect the presence of air-masses with stratospheric "fingerprints". Also trajectory and dispersion models are extensively used to detect the occurrence of STE. For example, Cui et al. (2009) used the particle dispersion model FLEXPART (Stohl et al., 2005) and the trajectory model LAGRANTO (Wernli and Davies, 1997) to identify stratospheric transport at the high-altitude Alpine site Jungfraujoch (Switzerland), while Tarasova et al. (2009) deployed 3D air-mass back-trajectories to trace the atmospheric transport at two high mountain measurement sites over the Alps and Caucasus. As pointed out by Bourqui (2006), trajectory-based approaches can provide a lower-bound estimate for STE flux, while dispersion models can provide slightly larger estimates. Typically, when used to detect STE at specific locations at the Earth's surface, all of these "observations-based" methodologies vary among different measurement sites, with respect to the number and types of stratospheric tracers available/considered, threshold values adopted, and often require a lot of time-consuming implementation to work. Moreover, it should be argued that none of the most diffused tracers have a "pure" stratospheric origin; for example, ^{7}Be and O_3 are affected by significant tropospheric sources. Furthermore, the compilation of proper long-term climatologies is very often hindered by the lack of long-term observations of "stratospheric" tracers.".

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmos. Chem. Phys., 5, 2461-2474, 2005.
Tarasova, O. A., Senik, I. A., Sosonkin, M. G., Cui, J., Staehelin, J., and Prévôt, A. S. H.: Surface ozone at the Caucasian site Kislovodsk high mountain station and the Swiss Alpine site Jungfraujoch: data analysis and trends (1990–2006). Atmos. Chem. Phys., 9, 4157-4175, 2009.

2. Representativity of a deep valley station

At several places in the text, it is suggested (e.g. when mentioning the "overpass" effect) that SIO may be missed at the surface stations because their measurement may not always be representative of the free troposphere at regional scale owing to local mountain meteorology. I think this concern is especially true for the NCO-P station, which is located in the bottom of a deep valley. Even in conditions of down-valley flow, it is likely that air has been in contact with the surface before reaching the observatory. Ozone in particular may have experienced deposition, and surface ozone concentrations may be lower than those encountered in the free troposphere. Valleys are indeed known to be net sinks for ozone (see e.g. Furger et al., Atmos. Env., 34, 1395-1412; Wotawa and Kromp-Kolb, Atmos. Env., 34, 1319-1322).

Even in the cited references (Bonasoni et al., 2010; Cristofanelli et al. 2010) little is said on the station representativeness at regional scale (except in the monsoon season at night). It would be worthy if this question could be briefly discussed somewhere in the paper (e.g. in Section 2 when the station are presented).

In contrast, I am much more confident in the regional representativeness of the mountain-top site Monte-Cimone (of course, apart from anabatic conditions) and the suitability of the site to detect deep stratospheric intrusion, although it is at much lower altitude.

The Reviewer raised an interesting point concerning the representativeness of NCO-P ozone observations. Unfortunately, no specific studies aimed at assessing these kinds of processes at NCO-P have been carried out so far. For this reason, in Sect. 4.1.3, the following sentences have been added, as well as two new references (Page 10, Lines 15-19): "Furthermore, for NCO-P, it should be considered that the station is located in a narrow valley. Thus, it is conceivable that, during the transport within the valley, O_3 (one of the stratospheric tracers considered by SIO) experiences deposition phenomena, thus decreasing the actual concentration that the stratospheric air-mass would have in the free troposphere (see, e.g., Furger et al., 2000; Wotawa and Kromp-Kolb, 2000).".

Furger, M., Dommen, J., Graber, W. K., Poggio, L., Prévôt, A. S., Emeis, S., Grell, G., Trickl, T., Gomiscek, B., Neininger, B., and Wotawa, G.: The VOTALP Mesolcina Valley Campaign 1996 – concept, background and some highlights, Atmos. Environ., 34, 1395-1412, 2000.

Wotawa, G., and Kromp-Kolb, H.: The research project VOTALP – general objectives and main results, Atmos. Environ., 34, 1319-1322, 2000.

3. SIO detection criteria too imprecise

In Section S1.4 (supplementary material), the SIO selection criteria are presented in a too vague and qualitative manner (and therefore the criteria appear to be subjective). For instance, what does "significant variation of daily P" mean? What is the threshold to consider the variation is significant? Further, is the current pressure daily mean compared to the value the day before? One could ask such questions for almost every items of the two lists. The authors must present their study in a reproducible way, and those criteria are central elements. This section should be rewritten in a much more rigourous and quantitative manner, with the concern of study reproducibility.

We apologize for having provided too little detail when describing the SIO criteria. The significant variations of the several parameters (TCO, P, ^{7}Be) are computed by the following methodology: first, a three-time repeated iteration of a 21-days running mean (the so-called Kolmogorov-Zurbenko filter, see Sebald et al., 2000) is applied to the daily average time series; then, residuals are computed by subtracting these latest values from the daily averages; residuals

which exceed the upper (for ^{7}Be) or upper and lower (for TCO and P) endpoints of the 95% confidence interval of the residuals distribution over the whole period are thus labelled as "significant variations".
In order to make the SIO selection methodology clearer and available in the main text, Sect. S1.4 has been rewritten including these detailed information and moved to Sect. 2.

Sebald, L., Treffeisen, R., Reimer, E., and Hies, T.: Spectral analysis of air pollutants. Part 2: ozone time series, Atmos. Environ., 34, 3503-3509, 2000.

4. Missing discussion on backtrajectory maximum duration
In this study, tropopause crossings are considered up to 5 (= 1+4) days prior the trajectories reach the target box. But if one goes sufficiently deep backward in time, any trajectory ending in the target box crossed the tropopause at some time in the past. On the contrary, if the trajectory maximum duration is reduced below some value, no SI at all is detected.
Actually, the target region can be found to be from 0% to 100% of the time under the influence of stratospheric intrusions, depending on the chosen trajectory maximum duration. This parameter appears to be of central importance in the STEFLUX tool. I think a sensitivity study to this parameter should be presented (especially in relation with the results (percentages) given in Section 4.2.1), or at least, the choice of 5 days (which obviously comes from the work by Skerlak et al., 2014) should be carefully discussed and justified.
This leads to a more general question: any sufficiently long-lived molecule in the troposphere resided in the stratosphere at some prior times. What is the typical lifetime of a stratospheric intrusion in the troposphere, and when should one consider the air mass composition as being no longer influenced by the stratosphere?
I think these points are crucial in this study and deserve thorough discussions.
The time aspect is always a difficult task in trajectory analyses. The typical lifetime of a stratospheric intrusion in the troposphere has been considered in several papers (e.g., Stohl et al., 2000; Bourqui and Trépanier, 2010; Trickl et al., 2014, 2016). Stohl et al. (2000) argue that, once brought into the troposphere, the stratospheric signature of an air-mass (i.e., low RH, high O_3) gets lost over the period of a few days, because it gets quasi-adiabatically stirred by large-scale cyclonic and anticyclonic disturbances. The choice of 5 days for trajectories is also presented in Bourqui and Trépanier (2010). It was found that the trajectory clusters for their case studies experienced three distinct phases during their descent from the stratosphere (namely crossing of the tropopause, free descent, and quasi-horizontal dispersion in the troposphere), and this whole process takes typically 4-5 days. Similarly, Trickl et al. (2016) define "stratospheric intrusion trajectories" as those initially residing in the stratosphere and descending during the following 5 days by more than 300 hPa into the troposphere. Finally, we could justify our choice with a simple order-of-magnitude calculation. If we suppose that an air parcel descends uniformly from the tropopause (~10 km) to the surface in 5 days, a back-of-the-envelope calculation for the corresponding vertical wind speed leads to 2.3 cm/s, which corresponds to the typical synoptic-scale vertical velocity. Therefore, we look at events which are in line with a "uniform" descent rate consistent with synoptic-scale weather. Obviously, if the descent takes place in a shorter time, the associated vertical wind speed would be higher.
A new sentence has been added to justify the choice of 5 days (Page 10, Lines 3-5): "The maximum value for Δt was chosen according to the typical lifetime values for a stratospheric

intrusion into the troposphere (see Stohl et al., 2000; Bourqui and Trépanier, 2010; Trickl et al., 2014, 2016).".

Bourqui, M. S., and Trépanier, P.-Y.: Descent of deep stratospheric intrusions during the IONS August 2006 campaign, J. Geophys. Res., 115, D18301, 2010.
Trickl, T., Vogelmann, H., Giehl, H., Scheel, H.-E., Sprenger, M., and Stohl, A.: How stratospheric are deep stratospheric intrusions?, Atmos. Chem. Phys., 14, 9941-9961, 2014.
Trickl, T., Vogelmann, H., Fix, A., Schäfler, A., Wirth, M., Calpini, B., Levrat, G., Romanens, G., Apituley, A., Wilson, K. M., Begbie, R., Reichardt, J., Vömel, H., and Sprenger, M.: How stratospheric are deep stratospheric intrusions? LUAMI 2008, Atmos. Chem. Phys., 16, 8791-8815, 2016.

5. Links with ENSO, QBO and sunspots poorly convincing
In Section 4.2.2, the authors claim that some IMFs are correlated with various indicators (of ENSO, QBO, solar activity), but I find that Figure 5 and 6 poorly support these results (at least when examined by eye). Could these correlations be demonstrated more clearly, for instance by means of scatterplots?
Beyond this, correlation is not causality. A correlation is interesting to consider only if one suspects some mechanism linking two quantities. In the text, the possible link between ENSO and STE is discussed, but to a much lesser extent the links with the QBO and the solar activity. Could the authors discuss or even speculate a bit more about this?
According to the Reviewer's comment, we modified the text as follows, to give a deeper indication on the correlations that exist between the parameters: (Page 12, Lines 9-11): "…that is weakly anti-correlated (r = -0.3) to the Quasi-Biennial Oscillation (QBO); the anti-correlation is maximized during post-monsoon and winter seasons (r = -0.5 and r = -0.4, respectively)." And (Page 12, Lines 22-23): "…IMF6 presents some periods of inverse variability with respect to the Multivariate ENSO Index…".
Moreover, the mechanisms for which QBO affects stratospheric circulation (and thus STE) are fully explained in Neu et al. (2014) and references therein. To better clarify this part in the manuscript, a new sentence and references have been added to the text (Page 12, Lines 14-17): "More generally, the mechanisms for which QBO affects the STE variability are both the direct modulation of the circulation through thermal wind balance, and the impact on the strength of the overturning circulation by altering the propagation and dissipation of planetary-scale waves, which enhance the meridional circulation and the cross-tropopause transport (Tung and Yang, 1994; Kinnersley and Tung, 1999; Neu et al., 2014).".
Furthermore, to better characterize the correlation with the solar activity, a sentence has been added, as well as new references (Page 12, Lines 30-32): "Signals of influence of the sunspot cycle in the upper troposphere-lower stratosphere have been indicated in several works (e.g., Labitzke and Van Loon, 1997a, b; Coughlin and Tung, 2004), suggesting that the association between the Sun and stratospheric parameters (e.g., O_3) is due to solar-induced changes in the atmospheric circulation.".

Kinnersley, J. S., and Tung, K. K.: Mechanisms for the Extratropical QBO in Circulation and Ozone, J. Atmos. Sci., 56, 1942-1962, 1999.

Labitzke, K., and Van Loon, H.: The signal of the 11-year sunspot cycle in the upper troposphere-lower stratosphere, Space Sci. Rev., 80, 393-410, 1997a.
Labitzke, K., and Van Loon, H.: Total ozone and the 11-yr sunspot cycle, J. Atmos. Sol.-Terr. Phys., 59, 9-19, 1997b.
Tung, K. K., and Yang, H.: Global QBO in circulation and Ozone. Part II: a simple mechanistic model, J. Atmos. Sci., 51, 2708-2721, 1994.

6. Balance between paper main body and supplementary material

The article main body is quite concise in it present form, and I think there is perhaps room for moving important elements from the supplementary material into the paper main body.
For instance, the criteria to detect SIO are of primary importance in the study and could appear in the article, as well as Table S1, and perhaps also Figures S4 and S5.

We agree with the Reviewer's suggestion of moving relevant part of the Supplementary Material into the main body of the paper. As reported in our response to Reviewer's comment #3, we moved Section S1.4 into Sect. 2. Also Table S1 has been moved into Sect. 4.1, and references to the tables have been updated throughout the text. On the other hand, we decided to keep Figures S4 and S5 into the Supplementary Material.

Specific comments

p.1, l.2: The use of upper-case letters suggests that "STEFLUX" is an acronym. In this case, could the authors make it explicit at least once in the abstract and in the main text body? If it is no acronym but a simple proper noun, I suggest one should write "Steflux".

STEFLUX stands for "Stratosphere-to-Troposphere Exchange Flux", it has now been mentioned in the Abstract and in the Introduction.

p.1, l.19: Please consider to change "relating" by "linking".

This sentence was changed in the Discussion paper version, being: "Furthermore, for the first time, by using the STEFLUX outputs, we investigate the potential impact of specific climate factors (i.e. ENSO, QBO and solar activity) on SI frequency variability over the Mediterranean basin and the Himalayas."

p.1, l.9-10: "show still"→"still show".

Done.

p.2, l.14, "anticyclonic": Do the author mean "cyclonic" instead?

We mean "anticyclonic", i.e., following the downward transport of air-masses already intruded deeply into the lower troposphere.

p.2, l.17, "due to anthropogenic emissions": I would specify: local or regional. Please also consider that local or regional biogenic emissions may also alter atmospheric composition with respect to the tropospheric background.

Done.

p.2, l.17: "make"→"makes".

Done.

p.2, l.27, "Many different methods are based on this combined approach (...) and vary considerably between different measurement sites.": These statements are supported by no literature reference. Could the author cite here a list of references or at least a review paper on the topic? What are those considerable variations between the method? Could the author be a bit more explicit? See also my general comment 1.

Please refer to our answer to Reviewer's comment #1.

p.2, l.27, "occurring over": reaching? detected?

