We would like to thank both reviewers for their positive comments. Our responses are given below in blue text.

Response to referee #1

The Sentence in line 85 beginning with And should be rewritten to make the meaning more obvious and the phrase high pure helium in line 165 should be changed to high purity helium.

The sentence has been changed to "So far, no PFC tracer release and sampling system suitable for a deployment onboard research aircraft in the free troposphere (FT) and the upper troposphere/lower stratosphere (UT/LS) have been reported.", and "pure" has been changed to "purity".

Response to referee #2

General Comments

This article documents a new tracer system designed to track air masses over time windows of more than a day, sufficient for deformation by advection and mixing with surrounding air. The system includes both the tracer release mechanism, an adsorption tube sampler (ATS) and a laboratory analytical system for the PFC tracers. The system was tested for the first time during the SHIVA campaing by releasing the tracer from a ship and sampling the air mass using ATS. The aircraft was also equipped with a gas chromatograph which could detect the tracer although not calibrated. This proved very useful when comparing to simulated tracer mixing ratios and demonstrating the traverse of the tracer plume. The utility of the tracer system is established in this paper including its capability to resolve structure in the tracer that has been released.

The scientific approach is thorough and the instrumentation is described well in the paper. The authors were very careful in their experimental approach, for example by showing that there had been no contamination of sampling tubes by using blanks.

The crossing of the plume on the first aircraft flight 5 hours after release produced an impressive quantitative correspondence between concentrations sampled and the simulated concentrations using the HYSPLIT Lagrangian model. As happens during field experiments, a sample was missed at the key time (07:10) when the flight track first crossed the tracer plume. However, the second measurement from the gas chromatograph (GHOST) sampled in this gap and provides convincing evidence that the aircraft did indeed cross the centre of the plume at this early stage. By using a coincident sample measured by both ATS and GHOST, the GHOST measurements were calibrated and show that the GHOST sample near the plume centre sampled a concentration in close accord with the maximum simulated by HYSPLIT.

However, this good fortune also highlights a potential difficulty in the utilisation of the tracer release and sampling system. High frequency sampling is required (without gaps) to resolve the tracer filaments typical of a tracer release and it is difficult to direct an aircraft across the tracer for a representative sample. The later intercepts on the first flight grazed the flanks of the plume (although the mixing ratio correspondence is still impressive) and the second flight almost missed the tracer plume altogether. A wider question not addressed in this paper (which focuses on the experimental methods) is whether the sampling strategy and frequency can be made sufficiently good to infer quantitative information about the evolution of an air mass. For example, in order to infer air mass dilution rate from the tracer data (or to refine a model estimate) it would be necessary to: (i) fly transects across the centre of the plume several times, knowing that the centre had not been missed and (ii) to resolve the shape of tracer concentration distribution on each transect. Clearly forecasting the tracer plume and executing an aircraft flight pattern are central to the potential success of the system. In the conclusions, the authors should include some discussion of how they envisage the new tracer system could be used in the quantification of atmospheric processes and the parameters involved in models. The big question is what we can learn with the new system that we could not know without it?

We agree with the referee, that a full exploitation of the tracer release and sampling system depend on high resolution forecasts and multiple high resolution sampling transects through the plume. This has not yet fully achieved in this initial tracer experiment during SHIVA. However, the tracer sampler has such a capability with up 64 samples per flight, a possible sampling frequency as low as 30 s, and a flexible in-flight control system as described in the text "Depending on different experimental purposes, the sampling flow, sampling time and the interval between two samples, etc., can be controlled through the touch screen as well.".

As mentioned by the referee, the focus of the paper is on the experimental technique. Regarding the applications of the tracer release and sampling system, we will add the following text before the last paragraph in the "Summary and conclusion": "The tracer experiment during SHIVA demonstrated the capability of the new tracer system PERTRAS to unambiguously identify and trace an air mass tagged with a PFC. This proof of concept study provides the practical basis for the application of the PFC tracer system in future aircraft experiments in the FT and UT/LS to quantify small-scale diffusion parameters and synoptic-scale flow distortion in this altitude region. Furthermore, the release of two different PFC tracers will be helpful to understand the mixing between distinct air masses, in particular near the tropopause. Moreover, the tagging of air masses with PFC tracers can significantly support Lagrangian experiments to study chemical process and aerosol transformations. In addition, investigations of transport and multiphase processes in convective clouds can be facilitated by tagging the air with a PFC tracer in the inflow region and sampling the tracer in the outflow region."

Overall the paper is well written and presents a tracer release system of unprecedented precision and I recommend publication subject to minor revisions.

Technical Corrections

1. abstract: "forecasted" -> "forecast" in both occurrences Both have been changed in the text.

2. p.6793: "parts per quadrillion" -> 10-15 Changed in the text.

3. The background is very low for the PMCP released, but the analytical system has great precision. Is there a possibility that PMCP release could increase the background during a campaign with several 30kg releases? Is this an insignificant contribution to global warming impact?

The present total atmospheric burden of PMCP in the atmosphere is about 500 t. Thus, a release of 100 kg PMCP would increase the atmospheric burden and background on a long term by about 0.02%. For the campaign region and period no increase should be expected since the air masses tagged with PFC will be advected out of this region in 2-3 days. Furthermore, different PFC tracer in subsequent experiments can be used to exclude the risk an interference of the tracer releases.

Therefore the release unit consists of three stainless steel tanks, allowing the release of three different PFC tracers (each ~10 kg maximum). Of course, as small amount of tracer as possible should be released, just sufficient to achieve the experiment aim. Usually the release of 10 kg PFC is sufficient taking advantage of the high sensitivity of the analytical techniques.