"occurring over" has been changed to "reaching".

p.2, l.33-34, "Moreover, ...": It seems that this potential application is not illustrated in the paper. Could the author justify this statement?

Because the O_3 mixing ratio is one of the parameters that are given at every point along the STE climatology trajectories (see Sect. 3.2), STEFLUX could also be deployed for linking O_3 concentrations deriving from SI to O_3 variations recorded at measurement sites. However, the aim of this paper is to present STEFLUX and compare it to the in situ methodologies deployed at two high-mountain stations, without giving indications on how the SI long-term variability has affected O_3 measurements at those sampling sites. Currently, this other potential application of STEFLUX is under study and will probably be part of a future work. We agree with the Reviewer that this statement could be misleading, if placed in the Introduction, thus we moved it in the Conclusions section, where more appropriate (Page 13, Lines 20-22): "Moreover, although not investigated in this work, STEFLUX might be deployed as a particularly relevant tool to investigate how SI long-term variability influences the atmospheric composition at these specific locations (e.g., by deploying the O_3 values that are available along each trajectory)".

p.3, l.5: "to it"→"on climate".

Done.

p.3, l.19-20: This statement is questionable and deserves further discussion. See my general comment 2.

Please refer to our answer to Reviewer's comment #2.

p.3, l.24-25, "starting at the measurement site": this is too imprecise, especially concerning the altitude. Was the true site altitude or the model surface altitude used to initialize the backtrajectories?

The starting altitude for NCO-P back-trajectories was 490 hPa, as also reported in Sect. S1.3, to minimize possible effects between the model and the real topography. However, in order to avoid misunderstandings, the sentence "starting at the measurement site" has been removed.

p.4, whole Section 3.1: even though the case study clarifies well what STEFLUX is (Sect. 3.2), Section 3.1 presenting the tool is confusing. Especially, it is hard to distinguish what comes from Skerlak et al. and what is specific to STEFLUX. Beyond this, a number of elements from Skerlak et al.'s methodology are mentioned in the text (trajectories extended 4 days prior to tropopause crossing; 3D labeling) but it seems these details are not needed in STEFLUX or at least in this

paper. If really not needed, these information items are confusing and should be removed. Otherwise, it should be explained why they are important. More generally, I think that the whole Section 3.1 should be rewritten and clarified.
According to the Reviewer's suggestion, Sect. 3.1 has been rewritten and clarified. Several details characterizing the input STE trajectories have been removed, not to create too much confusion. Additionally, the list of the output files produced by STEFLUX has been provided (Page 5, Lines 21-24): "STEFLUX produces several output files, which enclose: (i) the trajectory positions and timing found within the box, (ii) the first box crossing positions and timing for each trajectory, (iii) the tropopause crossing position and timing for each trajectory, (iv) the complete list of the trajectories that have crossed the box.".

p.4: title of Section 3.2 could be changed to "Illustrative case study".
Done.

p.4, l.29: The box centered at NCO-P is hardly visible in Fig.1b. Anyway, a reference to this Figure is not useful in this sentence, and mention to Fig.1b could be simply removed here.
Done.

p.4, l.30: "recorded" can be removed.
Done.

p.5, l.5: in the present form of the paper, the criteria are actually introduced in the supplementary material, not in Section 2. See also my general comment 6.
In the revised version of the manuscript, the criteria are fully explained in Sect. 2, see our answer to Reviewer's comments #3 and #6.

p.5, l.11 and ff.: it seems from these lines that there are three different output files from a STEFLUX run, but it is not clear what is in those files. This should be clarified (perhaps in Section 3.1).
Please refer to our answer above concerning Sect. 3.1.

p.5, l.17: "indicated in previous studies ..."→"identified as a preferred region for tropopause crossing in previous studies ...".
Done.

p.5, l.26, "they still maintained a stratospheric signature": poor expression, please rephrase.
The sentence has been changed to: "they still followed the stratospheric circulation steered by the subtropical jet stream".

p.5, l.33: the choice of an horizontal extension of 3°×3° should be justified briefly.
The choice of a 3°x3° horizontal extension was made after performing some sensitivity tests on this parameter. The chosen extension was the one presenting the best agreement with the SIO time series. However, it has to be noted that this parameter can be completely chosen by the user, adapting it to the very different situations (e.g., topography, surrounding regions) of the area

under study. The text has been modified as follows (Page 6, Line 32): "…site, after performing a sensitivity test on this parameter (not shown).".

p.6, l.2-3, "The selected time periods were the same as in Sect.2": please specify.
Done, a new sentence has been added (Page 7, Lines 1-2): "(i.e., March 2006–December 2013 for NCO-P and January 1998–December 2010 for Mt. Cimone)".

p.6, l.4: "a table listing ..."→"Table S1 listing ...".
This sentence has been modified to "Table 1 listing", since Table S1 has been moved into the main body of the text.

p.6, Section 4.1.1: What is the criterion to tag a day as SI day according to STEFLUX? Is only one box crossing at any moment of the day and of any duration needed? The author should specify this in this Section. (See also the corresponding comment from the Anonymous Referee #1.)
The criterion to tag a day as SI day is the threshold of at least two box crossings per day, independently on the time. Please see also our comment to Reviewer #1, since we have moved a sentence from Sect. 4.1.2 to Sect. 4.1.1.

p.6, l.11, "at the two measurement sites": not needed and a bit confusing, please remove.
Done.

p.6, l.24: "subtle" is unexpected as adjective for the inter-annual variability. Please rephrase.
The sentence has been rewritten (Page 7, Line 23): "Although the seasonality was a feature well captured by STEFLUX, the inter-annual variability was less clearly identifiable.".

p.7, l.1, "criteria coverage": please define. Is it the fraction of time when the data used in the criteria are simultaneously available? Every criterion does not use all the data: what does happen when one data is missing for one criterion but another criterion is fulfilled? Or none other fulfilled? Is the day tagged as SI/non-SI day or discarded? Please clarify.
As specified in the former Sect S1.4 in the Supplementary Material, and in Sect. 2 of the revised version, a day is selected as "SI day" if at least one criterion is fulfilled. Thus, a specific day can be simultaneously selected by different criteria, but the simultaneity is not strictly required for tagging the day as "SI day". The "criteria coverage" displayed in Fig. 2 is defined as the seasonally averaged percentage of available data from each criterion. The sentence in Page 7, Lines 9-11 has been rewritten to clarify this aspect: "Additionally, the seasonally averaged percentage of available data from each criterion (hereinafter referred to as "criteria coverage") is also reported in the plot (grey bars).".

p.7, l.22 and ff.: I had a hard time to understand those contingency tables. Considering for instance Table 1(a), does 55 means that during 55 SIO events, STEFLUX detected more than 50% of time of the episode as SI? Does 148 means that during 148 SIO events, STEFLUX detected less than 50% of time of the episode as SI? etc. Please explain a bit more how those numbers should be interpreted. See also the comment from the Anonymous Referee #1 concerning the definitions of accuracy and false alarm rate: how exactly are the presented scores calculated?

The contingency tables and the related description have been made clearer. We inserted formulas concerning the skill scores we presented, and we gave a description of each parameter composing Table 1. Please also refer to our comments to Reviewer #1.

p.8, l.2 and 5: the capture rates given in Table 2 (22-27%) are closer to one quarter than to one third.
Corrected.

p.8, l.20-24: In case of long travel time and high mixing, can one still consider the air mass as a stratospheric intrusion? See my general comment 4 on stratospheric intrusion lifetime.
Please refer to our answer to Reviewer's comment #4.

p.9, l.10: this again is related to my general comment 4: is it really relevant to be irrespective of the degree of mixing and dilution in the troposphere?
Please refer to our answer to Reviewer's comment #4.

p.9, l.23-25: could the author explain this statement?
This sentence has been rewritten as: "First, the location of the crossing is useful to determine the O_3 concentration of the air parcels at the start of their tropospheric path towards the target region.".

p.9, l.32: "If divided seasonally"→"Considering seasons separately"
Done.

p.9, l.33 and p.12 l.15: "southward of"→"south of"
Done.

p.11, l.12 "does not exhibit as"→"exhibits no"
Done.

p.11, l.18 "defined"→"user-defined"
Done.

p.11, l.19 "representative"→"illustrative"
Done.

p.12, l.17 "both of the"→"both"; "significant"→"statistically significant"
Done.

p.13, Appenzeller and Davies, 1992: insufficient reference.
Corrected.

p.16, Table 2: missing "(b)".
Corrected.

p.18, figure legend, l.3: ”Sect. 2“→”Sect. S1.4“. See my general comment 6.
Since the SIO criteria are now fully introduced in Sect. 2, we have not modified this caption.

p.19, figure 3: the STEFLUX and SIO panel columns could be interchanged, so that the panels (a-d) are numbered in the same order as in the text. Why do the box plots in the upper panel have no whiskers?
Figure 3 has been redrawn, with the SIO and STEFLUX columns interchanged, as they appear in the text (thus modified accordingly). The box plots for NCO-P have no whiskers (representing the 10^{th} and 90^{th} percentiles), because too few data were available for their calculation.

Supplementary material
p.1, l.18: do the authors mean gamma-spectroscopy?
Yes, corrected.

p.1, l.25: ”total column OF ozone“.
Done.

p.2, l.9-10: This sentence is not fully clear. What does ”centered at an ending altitude“ mean? Is 490hPa the real altitude of NCO-P? In the same vein in l.12, do the trajectories reach Mt. Cimone at its real altitude level? What is the corresponding pressure level?
The sentence has been changed to “starting at”. The choice of using 490 hPa, higher than the real altitude of NCO-P (5079 m a.s.l., or average pressure of 550 hPa), was made for minimizing possible effects between the model and the real topography. Similarly, back-trajectories at Mt. Cimone have been started at 2200 m. The text in the Supplementary Material has been changed accordingly.

p.2, Section S1.4 (SI selection criteria): see my general comments 3 and 6.
Section S1.4 has been integrated in Sect. 2.

STEFLUX, a tool for investigating stratospheric intrusions: application to two WMO/GAW global stations

Davide Putero[1,2], Paolo Cristofanelli[1], Michael Sprenger[2], Bojan Škerlak[2], Laura Tositti[3], and Paolo Bonasoni[1]

[1]CNR–ISAC, National Research Council of Italy – Institute of Atmospheric Sciences and Climate, via Gobetti 101, 40129, Bologna, Italy
[2]IAC–ETH, Institute for Atmospheric and Climate Science – ETH Zurich, Universitätstrasse 16, 8092, Zurich, Switzerland
[3]Dept. Of Chemistry "G. Ciamician", Alma Mater Studiorum University of Bologna, Via Selmi 2, 40126, Bologna, Italy

Correspondence to: D. Putero (d.putero@isac.cnr.it)

Abstract. Stratospheric intrusions (SI) are a topic of ongoing research, especially because of their ability to change the oxidation capacity of the troposphere and their contribution to tropospheric ozone levels. In this work, a novel tool called STEFLUX (Stratosphere-to-Troposphere Exchange Flux) is presented, discussed and used to provide a first long-term investigation of SI over two global hot-spot regions for climate change and air pollution: the southern Himalayas and the central Mediterranean basin. The main purpose of STEFLUX is to obtain a fast-computing and reliable identification of the SI occurring at a specific location and during a specified time window. It relies on a compiled stratosphere-to-troposphere exchange (STE) climatology, which makes use of the ERA-Interim reanalysis dataset from the ECMWF, as well as a refined version of a well-established Lagrangian methodology. STEFLUX results are hereby compared to the SI observations (SIO) at two high-mountain WMO/GAW global stations in these climate hot-spots, i.e., the Nepal Climate Observatory-Pyramid (NCO-P, 5079 m a.s.l.) and Mt. Cimone (2165 m a.s.l.), which are often affected by SI events. Compared to the observational datasets at the two specific measurement sites, STEFLUX is able to detect SI on a regional scale. Furthermore, it has the advantage of retaining additional information concerning the pathway of stratospheric-affected air-masses, such as the location of tropopause crossing and other meteorological parameters along the trajectories. However, STEFLUX neglects mixing and dilution that air-masses undergo along their transport within the troposphere. Therefore, the regional-scale STEFLUX events cannot be expected to perfectly reproduce the point measurements at NCO-P and Mt. Cimone, which are also affected by small-scale (orographic) circulations. Still, the SI seasonal variability according to SIO and STEFLUX agree fairly well. By exploiting the fact that the ERA-Interim reanalysis extends back to 1979, the long-term climatology of SI at NCO-P and Mt. Cimone is also assessed in this work. The analysis of the 35-year record at both stations denies the existence of any significant trend in the SI frequency, except for winter seasons at NCO-P. Furthermore, for the first time, by using the STEFLUX outputs, we investigate the potential impact of specific climate factors (i.e. ENSO, QBO and solar activity) on SI frequency variability over the Mediterranean basin and the Himalayas.

1 Introduction

Stratosphere-to-troposphere exchange (STE) represents one of the natural processes that have substantial impacts on meteorology and atmospheric chemistry, and is an important aspect of climate change (Appenzeller and Davies, 1992; Holton et al., 1995; Stohl et al., 2003; Stevenson et al., 2006). The definition of STE encompasses a two-way air-mass transport: the downward transport from the stratosphere to the troposphere (STT) and the upward transport from the troposphere to the stratosphere (TST). A specific type of STT is called stratospheric intrusion (SI), which we hereby define as the downward transport of stratospheric air-masses relatively deep into the troposphere (as done in Cristofanelli et al., 2006). SI are capable of changing the oxidation capacity of the troposphere (Gauss et al., 2003) and their contribution to the ozone (O_3) levels in the troposphere has been estimated to be as large as the net photochemical production (Roelofs et al., 1997), although models ~~show still~~ still show large uncertainties in the estimates (e.g., Stevenson et al., 2006; Young et al., 2013). As pointed out in many studies (e.g., Reed, 1955; Appenzeller and Davies, 1992; Lamarque and Hess, 1994; Holton et al., 1995; Appenzeller et al., 1996; Stohl et al., 2003; Cooper et al., 2005; Sprenger et al., 2007) SI can be caused by different mechanisms and are typically associated to distinct synoptic- and meso-scale features: tropopause folds and cutoff lows, subtropical jet streams and streaks, potential vorticity (PV) streamers, upper-level fronts and anticyclonic areas.

High mountain stations are appropriate sites for investigating the transport of stratospheric air-masses into the troposphere, because stratospheric air-masses can already be identified at mid-tropospheric levels. Furthermore, they are less influenced by polluted air-masses due to local or regional anthropogenic emissions (Stohl et al., 2000), which ~~make~~ makes the SI detection more straightforward. Several studies have been carried out in the past, to assess the influence of SI at high-altitude remote sites, which also represent ideal locations for studying the background conditions of the troposphere (e.g., Stohl et al., 2000; Cristofanelli et al., 2006; Ordóñez et al., 2007; Cristofanelli et al., 2010; Trickl et al., 2010; Lin et al., 2012). Usually, stratospheric influence is detected at a measurement site by analyzing the variability of in situ "stratospheric" ~~tracers,~~ observations (e.g., relative humidity, ^{7}Be, ^{10}Be, O_3, atmospheric pressure variability) and profiling datasets (radio/ozone-sondes), coupled with the analysis of satellite (e.g., total column of ozone) and various kinds of numerical weather prediction (NWP) model products fields~~(i. e., back-trajectories, or passive stratospheric tracers implemented in high resolution models).~~ . Many different methods, as thoroughly reviewed in Stohl et al. (2003), are based on this combined approach. ~~Typically, they differ~~ Stohl et al. (2000) deployed a detection algorithm based on the in situ variation of experimental data and simulations with a passive stratospheric tracer. Similarly, other studies analyzed STE by coupling experimental data and back-trajectories (e.g., Cristofanelli et al., 2006, 2010; Trickl et al., 2010). Usually, specific threshold values are applied to in situ tracers' variability to detect the presence of air-masses with stratospheric "fingerprints". Also trajectory and dispersion models are extensively used to detect the occurrence of STE. For example, Cui et al. (2009) used the particle dispersion model FLEXPART (Stohl et al., 2005) and the trajectory model LAGRANTO (Wernli and Davies, 1997) to identify stratospheric transport at the high-altitude Alpine site Jungfraujoch (Switzerland), while Tarasova et al. (2009) deployed 3D air-mass back-trajectories to trace the atmospheric transport at two high mountain measurement sites over the Alps and Caucasus. As pointed out by Bourqui (2006), trajectory-based approaches can provide a lower-bound estimate for STE flux, while dispersion models can provide slightly larger estimates. Typically, when used to detect STE at specific locations at the Earth's surface, all of these "observations-based" methodologies vary among different measurement sites, with respect to the number and types of stratospheric tracers ~~observed, and vary considerably between different measurement sites.~~ available/considered, threshold values adopted, and often require a lot of time-consuming implementation to work. Moreover, it should be argued that none of the most diffused tracers have a "pure" stratospheric origin; for example, ^{7}Be and O_3 are affected by significant tropospheric sources. Furthermore, the compilation of proper long-term climatologies is very often hindered by the lack of long-term observations of "stratospheric" tracers.

In this work we present a novel tool, which aims at objectively identifying SI ~~occurring over~~ reaching a "target" geographical region and during a specific time window. The tool, called STEFLUX (Stratosphere-to-Troposphere Exchange Flux), is a relatively fast-computing algorithm which makes use of the pre-computed trajectories composing the STE climatology by Škerlak et al. (2014). This climatology is available from 1979 and continuously updated. The Lagrangian approach, on which it is based, has been extensively used in previous studies (e.g., Wernli and Bourqui, 2002; Sprenger and Wernli, 2003; Bourqui, 2006; Sprenger et al., 2007; Škerlak et al., 2014), and has been confirmed to effectively identify SI events and to reproduce several of their related aspects. ~~Another~~ Its computational speed and user-friendly approach (it is sufficient to specify only a few parameters to work) make it suitable for obtaining a quick and reliable estimate of the SI occurred at a specific place over the desired time window (including long periods which would otherwise require a lot of time-consuming calculations). A potential use of STEFLUX is to identify the SI occurrence in locations where a detection based on observational data is not available. ~~Moreover, it might be deployed as a particularly relevant tool to investigate how SI long-term variability influences the atmospheric composition.~~ To evaluate the STEFLUX skills in identifying the SI events, we hereby compare its outputs with the SI identification based on observations at two high-altitude World Meteorological Organization/Global Atmosphere Watch (WMO/GAW) global stations in Asia (Nepal) and Europe (Italy). Then, we use STEFLUX to provide a first investigation on the long-term (i.e., 1979–2013) variability of SI occurrence at these measurement sites, that are indeed representative of the lower troposphere of two hot-spot regions for climate change and anthropogenic impacts ~~to it~~ on climate (Monks et al., 2009): the central Mediterranean basin and the southern Himalayas. In particular, we provide a first assessment of possible impact of large-scale climate processess (i.e. ENSO, QBO, solar activity) in modulating the long-term SI variability.

The paper is structured as follows: in Sect. 2 we define the selection methodologies that were used for identifying SI events at the two measurement sites; in Sect. 3 we describe in detail the STEFLUX tool, along with a case study to show a potential application. STEFLUX time series are then compared to the in situ measurements in Sect. 4, followed by a critical discussion about the benefits and restrictions of the tool. Furthermore, trends and periodicities of the long-term SI time series at NCO-P and Mt. Cimone are assessed. Finally, Sect. 5 summarizes the main results of the study.

2 Experimental datasets

Datasets of daily SI occurrences are available at two high-altitude WMO/GAW global stations, i.e., the Nepal Climate Observatory-Pyramid (NCO-P, 5079 m a.s.l., Nepal) and Mt. Cimone (2165 m a.s.l., Italy) since 2006 and 1998, respectively. In this section, a brief description of the two measurement sites is provided, together with the description of the methodology used to detect SI events based on the analysis of in situ stratospheric tracers' variability (coupled with additional model data). Hereinafter, these datasets will be referred to as "Stratospheric Intrusions Observations (SIO)". ~~Extended technical~~ Technical details on the ~~two methodologies applied~~ different parameters considered are given in the Supplementary Material and in the papers by Cristofanelli et al. (2006, 2010).

NCO-P (27.95° N, 86.82° E) is located in the southern Himalayas, near the base camp of Mt. Everest, in the Khumbu Valley, Nepal. This station is far away from anthropogenic sources, thus it can be considered representative of the background conditions of the high Himalayas and the free troposphere (especially during night-time). Further details on the measurement site and on the instrumental setup are given in Cristofanelli et al. (2010). To account for days likely affected by SI events at NCO-P, a specifically designed statistical methodology was applied to the time series of observed and modeled variables. The parameters used consisted of in situ measurements (O_3, atmospheric pressure – P and relative humidity – RH), satellite observations (total column of O_3 – TCO, as retrieved by the OMI – Ozone Monitoring Instrument) and NWP-based back-trajectories ~~(starting at the measurement site, by using LAGRANTO, see Sprenger and Wernli, 2015; Wernli and Davies, 1997).~~ (by using LAGRANTO, see Sprenger and Wernli, 2015; Wernli and Davies, 1997). The methodology is composed of four different criteria; at least one must be satisfied to identify a day as likely influenced by SI:

1. significant variations of daily P values and presence of back-trajectories with values of PV > 1.6 pvu;

2. significant daily TCO increases and presence of back-trajectories with values of PV > 1.6 pvu;

3. significant variations of daily P values and significant TCO daily increases;

4. presence of RH values lower than 60% and significant negative correlation O3-RH and daily O_3 maximum higher than the seasonal value and significant variation of daily P, PV or TCO values (this last criterion was introduced for taking into account the possible role of downward valley winds in transporting air-masses from aloft).

The significant variations are obtained as follows: first, a three-time repeated iteration of a 21-days running mean (the so-called Kolmogorov-Zurbenko filter, see Sebald et al., 2000) is applied to the daily average time series, and residuals are calculated by subtracting these values from the daily averages; then, it is checked whether residuals exceed the upper or lower endpoints of the 95% confidence interval of the residuals distribution over the whole period. For this work, the period of study considered for NCO-P spans from March 2006 to December 2013.

Mt. Cimone (44.19° N, 10.70° E) is the highest peak of the Italian northern Apennines. The observations carried out at this sampling site can be considered representative of the free tropospheric conditions for most of the year, while during warm periods the station can be affected by thermal and convective transport of planetary boundary layer (PBL) air. Other details about Mt. Cimone and the instrumental setup can be found in Cristofanelli et al. (2015) and references therein. ~~The methodology used to investigate the influence of SI at Mt. Cimone is based on a statistical method similar to that applied to~~ Similarly to NCO-P, ~~and encloses the variability of in situ , satellite and modeled variables. In situ~~ in situ measurements of Beryllium-7 (^7Be) and RH have been considered, as well as satellite observations (~~total column O_3~~TCO, as deduced by

TOMS–Total Ozone Mapping Spectrometer–and OMI overpass data) and PV of air-masses reaching the sampling site (by the analysis of 7-day FLEXTRA back-trajectories, Stohl et al., 1995). The statistical method for identifying SI days is based on the following four criteria:

1. significant daily TCO increases and presence of back-trajectories with values of PV > 1.6 pvu;
2. significant daily ^{7}Be increases and presence of back-trajectories with values of PV > 1.6 pvu;
3. presence of RH values lower than 40% and presence of back-trajectories with values of PV > 1.6 pvu;
4. presence of RH values lower than 40% and significant TCO daily value increases.

Again, the significant variations are defined in the same way as done for NCO-P, and at least one criterion must be valid for tagging the selected day as influenced by SI. The period of study for Mt. Cimone spans from January 1998 to December 2010.

SIO at these two measurement sites provide a unique opportunity to test the capacity of STEFLUX in reproducing the main features of SI occurrences (frequency, seasonality, long-term variations) at two locations representative of the Northern Hemisphere midlatitudes and subtropics.

3 The STEFLUX tool

3.1 Description of the tool

The main purpose of STEFLUX is to obtain a fast-computing and reliable estimation of SI occurring at a specific location. The database used as input relies on the trajectories from the STE climatology presented in Škerlak et al. (2014), which makes use of the ERA-Interim reanalysis dataset from the ECMWF (Dee et al., 2011), as well as a refined version of a well-established Lagrangian methodology (Wernli and Bourqui, 2002), to calculate mass and ozone fluxes across the tropopause and several pressure surfaces. ~~Basically, from a large set of global trajectories available each day, only the ones that cross the tropopause (defined as the 2 pvu/380 K surface) within the first 24 hours are retained. These are then extended backward and forward in time for additional 4 days, creating 9-day long STE trajectories. Furthermore, the application of a 3-D labelling algorithm is used to distinguish points of stratospheric and tropospheric nature, as well as potential low-level PV anomalies due to friction (e.g., near mountains).~~ For further details, please refer to Škerlak et al. (2014).

~~Based on this STE climatology,~~ STEFLUX detects the air parcels originating in the stratosphere and entering a tropospheric 3D target box during a specific time window. For this reason, several parameters need to be defined for the STEFLUX tool to work: (i) the time period for which the analysis should be carried out, and (ii) the ~~geographical region of interest, i.e., a target box must be specified~~ target box by means of its longitude and latitude boundaries and by its vertical extension from the surface up to the top boundary (defined as a pressure level, in hPa). On request, the PBL height can be used as top boundary of the target box: this option takes into account the ERA-Interim PBL height, which is a parameter also available along each STE trajectory. ~~This option differs from the "deep STT" events from the STE climatology (Škerlak et al., 2014), because it also considers the~~

~~trajectories that are advected sideways into the target box.~~ In addition to this, another optional parameter allows the temporal resolution for the STE trajectories to be increased from its default value (6 h) up to 1 h. STEFLUX produces several output files, which enclose: (i) the trajectory positions and timing found within the box, (ii) the first box crossing positions and timing for each trajectory, (iii) the tropopause crossing position and timing for each trajectory, (iv) the complete list of the trajectories that have crossed the box.

3.2 ~~Case~~ Illustrative case study

To present an application of STEFLUX and its output files, a SI case in January 2007 is discussed. The period of study coincides with a case study discussed in Bracci et al. (2012), where the SI event is strongly related to the subtropical jet sream. We defined a box with horizontal extension 85–88° E and 26–29° N, centered at NCO-P~~(see Fig. 1b).~~. The top boundary of the box was taken at 550 hPa (the average ~~recorded~~ pressure at the station)~~and we then ran~~, while the STEFLUX tool was run for the time period 9–25 January 2007. All STE trajectories from the climatology by Škerlak et al. (2014), introduced in Sect. 3.1, were analyzed and their crossings of the target box boundaries were determined.

Figure 1a shows as a time series the number of the daily crossings (derived from the STE trajectories) in the box, according to STEFLUX. Additionally, the daily averaged values for O_3 and RH at NCO-P are shown, and each day is also marked as selected or not by the SIO methodology, according to the criteria introduced in Sect. 2. The study period was characterized by the presence of clean and dry (daily RH average always below 40%) air-masses; moreover, the double-jet structure of wind speed at 250 hPa (contour lines in Fig. 1b) indicated the presence of the subtropical jet stream over South Asia and the Himalayas. SIO methodology identified a likely SI event, spanning from 13 to 17 January 2007 (the missing record of O_3 during 15 January was responsible for the gap in the SIO time series). This was confirmed by STEFLUX: from 12 to 24 January 2007, several STE trajectories passed through the target box; at its peak, 25 daily crossings were counted. This time series is directly based on two STEFLUX output files, which list the first and the entire trajectory positions found within the target box. In addition to the crossing time, the list of variables includes the position (longitude/latitude/pressure) and several meteorological parameters (e.g., potential temperature, specific and relative humidity, PBL height) at each point within the target box.

The first positions of the air parcels after entering the target box are marked in Fig. 1b as blue dots; additionally, the positions where the air parcels crossed the dynamical tropopause are shown as green dots. Interestingly, these crossings are mainly clustered into a region over North Africa, which is ~~indicated~~ identified as a preferred region for tropopause crossing in previous studies (Sprenger and Wernli, 2003; Škerlak et al., 2014), and also in this paper (see Sect. 4.2.1)~~as a preferred region for tropopause crossing~~. The exact times and positions of the tropopause crossings are saved by STEFLUX in a third output file, together with atmospheric pressure, potential temperature and O_3 concentration. The output files from STEFLUX allow the history of the STE air parcels to be studied along their way from the stratosphere to the target box. As an example, Fig. 1b reports all the trajectories from the tropopause crossings (green dots) to the target box (blue dots). These trajectories are colored according to their PV value, i.e., points below 2 pvu (magenta) and points greater than 2 pvu (black). Fig. 1c shows the time-height evolution of the trajectories, where time is given relative to the arrival time in the target box. Additionally, the top boundary height of the target box (550 hPa) is reported in the figure (red horizontal line). It is discernible that most of air parcels slowly descended until 72–48 h before they reached the box. Previously, they were characterized by PV values above 2 pvu (black line), i.e., they still ~~maintained a stratospheric signature, and followed the shape of~~ followed the stratospheric circulation steered by the subtropical jet stream (as also discernible from the contour lines of wind speed at 250 hPa in Fig. 1b, which mark the double-jet structure). A rapid descent then set in before their arrival; hence, the PV falls below 2 pvu, indicating that the air parcels crossed the dynamical tropopause.

4 Results

4.1 STEFLUX vs SIO

In this section, the SI occurrences from STEFLUX are compared to the ones from SIO at the two WMO/GAW global stations (see Sect. 2). For both stations, STEFLUX was run by setting a target box with a horizontal extension of $3° \times 3°$ around the measurement site, after performing a sensitivity test on this parameter (not shown). Vertically, the box extended up to 550 hPa for NCO-P and 790 hPa for Mt. Cimone, respectively, corresponding to the average pressure level recorded at each station throughout the year. The selected time periods were the same as in Sect. 2 ~~,~~ (i.e., March 2006–December 2013 for NCO-P and January 1998–December 2010 for Mt. Cimone), but the temporal resolution for the STE trajectories was increased to 1 hour (see ~~the Supplementary Material for a table~~ Table 1 listing all of the input parameters).

The aim of this part is twofold: first, we would like to see how STEFLUX compares to SIO (in Sect. 4.1.1 and 4.1.2). This comparison, which turns out to be not a perfect match, will lead to a critical discussion of what can and cannot be expected from STEFLUX. Hence, it is paramount to understand that both STEFLUX and SIO have complementary strong and weak points in identifying SI and thus may not be exactly compared one-to-one (as discussed in Sect. 4.1.3).

4.1.1 Seasonal comparison and inter-annual variability

The seasonal frequency (in %) of SI days~~within each season~~, derived from measurements (SIO), is presented in Fig. 2 as a red line. Additionally, the ~~seasonal percentage of days with available data at the two measurement sites~~ seasonally averaged percentage of available data from each criterion (hereinafter referred to as “criteria coverage”) is also reported in the plot (grey bars). Note that the season definition slightly differs for the two sites: at NCO-P it consists of a dry (winter – DJF), a wet (monsoon – JJAS) and two transition seasons (pre-monsoon – MAM and post-monsoon – ON) (see Bonasoni et al., 2010), while for Mt. Cimone the “classic” Northern Hemispheric definition is chosen (winter – DJF, spring – MAM, summer – JJA, and autumn – SON). A clear seasonality characterized the SIO frequency at both stations, as highlighted by the seasonally averaged SI frequencies obtained. For NCO-P (Fig. 3~~c~~a), a maximum was discernible in winter and a minimum during the monsoon season, while for Mt. Cimone (Fig. 3~~d~~b) high SI values were found in winter and spring and a minimum in summer. We computed the same time series of seasonal frequencies using STEFLUX (blue lines in Fig. 2), with a threshold of at least 2 box crossings per day, in order to retain robust information only and to discharge “erratic” events. It showed a clear average seasonality at both stations (Fig. 3~~a,b~~c,d), comparable and consistent with that from SIO, especially for NCO-P. At Mt. Cimone, the average annual variation for STEFLUX was more pronounced than the one derived from SIO. This was due to an overestimation of the STEFLUX average frequency for November–January and to an underestimation for June–July. However, the observed seasonality is in line with previous works (e.g., Trickl et al., 2010; Škerlak et al., 2014).

Although the seasonality was a feature well captured by STEFLUX, the ~~representation of the~~ inter-annual variability ~~turned out to be much more subtle~~was less clearly identifiable. Concerning NCO-P (Fig. 2a), the correlation between the two seasonal time series was rather high (Pearson's $r = 0.7$), but for specific years STEFLUX and SIO results evidently differed in the amplitude and timing of the annual peaks. In particular, SIO showed much higher SI frequency than STEFLUX for post-monsoons during 2010–2012. When comparing seasons individually, the correlation coefficient was satisfying for pre-monsoon, monsoon and post-monsoon ($r = 0.5$ on average). For winter, the two time series are even anti-correlated ($r = -0.4$). This can be attributed to a significant decrease of SI detection by SIO during winter for the years 2009–2010. Moving to Mt. Cimone (Fig. 2b), the correlation between the SIO and STEFLUX time series was still high ($r = 0.7$). Also the individual comparison of the seasons gave satisfying results (r varied between 0.4 and 0.5). It has to be noted the evident decrease in SI detection coverage at Mt. Cimone during the period 2006–2011, related to a lower availability of ^{7}Be observations at this sampling site (measurements were stopped in 2012, see Tositti et al., 2014). To investigate whether the low SIO coverage hindered the comparison with STEFLUX, the same analysis was also performed by limiting the Mt. Cimone dataset to the period 1998–2004, i.e., when the criteria coverage was greater than 90% for all of the seasons. However, the results did not significantly differ (see Table S1 in the Supplementary Material).

4.1.2 Event-based comparison

In this section, we extend the comparison to a higher temporal resolution, i.e., by considering single SI events. More specifically, in this study, a SI event was defined (for both STEFLUX and SIO) as the aggregation of contiguous SI days~~(with a threshold of at least 2 box crossings per day for STEFLUX, in order to retain robust information only and to discharge "erratic" events)~~. Furthermore, cases in which two distinct SI events were separated by a single no-SI day were treated like a single event covering the entire period. Generally, an event-based comparison between modelled and observed SI events and experimental detection is a very challenging task, as pointed out by previous investigations, concerning the transport and mixing of stratospheric air deep into the troposphere (e.g., Meloen et al., 2003; Cui et al., 2009; Bracci et al., 2012).

For NCO-P, based upon the SIO criteria, a total of 203 SI events (361 days influenced by SI, representing 13% of the period) were identified, with duration ranging from 1 to 14 days, and average length of 1.9 days. On the other hand, STEFLUX identified 155 SI events (376 days, 13%), with duration ranging from 1 to 10 days (average length: 2.6 days). At Mt. Cimone, 299 SI events (433 days, 9%) were identified by the SIO methodology (with duration ranging from 1 to 8 days, and average length of 1.6 days), while STEFLUX yielded 237 SI events (491 days, 10%) that lasted from 1 to 10 days (with an average length of 2.2 days).

To assess the STEFLUX performance, the approach presented in Cui et al. (2009) was followed. First, all SI events as retrieved by SIO were considered, and then it was checked whether at least 50% of the duration of each SIO event was confirmed by STEFLUX. If this was the case, STEFLUX was considered able to capture the selected SIO event. Hereinafter, we will refer to this comparison as "SIO vs STEFLUX". Vice-versa, the "STEFLUX vs SIO" comparison checked if a SI event (as defined by STEFLUX) was confirmed by the SIO dataset.

As an overview of the results of this comparison, we computed contingency tables (Table 2). In these 2×2 tables, each entry encloses a list of SI or no-SI events, as defined by the considered methodology (STEFLUX and SIO). From the contingency tables it is possible to evaluate several skill scores, which are useful to measure the skill of one method in identifying SI events compared with the other one. The accuracy (ACC), false alarm ratio (FAR), and probability of false detection (POFD) skill scores are defined, according to Thornes and Stephenson (2001) and Wilks (2006), as:

$$ACC = \frac{A+D}{A+B+C+D} \tag{1}$$

$$FAR = \frac{B}{A+B} \tag{2}$$

$$POFD = \frac{B}{B+D} \tag{3}$$

where, for each contingency table, A represents the number of SI events selected by both methodologies (STEFLUX and SIO); B represents the number of events selected as SI by the first methodology but as no-SI by the second one; C represents the number of events selected as no-SI by the first methodology but as SI by the second one; and D represents the number of no-SI events selected by both methodologies. All four contingency tables ~~present~~ give identical values of ~~accuracy (0.6) and a false alarm rate between 0.7 and 0.8. An additional important parameter , also given in Table 2, is the Odds Ratio Skill Score (ORSS, see Thornes and Stephenson, 2001).~~ACC (0.58) and POFD (0.45), while FAR varies between 0.73 and 0.78. The rather high FAR values and low POFD values can be partially explained by considering that the occurrence of SI is a relatively "unlikely" event with respect to the occurrence of no-SI. Also for taking into account this point, we considered an additional parameter (i.e., the Odds Ratio Skill Score, ORSS, see Thornes and Stephenson, 2001), which is not influenced by the marginal totals (i.e., A+C and B+D). This parameter is defined as:

$$ORSS = \frac{A \times D - B \times C}{A \times D + B \times C} \tag{4}$$

The ORSS varies between -1 and +1, where a score of 1 represents perfect skill and a score of 0 indicates no skill; negative values imply that values of one series are opposite to what observed by the other one. Also reported in each table is the minimum ORSS required to have real skill at the 99% confidence level (see Thornes and Stephenson, 2001). All of our combinations indicate that the agreement between the methodologies is not due to chance (i.e., is statistically significant). Mt. Cimone showed higher scores than NCO-P. This can be explained by the location of NCO-P: it is placed at the bottom of a narrow valley (see Bonasoni et al., 2010) and therefore subgrid-scale processes (e.g., PBL entrainment and thermally driven valley winds, not reproduced by the trajectories analyzed by STEFLUX) play an important role in transporting stratospheric air-masses from the free troposphere to the surface (Cristofanelli et al., 2010).

Table 3a focuses on the “SIO vs STEFLUX” comparison, as a function of the length of the different events. STEFLUX captured ~~almost one third~~ one quarter of the measured events for NCO-P and Mt. Cimone. The highest agreement was found for 2-day events at NCO-P (42%), and for 3-day events at Mt. Cimone (39%). In particular, all the longest events were confirmed by STEFLUX at NCO-P (9- and 14-day long events), while at Mt. Cimone only 2 of 4 events longer than 7 days were captured. Finally, the “STEFLUX vs SIO” approach is correspondingly assessed in Table 3b. ~~Nearly one third~~ One quarter of the SI events observed by STEFLUX were confirmed by SIO at NCO-P and Mt. Cimone (25% and 22%, respectively). Again, the maximum agreement was found for events that lasted 2 days, while the minimum agreement was assessed for 1-day long events.

4.1.3 STEFLUX and SIO: strong and weak points

Several possible reasons can explain the mismatch between the SIO and STEFLUX time series. For instance, STEFLUX is not fully able to capture subgrid-scale processes (like convection, turbulent diffusion and mixing) along the path from the stratosphere to the target region. This deficiency becomes particularly pronounced over mountainous measurement sites, mostly because of the complex topography and the associated small-scale thermally and dynamically driven circulations that characterize the area. As shown in Bracci et al. (2012), it is common that stratospheric air-masses reach the upper tropospheric layers over NCO-P, without directly arriving at the station altitude. Then, the air is trapped and mixed within the PBL and thus brought to the measurement station. It was for this reason that a specific criterion was introduced in the SIO detection methodology at NCO-P (see criterion ~~(iv) in the Supplementary Material~~4 in Sect. 2). It is worth noting that for NCO-P the largest bias between STEFLUX and SIO was observed when this criterion dominated the detection of SI (post-monsoons 2010 and 2011). Since the mixing processes might take several hours, this could be the reason of the lower agreement between SIO and STEFLUX at NCO-P.

In case of long travel times from the tropopause to the target region, we expect a stronger impact of mixing and dilution processes on the air-mass properties. Hence, when a SI actually affects a specific region, the SIO criteria might not be able to detect it, because mixing and dilution with tropospheric air-masses could lower stratospheric tracers concentrations below the thresholds used for detection. Then, we computed the travel times (hereinafter called Δt, expressed in hours) between the tropopause crossing and the first box crossing for each SI event. To evaluate the possible dependence from the travel time as a function of seasons, we sorted Δt into five categories (from 0 to 120 h, divided into 24 h intervals), and then we calculated the seasonal occurrence and the annual variation of each category (Fig. S1 and S2 in the Supplementary Material). The maximum value for Δt was chosen according to the typical lifetime values for a stratospheric intrusion into the troposphere (see Stohl et al., 2000; Bourqui and Trépanier, 2010; Trickl et al., 2014, 2016). On average, one third (32% and 30% for NCO-P and Mt. Cimone, respectively) of the SI events identified by STEFLUX presented maximum travel times (96 h $\leq \Delta t <$ 120 h). Furthermore, SI events characterized by relatively long (i.e., $\Delta t \geq$ 72 h) travel times usually dominated all the seasons. This suggests that a significant impact of dilution/turbulence small scale processes along stratospheric air-mass transport is likely and might explain part of the mismatch between STEFLUX and SIO. This hypothesis was further confirmed by analyzing the events seen by STEFLUX, but not confirmed by SIO, as a function of Δt: most of them (86% and 88% for NCO-P and Mt. Cimone, respectively) were characterized by medium/long travel times (i.e., $\Delta t \geq 48$ h).

A further point of discrepancy between STEFLUX and SIO results is related to the "overpasses" phenomenon, i.e., air-masses that overpass the station at altitudes high enough that there is no indication in the measurements record (but might be observed by STEFLUX). Indeed, during a study conducted at the Zugspitze (Germany, 2962 m a.s.l.), Trickl et al. (2010) showed that overpasses explained nearly the 20% of occurrences that were not identified by the observations. Furthermore, for NCO-P, it should be considered that the station is located in a narrow valley. Thus, it is conceivable that, during the transport within the valley, O_3 (one of the stratospheric tracers considered by SIO) experiences deposition phenomena, thus decreasing the actual concentration that the stratospheric air-mass would have in the free troposphere (see, e.g., Furger et al., 2000; Wotawa and Kromp-Kolb, 2000).

In summary, although correctly depicting the typical seasonal variability of SI frequency at NCO-P and Mt. Cimone, the STEFLUX and SIO time series differ for several reasons. These differences point out that the complete approach to study and assess SI is to deploy together modeling tools and observations, because they are complementary and address together several scientific questions. Especially, in situ observations have the advantage of capturing short and transient SI events associated to transport processes occurring at subgrid scales, while STEFLUX has the advantage of detecting the arrival of stratospheric-affected air-masses, irrespective of the degree of mixing and dilution along the transport within the troposphere.

4.2 Long-term evaluation of SI occurrences at the two measurement sites

4.2.1 SI events climatology

In this section we present a climatology of SI events, as defined in Sect. 4.1.2, for the whole STEFLUX dataset (back to 1979, i.e., when the trajectories from the ERA-Interim reanalysis were first available). This allowed us to cover a 35-year period (1979–2013) of monthly SI frequency values. In total, we obtained 673 SI events at NCO-P (representing 13% of the period), with an average length of 2.6 days. For Mt. Cimone, the number of SI events was lower (592, representing 9% of the period), as well as the average duration of an event, i.e., 2.1 days. The percentage of events with length equal to or less than 4 days was very high for both stations (86% and 93% considering all data, for NCO-P and Mt. Cimone, respectively), with peaks up to 98% and 100% in the summer season. On the other hand, longer events were observed during winter at NCO-P and were more equally spread throughout the rest of the year at Mt. Cimone. The seasonal cycle was also computed for these two longer time series (see Fig. S3 in the Supplementary Material); the seasonality considering all monthly data was confirmed and comparable to that obtained in Fig. 3, for both measurement sites.

An important aspect of SI is where and when the SI trajectories actually crossed the tropopause. First, the location of the crossing ~~determines~~ is useful to determine the O_3 concentration of the air parcels ~~exactly before they continue their path in the troposphere~~ at the start of their tropospheric path towards the target region. Second, as mentioned in Sect. 4.1.3, we can expect that a longer time since the tropopause crossing goes along with enhanced dilution until its arrival in the target region, although keeping in mind that the diluting processes along the path can be highly transient and nonlinear in time. As introduced in Sect. 3.2, STEFLUX allows the position and time of the tropopause crossings to be analyzed. This allowed us to compute a tropopause crossing density plot for each measurement site over the entire 35-year period, as presented in Fig. 4. For NCO-P (Fig. 4a), the tropopause crossings associated with SI events are predominantly found over two areas, i.e., Central Asia and Northeast Africa, close to the Mediterranean Sea. On the other hand, no prevalent locations characterize the tropopause crossing for SI events at Mt. Cimone (Fig. 4b), where the crossings are spread over a large area extending from North America to the northern Europe. ~~If divided seasonally~~Considering seasons separately, a small cluster emerges ~~southward~~ south of Greenland during winter, while the other seasons still maintain the crossings spread over a larger area (not shown). These results agree with previous climatological studies (Sprenger and Wernli, 2003; Škerlak et al., 2014), which indicated that the tropopause crossing predominantly occurs over the Atlantic and Pacific storm track regions (in winter, spring and autumn), over the Mediterranean (in winter and spring) and over southeastern Europe and Central Asia (in summer). The tropopause crossing locations were then categorized according to Δt (defined in Sect. 4.1.3), see Fig. S4 and S5 in the Supplementary Material. For NCO-P, the Central Asia zone of tropopause crossing was pronounced for all Δt categories, up to 96 h, while the Northeast Africa cluster was absent for events with low Δt and clearly discernible for events with 48 h $\leq \Delta$t $<$72 h, stressing the importance, for the southern Himalayas, of the fast transport of stratospheric air-masses embedded within the subtropical jet stream (Bracci et al., 2012). As stated above, the fact that SI events at Mt. Cimone showed no preferred tropopause crossing locations was confirmed for all the different Δt categories.

4.2.2 SI frequency long-term trends and variability

To detect potential trends in the SI frequencies, we adopted the same STEFLUX climatological datasets presented in Sect. 4.2.1. Trends were calculated by using the Theil-Sen (Theil, 1950; Sen, 1968) regression method implemented in the Openair software (Carslaw and Ropkins, 2012), after having deseasonalized the time series. No significant trends in the SI frequency were discernible for the stations, which both showed only a weak increase of 0.03 % yr^{-1}. In addition to this estimation, we repeated the trend analysis based on seasonal SI frequencies. The only significant ($p < 0.1$) positive trend was found for winter at NCO-P (0.18 % yr^{-1}). The lack of an overall trend in SI events was in line with previous findings, such as Sprenger and Wernli (2003) and Škerlak et al. (2014).

The long-term variability of SI frequency at NCO-P and Mt. Cimone was further analyzed with respect to potential oscillations and periodicities. To this aim, we applied the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise method (CEEMDAN, Torres et al., 2011). This technique is an improved version of the original EMD method (Huang et al., 1998) based on the Hilbert-Huang transform and practical for non-linear and non-stationary time series. It aims at subtracting several components (i.e., the so-called Intrinsic Mode Functions, IMFs) from the original signal, each of which explains a different cyclic variation, and a residual, which represents the overall trend in the original time series. The method has been recently used in atmospheric and climatic studies (e.g., Coughlin and Tung, 2004; Xu et al., 2016), but none of these regarded trends in SI yet. For both sites, the time series could be decomposed into 7 IMFs, with very different time scales (Fig. 5 for NCO-P and Fig. 6 for Mt. Cimone). Since we were particularly interested in long-term variations, we neglected high-frequency oscillations with characteristic periods shorter than one year. Apart from an evident seasonal cycle (Fig. 5b), the time series at NCO-P presents an IMF with a clear period of 28 months (IMF5, Fig. 5c) that is weakly anti-correlated (r = -0.3) to the Quasi-Biennial Oscillation (QBO); the anti-correlation is maximized during post-monsoon and winter seasons (r = -0.5 and r = -0.4, respectively). In this work we adopted the QBO index (blue line in Fig. 5c) for comparison, archived by the Free University of Berlin (http://www.geo.fu-berlin.de/met/ag/strat/produkte/qbo/qbo.dat). It refers to the monthly equatorial zonal wind at 50 hPa. Signals relating STE and QBO were found by Hsu and Prather (2009), who indicated that 20% of the inter-annual STE variance in the Northern Hemisphere can be explained by the QBO. More generally, the mechanisms for which QBO affects the STE variability are both the direct modulation of the circulation through thermal wind balance, and the impact on the strength of the overturning circulation by altering the propagation and dissipation of planetary-scale waves (Tung and Yang, 1994; Kinnersley and Tung, 1999; Neu et al., 2014). IMF6 (Fig. 5d) exhibits two peaks in the power spectrum, corresponding to periods of 3.5 and 5.8 years (not shown), potentially indicating an influence from the El-Niño/Southern Oscillation (ENSO). In fact, ENSO has been found to have an impact on the STE variability via the induced anomalous strong convective activity in the tropics (James et al., 2003). Moreover, a strong correlation between STE and ENSO was found by Zeng and Pyle (2005) and Voulgarakis et al. (2011), with the total modeled STE maximized during El Niño and minimized during La Niña years. The link is probably caused by modulations of the subtropical jet. In our work, IMF6 ~~reveals a slight anti-correlation with~~ presents some periods of inverse variability with respect to the Multivariate ENSO Index (MEI, Wolter and Timlin, 1993), included as the red line in Fig. 5d. Similar relations were also reported in Neu et al. (2014), with the STE flux increased during El Niño/easterly-shear QBO, because of the strengthening of the stratospheric overturning circulation and the intensified transport of air from the ozone maximum poleward and downward to midlatitudes. Conversely, La Niña/westerly QBO phases are associated with a weakening of the circulation and hence reduced STE flux. The last IMF (IMF7, Fig. 5e) shows a period of nearly 10 years, possibly related to the solar cycle. The time series of the 13-month smoothed monthly total sunspot number (orange line in Fig. 5e, retrieved by the Royal Observatory of Belgium, http://www.sidc.be/silso/datafiles) is positively correlated with IMF7 (r = 0.7). Signals of influence of the sunspot cycle in the upper troposphere-lower stratosphere have been indicated in several works (e.g., Labitzke and Van Loon, 1997a, b; Coughlin and Tung, 2004), suggesting that the association between the Sun and stratospheric parameters (e.g., O_3) is due to solar-induced changes in the atmospheric circulation.

For Mt. Cimone the situation was different and more difficult to understand. The seasonal cycle is still present (IMF4, Fig. 6b), but IMF5 (Fig. 6c) ~~does not exhibit a clear period~~exhibits no clear periodicity, presenting peaks at 20, 28 and 35 months. In this case, the comparison with the QBO signal does not highlight any evident relation. Variations in the peak amplitudes of IMF6 (Fig. 6d) are more regular, with two prominent oscillations with periods of 4.4 (ENSO) and 2.9 years. The last IMF (Fig. 6e) results in a characteristic period of 11.6 years, which most likely is related to the solar cycle, as indicated by the high correlation (r = 0.8) with the sunspot number (orange line in Fig. 6e).

5 Conclusions

In this work we presented a novel methodology (STEFLUX) to evaluate SI in a ~~defined~~ user-defined region, by using as input a Lagrangian STE climatology derived from the ERA-Interim reanalysis. Besides having shown ~~a representative~~ an illustrative case study (Sect. 3.2) as a typical STEFLUX application, we investigated STEFLUX skills in detecting SI by comparing its time series with corresponding long-term SI time series derived from observational datasets (SIO, see Sect. 2). The analysis was performed in two very different areas of the world, i.e., the southern Himalayas (NCO-P) and the central Mediterranean basin (Mt. Cimone), which represent two global hot-spots for what concerns air pollution and climate change, and are often affected by SI.

Our results showed that STEFLUX correctly represented the typical seasonal cycles of SI frequencies over these two areas, with the highest occurrence of SI in winter for NCO-P, and in winter–spring for Mt. Cimone. For both sites, the lowest SI occurrence was recorded during summer (i.e., the monsoon for NCO-P). STEFLUX had real skill (higher for Mt. Cimone than NCO-P) in detecting single events at both regions, especially for robust (i.e., longer than 1 day) events. The identification of short events was more problematic; this is in agreement with a similar study by Cui et al. (2009), who reported considerable difficulties for two Lagrangian models in capturing "inconspicuous" SI events. This issue is also reflected by a low agreement in the evaluation of the inter-annual variability of SI frequency (especially for NCO-P during winter). Hence, STEFLUX should not be regarded as a tool to exactly reproduce SI occurrences at specific measurement sites, which typically are strongly affected by rather small-scale circulations. Instead, its premium application is in determining the SI input at a larger, regional scale. Moreover, although not investigated in this work, STEFLUX might be deployed as a particularly relevant tool to investigate how SI long-term variability influences the atmospheric composition at these specific locations (e.g., by deploying the O_3 values that are available along each trajectory).

The observed mismatch between STEFLUX and SIO is due to several reasons, such as the absence of representation of subgrid-scale processes in STEFLUX (e.g., convection, turbulent diffusion and mixing), along the path from the stratosphere to the target region. Last but not least, one should consider that these subgrid-scale processes can also lead to "local" or transient SI events captured by a single measurement point, which cannot be considered representative/significant for a whole region. In addition to this, another reason for the mismatch might be mixing and dilution processes occurring within air-masses from the tropopause crossing to the target region, expected to be maximized for greater travel times. Lastly, as also demonstrated in previous studies, the "overpasses" phenomenon might have a not negligible impact. All of these discrepancies point out that a combination of modeling outputs (e.g., STEFLUX) and in situ observations is still needed to completely study, characterize and evaluate the occurrence of SI.

Another important feature of STEFLUX is its capability of climatologically assessing the SI occurrence at the chosen site, since the ERA-Interim reanalysis extends back to 1979. In this study, it allowed us to obtain a 35-year time series of SI events at NCO-P and Mt. Cimone, which affected 13% and 9% of the period, respectively. The tropopause crossings during the whole 35-y period, provided by STEFLUX, were further analyzed. NCO-P showed two main cluster regions, i.e., Central Asia (maximized for events with short travel times between the tropopause and the target box, Δt) and northern Africa (which had its maximum for events with 48 h $\leq \Delta$t $<$72 h). On the other hand, no preferred locations characterized Mt. Cimone, except for a small cluster ~~southward~~ south of Greenland during winter. We then evaluated trends on these long time series: no trends in the SI occurrence were discernible for both ~~of the~~ measurement sites, and the only statistically significant trend was observed for winter at NCO-P (0.18 % yr^{-1}). Furthermore, for the first time the CEEMDAN analysis has been performed on the time series characterizing these two hot-spot areas, in order to evaluate periodicities and their possible relation to climate factors. At NCO-P, signs of influence from the Quasi-Biennial Oscillation (QBO), the El-Niño/Southern Oscillation (ENSO) and the solar cycle were found, while Mt. Cimone exhibited relevant relations with ENSO and the solar cycle (high correlation with the sunspot number). These results indicate the possible impact of anthropogenic climate change on SI occurrence via changes in the ENSO and QBO regimes.

STEFLUX availability

The STEFLUX outputs are available on request by writing an e-mail to the authors, specifying the box characteristics and the period of study chosen.

Acknowledgements. The authors thank the ECMWF and MeteoSwiss for providing access to the meteorological data, the technical and logistic staff at NCO-P, F. Calzolari and F. Roccato (CNR–ISAC) for the technical support at Mt. Cimone, as well as the “Magera team” (C. Magera, P. Giambi and N. Gherardini) and the Italian Air Force (CAMM Monte Cimone) for the valuable co-operation. Furthermore, the authors thank the NOAA–CPC for the access to the MEI index, the Free University of Berlin for the QBO index, the Royal Observatory of Belgium for providing the sunspot number, and S. Pfahl for the useful discussions.

References

Appenzeller, C. and Davies, H. C.: Structure of stratospheric intrusions into the troposphere, Nature, 358, 570–572, 1992.

Appenzeller, C., Davies, H., and Norton, W.: Fragmentation of stratospheric intrusions, Journal of Geophysical Research: Atmospheres, 101, 1435–1456, 1996.

Bonasoni, P., Laj, P., Marinoni, A., Sprenger, M., Angelini, F., Arduini, J., Bonafè, U., Calzolari, F., Colombo, T., Decesari, S., et al.: Atmospheric Brown Clouds in the Himalayas: first two years of continuous observations at the Nepal Climate Observatory-Pyramid (5079 m), Atmospheric Chemistry and Physics, 10, 7515–7531, 2010.

Bourqui, M.: Stratosphere-troposphere exchange from the Lagrangian perspective: a case study and method sensitivities, Atmospheric Chemistry and Physics, 6, 2651–2670, 2006.

Bourqui, M. and Trépanier, P.-Y.: Descent of deep stratospheric intrusions during the IONS August 2006 campaign, Journal of Geophysical Research: Atmospheres, 115, 2010.

Bracci, A., Cristofanelli, P., Sprenger, M., Bonafè, U., Calzolari, F., Duchi, R., Laj, P., Marinoni, A., Roccato, F., Vuillermoz, E., et al.: Transport of stratospheric air masses to the Nepal Climate Observatory-Pyramid (Himalaya; 5079 m MSL): a synoptic-scale investigation, Journal of Applied Meteorology and Climatology, 51, 1489–1507, 2012.

Carslaw, D. C. and Ropkins, K.: Openair—an R package for air quality data analysis, Environmental Modelling & Software, 27, 52–61, 2012.

Cooper, O., Stohl, A., Hübler, G., Hsie, E., Parrish, D., Tuck, A., Kiladis, G., Oltmans, S., Johnson, B., Shapiro, M., et al.: Direct transport of midlatitude stratospheric ozone into the lower troposphere and marine boundary layer of the tropical Pacific Ocean, Journal of Geophysical Research: Atmospheres, 110, 2005.

Coughlin, K. and Tung, K.-K.: 11-year solar cycle in the stratosphere extracted by the empirical mode decomposition method, Advances in Space Research, 34, 323–329, 2004.

Cristofanelli, P., Bonasoni, P., Tositti, L., Bonafe, U., Calzolari, F., Evangelisti, F., Sandrini, S., and Stohl, A.: A 6-year analysis of stratospheric intrusions and their influence on ozone at Mt. Cimone (2165 m above sea level), Journal of Geophysical Research: Atmospheres, 111, 2006.

Cristofanelli, P., Bracci, A., Sprenger, M., Marinoni, A., Bonafè, U., Calzolari, F., Duchi, R., Laj, P., Pichon, J., Roccato, F., et al.: Tropospheric ozone variations at the Nepal Climate Observatory-Pyramid (Himalayas, 5079 m asl) and influence of deep stratospheric intrusion events, Atmospheric Chemistry and Physics, 10, 6537–6549, 2010.

Cristofanelli, P., Scheel, H.-E., Steinbacher, M., Saliba, M., Azzopardi, F., Ellul, R., Fröhlich, M., Tositti, L., Brattich, E., Maione, M., et al.: Long-term surface ozone variability at Mt. Cimone WMO/GAW global station (2165 m asl, Italy), Atmospheric Environment, 101, 23–33, 2015.

Cui, J., Sprenger, M., Staehelin, J., Siegrist, A., Kunz, M., Henne, S., and Steinbacher, M.: Impact of stratospheric intrusions and intercontinental transport on ozone at Jungfraujoch in 2005: comparison and validation of two Lagrangian approaches, Atmospheric Chemistry and Physics, 9, 3371–3383, 2009.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis:

configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi:10.1002/qj.828, http://dx.doi.org/10.1002/qj.828, 2011.

Furger, M., Dommen, J., Graber, W. K., Poggio, L., Prévôt, A. S., Emeis, S., Grell, G., Trickl, T., Gomiscek, B., Neininger, B., et al.: The VOTALP Mesolcina Valley Campaign 1996–concept, background and some highlights, Atmospheric Environment, 34, 1395–1412, 2000.

Gauss, M., Myhre, G., Pitari, G., Prather, M., Isaksen, I., Berntsen, T., Brasseur, G. P., Dentener, F., Derwent, R., Hauglustaine, D., et al.: Radiative forcing in the 21st century due to ozone changes in the troposphere and the lower stratosphere, Journal of Geophysical Research: Atmospheres, 108, 2003.

Holton, J. R., Haynes, P. H., McIntyre, M. E., Douglass, A. R., Rood, R. B., and Pfister, L.: Stratosphere-troposphere exchange, Reviews of Geophysics, 33, 403–439, 1995.

Hsu, J. and Prather, M. J.: Stratospheric variability and tropospheric ozone, Journal of Geophysical Research: Atmospheres, 114, 2009.

Huang, N. E., Shen, Z., Long, S. R., Wu, M. C., Shih, H. H., Zheng, Q., Yen, N.-C., Tung, C. C., and Liu, H. H.: The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis, in: Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, vol. 454, pp. 903–995, The Royal Society, 1998.

James, P., Stohl, A., Forster, C., Eckhardt, S., Seibert, P., and Frank, A.: A 15-year climatology of stratosphere–troposphere exchange with a Lagrangian particle dispersion model 2. Mean climate and seasonal variability, Journal of Geophysical Research: Atmospheres, 108, 2003.

Kinnersley, J. S. and Tung, K. K.: Mechanisms for the extratropical QBO in circulation and ozone, Journal of the atmospheric sciences, 56, 1942–1962, 1999.

Labitzke, K. and Van Loon, H.: The signal of the 11-year sunspot cycle in the upper troposphere-lower stratosphere, Space Science Reviews, 80, 393–410, 1997a.

Labitzke, K. and Van Loon, H.: Total ozone and the 11-yr sunspot cycle, Journal of Atmospheric and Solar-Terrestrial Physics, 59, 9–19, 1997b.

Lamarque, J.-F. and Hess, P. G.: Cross-tropopause mass exchange and potential vorticity budget in a simulated tropopause folding, Journal of the atmospheric sciences, 51, 2246–2269, 1994.

Lin, M., Fiore, A. M., Cooper, O. R., Horowitz, L. W., Langford, A. O., Levy, H., Johnson, B. J., Naik, V., Oltmans, S. J., and Senff, C. J.: Springtime high surface ozone events over the western United States: Quantifying the role of stratospheric intrusions, Journal of Geophysical Research: Atmospheres, 117, 2012.

Meloen, J., Siegmund, P., Van Velthoven, P., Kelder, H., Sprenger, M., Wernli, H., Kentarchos, A., Roelofs, G., Feichter, J., Land, C., et al.: Stratosphere-troposphere exchange: A model and method intercomparison, Journal of Geophysical Research: Atmospheres, 108, 2003.

Monks, P., Granier, C., Fuzzi, S., Stohl, A., Williams, M., Akimoto, H., Amann, M., Baklanov, A., Baltensperger, U., Bey, I., et al.: Atmospheric composition change–global and regional air quality, Atmospheric Environment, 43, 5268–5350, 2009.

Neu, J. L., Flury, T., Manney, G. L., Santee, M. L., Livesey, N. J., and Worden, J.: Tropospheric ozone variations governed by changes in stratospheric circulation, Nature Geoscience, 7, 340–344, 2014.

Ordóñez, C., Brunner, D., Staehelin, J., Hadjinicolaou, P., Pyle, J., Jonas, M., Wernli, H., and Prévôt, A.: Strong influence of lowermost stratospheric ozone on lower tropospheric background ozone changes over Europe, Geophysical Research Letters, 34, 2007.

Reed, R. J.: A study of a characteristic type of upper-level frontogenesis, Journal of Meteorology, 12, 226–237, 1955.

Roelofs, G.-J., Lelieveld, J., and Dorland, R.: A three-dimensional chemistry/general circulation model simulation of anthropogenically derived ozone in the troposphere and its radiative climate forcing, Journal of Geophysical Research: Atmospheres, 102, 23 389–23 401, 1997.

Sebald, L., Treffeisen, R., Reimer, E., and Hies, T.: Spectral analysis of air pollutants. Part 2: ozone time series, Atmospheric Environment, 34, 3503–3509, 2000.

Sen, P. K.: Estimates of the regression coefficient based on Kendall's tau, Journal of the American Statistical Association, 63, 1379–1389, 1968.

Škerlak, B., Sprenger, M., and Wernli, H.: A global climatology of stratosphere–troposphere exchange using the ERA-Interim data set from 1979 to 2011, Atmospheric Chemistry and Physics, 14, 913–937, 2014.

Sprenger, M. and Wernli, H.: A Northern Hemispheric climatology of cross-tropopause exchange for the ERA15 time period (1979–1993), Journal of Geophysical Research: Atmospheres, 108, 2003.

Sprenger, M. and Wernli, H.: The LAGRANTO Lagrangian analysis tool – version 2.0, Geoscientific Model Development, 8, 2569–2586, 2015.

Sprenger, M., Wernli, H., and Bourqui, M.: Stratosphere-troposphere exchange and its relation to potential vorticity streamers and cutoffs near the extratropical tropopause, Journal of the atmospheric sciences, 64, 1587–1602, 2007.

Stevenson, D., Dentener, F., Schultz, M., Ellingsen, K., Van Noije, T., Wild, O., Zeng, G., Amann, M., Atherton, C., Bell, N., et al.: Multimodel ensemble simulations of present-day and near-future tropospheric ozone, Journal of Geophysical Research: Atmospheres, 111, 2006.

Stohl, A., Wotawa, G., Seibert, P., and Kromp-Kolb, H.: Interpolation errors in wind fields as a function of spatial and temporal resolution and their impact on different types of kinematic trajectories, Journal of Applied Meteorology, 34, 2149–2165, 1995.

Stohl, A., Spichtinger-Rakowsky, N., Bonasoni, P., Feldmann, H., Memmesheimer, M., Scheel, H., Trickl, T., Hübener, S., Ringer, W., and Mandl, M.: The influence of stratospheric intrusions on alpine ozone concentrations, Atmospheric Environment, 34, 1323–1354, 2000.

Stohl, A., Bonasoni, P., Cristofanelli, P., Collins, W., Feichter, J., Frank, A., Forster, C., Gerasopoulos, E., Gäggeler, H., James, P., et al.: Stratosphere-troposphere exchange: A review, and what we have learned from STACCATO, Journal of Geophysical Research: Atmospheres, 108, 2003.

Stohl, A., Forster, C., Frank, A., Seibert, P., and Wotawa, G.: Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2, Atmospheric Chemistry and Physics, 5, 2461–2474, 2005.

Tarasova, O. A., Senik, I. A., Sosonkin, M. G., Cui, J., Staehelin, J., and Prévôt, A. S. H.: Surface ozone at the Caucasian site Kislovodsk High Mountain Station and the Swiss Alpine site Jungfraujoch: data analysis and trends (1990–2006), Atmospheric Chemistry and Physics, 9, 4157–4175, 2009.

Theil, H.: A rank-invariant method of linear and polynomial regression analysis, Part 3, in: Proceedings of Koninklijke Nederlandse Akademie van Weinenschatpen A, vol. 53, pp. 1397–1412, 1950.

Thornes, J. E. and Stephenson, D. B.: How to judge the quality and value of weather forecast products, Meteorological Applications, 8, 307–314, 2001.

Torres, M. E., Colominas, M. A., Schlotthauer, G., and Flandrin, P.: A complete ensemble empirical mode decomposition with adaptive noise, in: Acoustics, speech and signal processing (ICASSP), 2011 IEEE international conference on, pp. 4144–4147, IEEE, 2011.

Tositti, L., Brattich, E., Cinelli, G., and Baldacci, D.: 12 years of 7 Be and 210 Pb in Mt. Cimone, and their correlation with meteorological parameters, Atmospheric Environment, 87, 108–122, 2014.

Trickl, T., Feldmann, H., Kanter, H.-J., Scheel, H.-E., Sprenger, M., Stohl, A., and Wernli, H.: Forecasted deep stratospheric intrusions over Central Europe: case studies and climatologies, Atmospheric Chemistry and Physics, 10, 499–524, 2010.

Trickl, T., Vogelmann, H., Giehl, H., Scheel, H.-E., Sprenger, M., and Stohl, A.: How stratospheric are deep stratospheric intrusions?, Atmospheric Chemistry and Physics, 14, 9941–9961, 2014.

Trickl, T., Vogelmann, H., Fix, A., Schäfler, A., Wirth, M., Calpini, B., Levrat, G., Romanens, G., Apituley, A., Wilson, K. M., Begbie, R., Reichardt, J., Vömel, H., and Sprenger, M.: How stratospheric are deep stratospheric intrusions? LUAMI 2008, Atmospheric Chemistry and Physics, 16, 8791–8815, 2016.

Tung, K. and Yang, H.: Global QBO in circulation and ozone. Part II: A simple mechanistic model, Journal of the atmospheric sciences, 51, 2708–2721, 1994.

Voulgarakis, A., Hadjinicolaou, P., and Pyle, J.: Increases in global tropospheric ozone following an El Nino event: examining stratospheric ozone variability as a potential driver, Atmospheric Science Letters, 12, 228–232, 2011.

Wernli, B. H. and Davies, H. C.: A Lagrangian-based analysis of extratropical cyclones. I: The method and some applications, Quarterly Journal of the Royal Meteorological Society, 123, 467–489, 1997.

Wernli, H. and Bourqui, M.: A Lagrangian "1-year climatology" of (deep) cross-tropopause exchange in the extratropical Northern Hemisphere, Journal of Geophysical Research: Atmospheres, 107, 2002.

Wilks, D. S.: Statistical methods in the atmospheric sciences: an introduction, 2nd edition, Academic press, 2006.

Wolter, K. and Timlin, M. S.: Monitoring ENSO in COADS with a seasonally adjusted principal component index, in: Proc. of the 17th Climate Diagnostics Workshop, pp. 52–7, 1993.

Wotawa, G. and Kromp-Kolb, H.: The research project VOTALP–general objectives and main results, Atmospheric Environment, 34, 1319–1322, 2000.

Xu, W., Lin, W., Xu, X., Tang, J., Huang, J., Wu, H., and Zhang, X.: Long-term trends of surface ozone and its influencing factors at the Mt. Waliguan GAW station, China–Part 1: Overall trends and characteristics, Atmospheric Chemistry and Physics, 16, 6191–6205, 2016.

Young, P., Archibald, A., Bowman, K., Lamarque, J.-F., Naik, V., Stevenson, D., Tilmes, S., Voulgarakis, A., Wild, O., Bergmann, D., et al.: Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmospheric Chemistry and Physics, 13, 2063–2090, 2013.

Zeng, G. and Pyle, J. A.: Influence of El Nino Southern Oscillation on stratosphere/troposphere exchange and the global tropospheric ozone budget, Geophysical research letters, 32, 2005.

Table 1. ~~2×2 contingency tables~~ Input parameters for ~~the comparisons of the SI events time series, i.e., identified by the SIO and~~ STEFLUX ~~approaches,~~ for ~~NCO-P (a, b) and Mt. Cimone (c, d). For each table~~ the ~~Odds Ratio Skill Score~~ comparison with in situ measurements (~~ORSS~~SIO)~~is also reported, along with the minimum ORSS required to have real skill at the 99% confidence level (in parentheses, see Thornes and Stephenson, 2001)~~.

Parameter	NCO-P	Mt. Cimone
Lat_min, Lat_max	26° N, 29° N	43° N, 46° N
Lon_min, Lon_max	85° E, 88° E	9° E, 12° E
Box_top	550 hPa	790 hPa
Time span	01 Mar. 2006–31 Dec. 2013	01 Jan. 1998–31 Dec. 2010
Temporal resolution	1 h	1 h

Table 2. 2×2 contingency tables for the comparisons of the SI events time series, i.e., identified by the SIO and STEFLUX approaches, for NCO-P (a, b) and Mt. Cimone (c, d). For each table the Odds Ratio Skill Score (ORSS) is also reported, along with the minimum ORSS required to have real skill at the 99% confidence level (in parentheses, see Thornes and Stephenson, 2001). Capital letters are defined as follows: A indicates the number of SI events selected by both methodologies (STEFLUX and SIO); B represents the number of events selected as SI by the first methodology but as no-SI by the second one; C represents the number of events selected as no-SI by the first methodology but as SI by the second one; and D represents the number of no-SI events selected by both methodologies.

NCO-P							
(a)	"SIO vs STEFLUX"			(b)	"STEFLUX vs SIO"		
		STEFLUX				SIO	
		SI	no-SI			SI	no-SI
SIO	SI	A = 55	B = 148	STEFLUX	SI	A = 39	B = 116
	no-SI	C = 23	D = 181		no-SI	C = 16	D = 140
ORSS	0.49 (0.35)			ORSS	0.49 (0.35)		
Mt. Cimone							
(c)	"SIO vs STEFLUX"			(d)	"STEFLUX vs SIO"		
		STEFLUX				SIO	
		SI	no-SI			SI	no-SI
SIO	SI	A = 73	B = 226	STEFLUX	SI	A = 52	B = 185
	no-SI	C = 25	D = 275		no-SI	C = 13	D = 225
ORSS	0.56 (0.35)			ORSS	0.66 (0.35)		

Table 3. (a) "SIO vs STEFLUX": agreement between STEFLUX and the measured SI events (SIO), and (b) "STEFLUX vs SIO": agreement between the measured and the modeled (by using STEFLUX) SI events, as a function of the different length of the SI events.

(a)	NCO-P		Mt. Cimone	
SI event duration	SI events by SIO	STEFLUX	SI events by SIO	STEFLUX
1 day	117	25 (22%)	217	45 (21%)
2 days	41	17 (42%)	36	11 (31%)
3 days	20	4 (20%)	28	11 (39%)
≥4 days	25	9 (36%)	18	6 (33%)
Total	203	55 (27%)	299	73 (24%)

(b)	NCO-P		Mt. Cimone	
SI event duration	SI events by STEFLUX	SIO	SI events by STEFLUX	SIO
1 day	55	7 (13%)	100	15 (15%)
2 days	36	15 (42%)	73	23 (31%)
3 days	22	7 (32%)	28	5 (18%)
≥4 days	42	10 (24%)	36	9 (25%)
Total	155	39 (25%)	237	52 (22%)

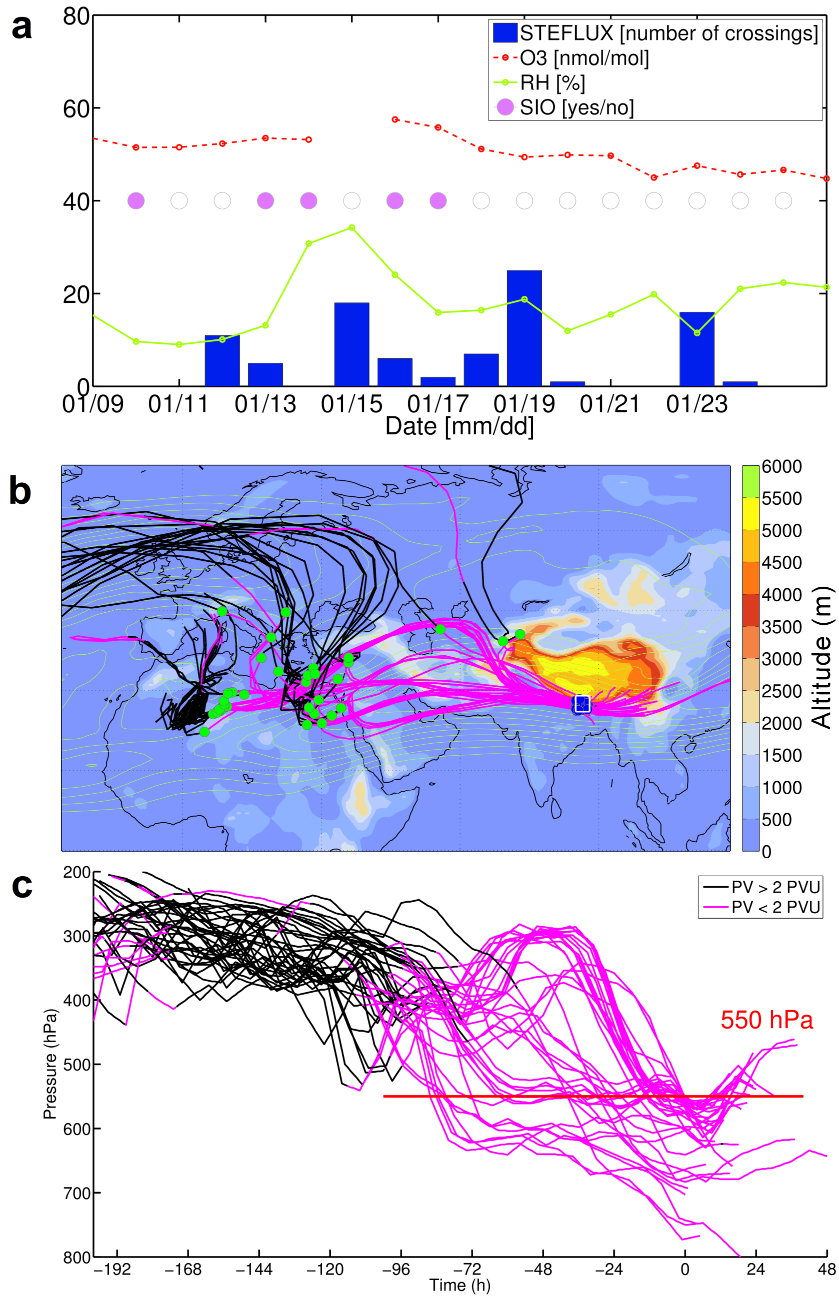

Figure 1. Example of application of STEFLUX, in a target box around NCO-P position over the period 09/01/2007–25/01/2007. Panel a shows the daily averages for O_3 and RH at NCO-P, the days selected by the SIO methodology (see Sect. 2) and the number of STE trajectory points crossing the box. The STE trajectories are also displayed entirely in panel b, where blue dots indicate the first crossings of the target box and green dots identify the tropopause crossing locations. Contour lines indicate the wind speed at 250 hPa, averaged over the case study period. Panel c shows the temporal-height evolution of the selected trajectories (0 is the time of arrival into the target box), as a function of the PV value along each trajectory point.

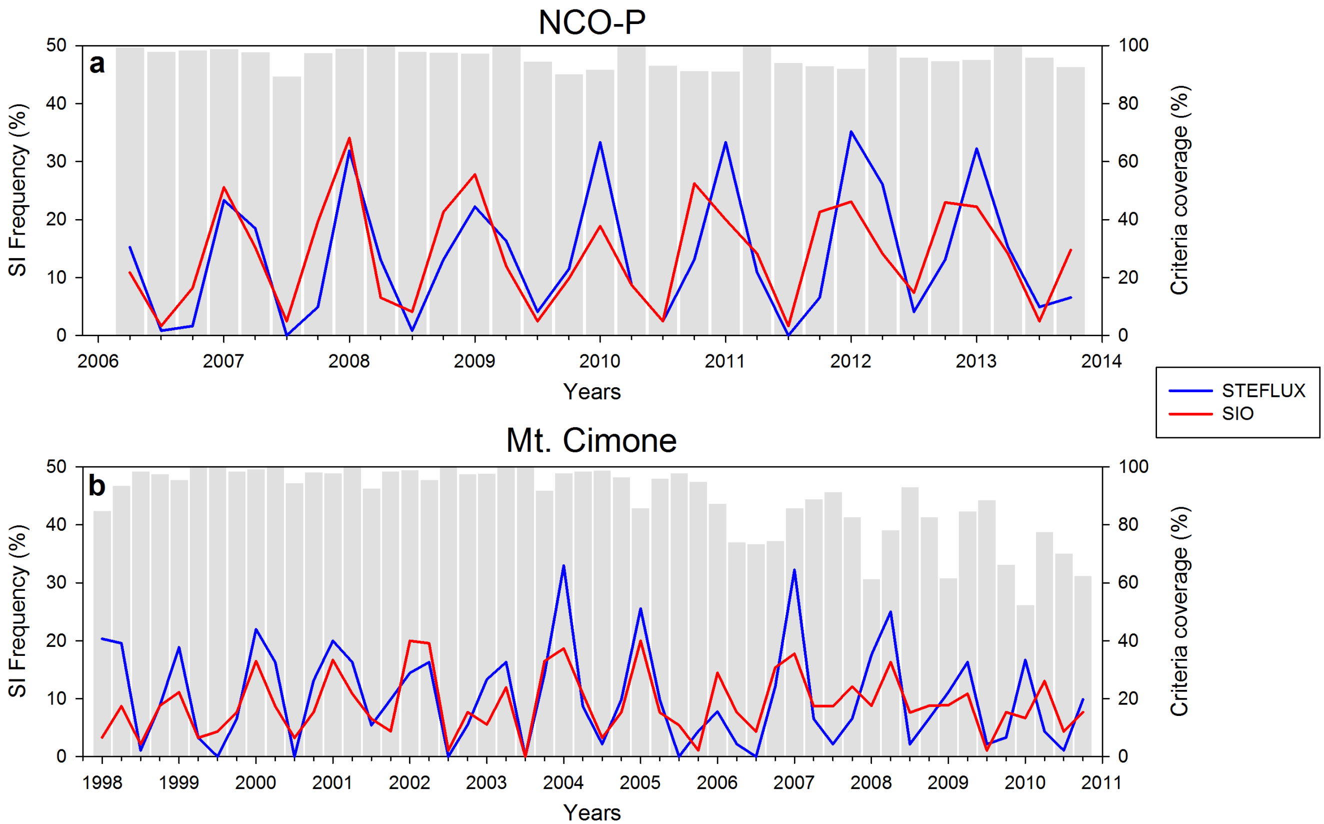

Figure 2. Seasonal graph of SI frequency at NCO-P (panel a) and Mt. Cimone (panel b). Blue lines indicate the STEFLUX outputs (for the configuration parameters see the Supplementary Material), while the red ones represent the results from the application of the criteria presented in Sect. 2 (SIO). Also shown in the plot is the percentage of criteria coverage (grey bars) for each season.

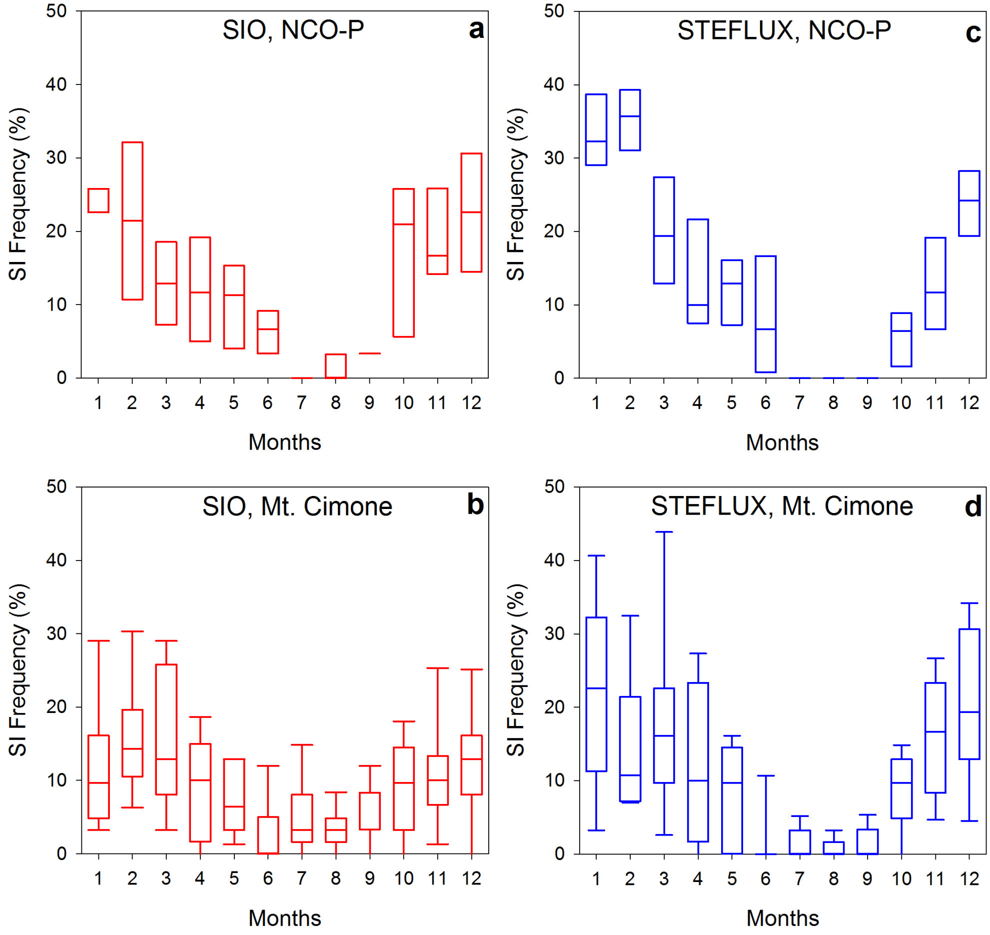

Figure 3. Box-whiskers plot of the annual variation of SI frequency at NCO-P (panels a,c) and Mt. Cimone (panels b,d). For both sites, the ~~STEFLUX~~ SIO (panels a,b) and ~~SIO~~ STEFLUX (panels c,d) values are presented.

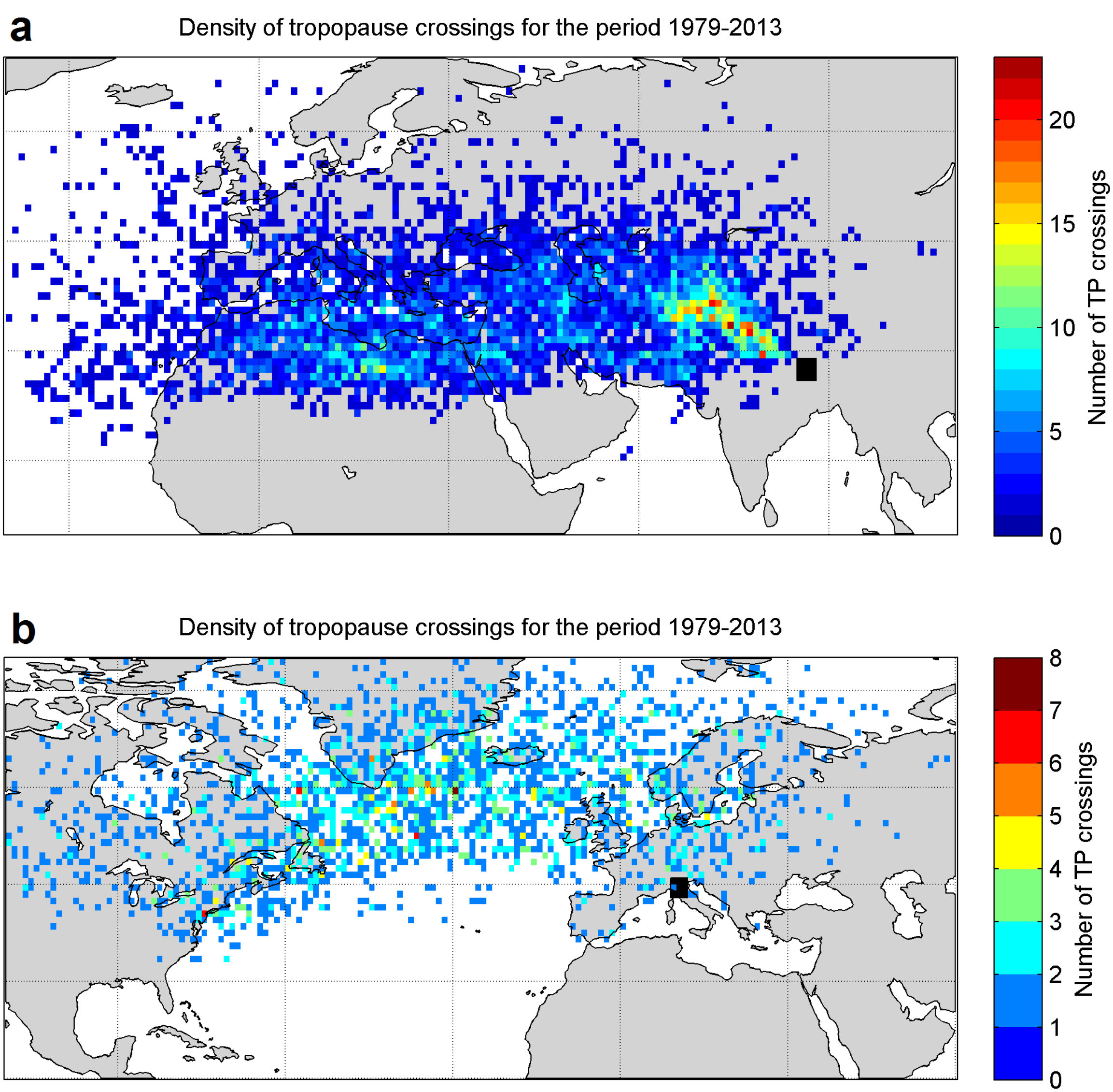

Figure 4. Density of tropopause (TP) crossings for the period 1979–2013, for NCO-P (panel a) and Mt. Cimone (b). Values for both figures were aggregated on a $1^\circ \times 1^\circ$ horizontal grid. In both panels, the black square indicates the horizontal extension of the STEFLUX box.

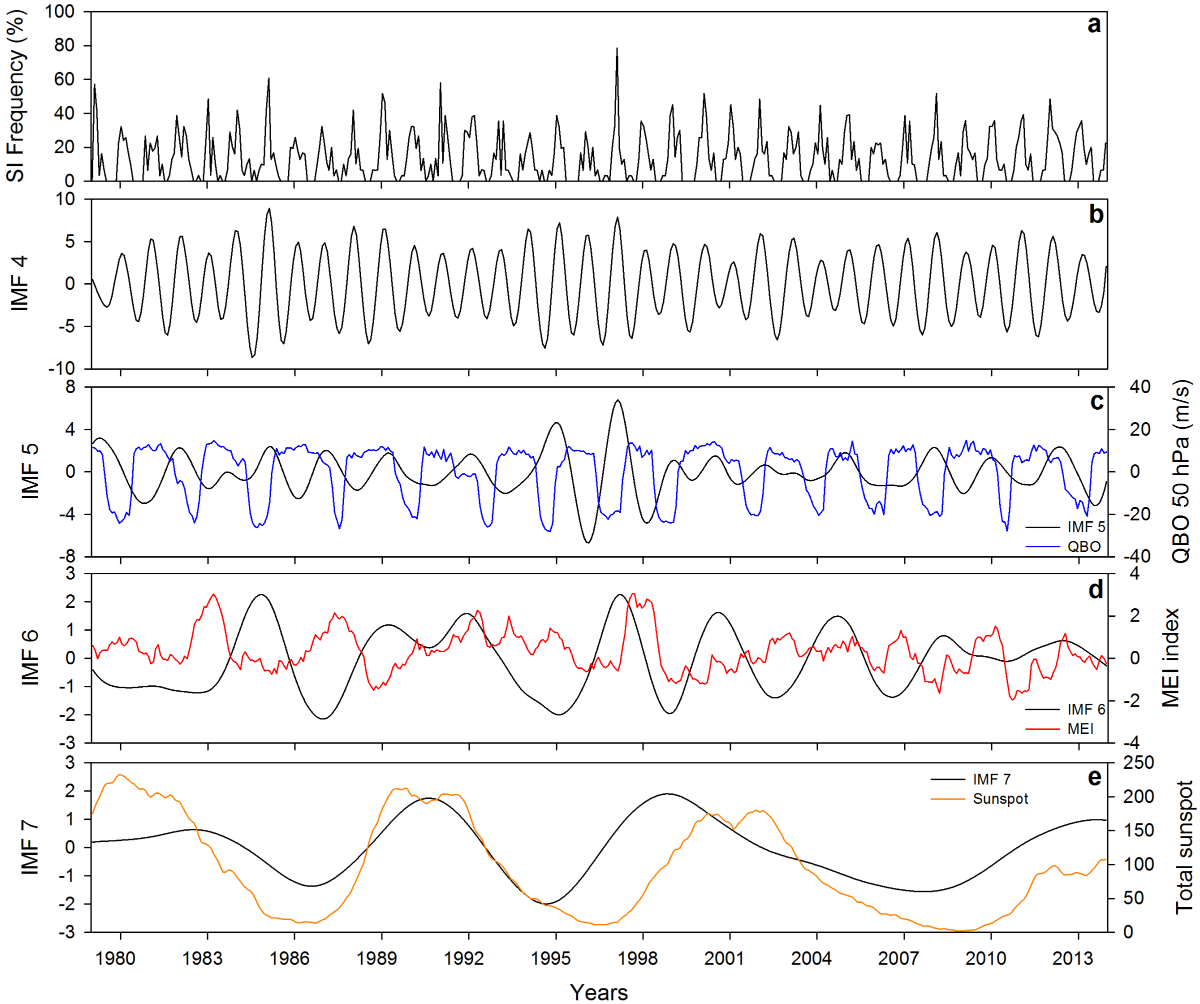

Figure 5. Time series of monthly averaged SI frequency, as retrieved by STEFLUX, at NCO-P (panel a), and some of its IMFs (i.e., IMF4–7, panels b–e, respectively) resulting from the application of the CEEMDAN analysis. The blue line in panel c represents the equatorial zonal wind at 50 hPa, used as a measure of the QBO signal, the red line in panel d depicts the Multivariate ENSO Index (MEI) and the orange line in panel e indicates the 13-month smoothed monthly total sunspot number.

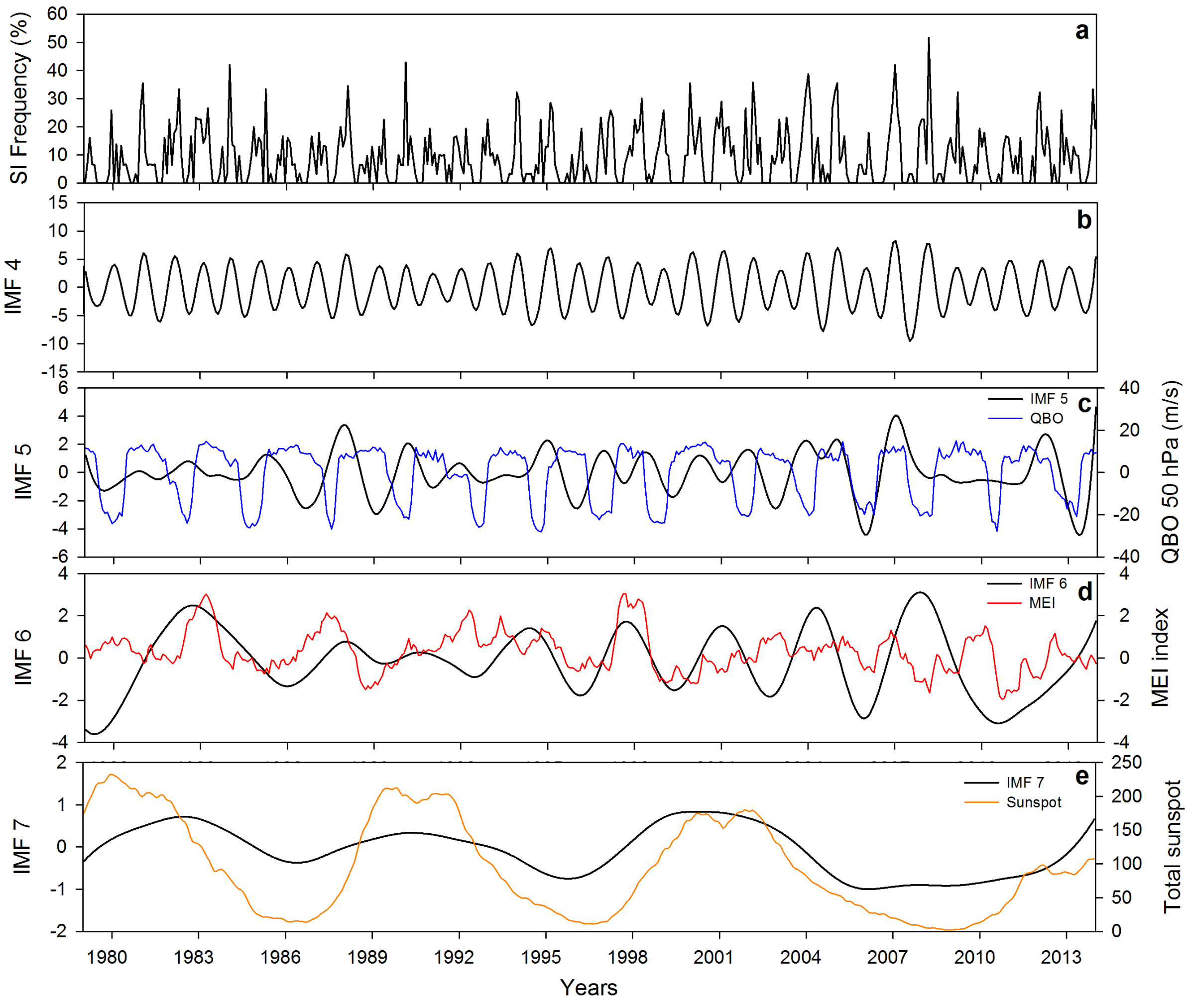

Figure 6. Same as Fig. 5, for Mt. Cimone time series.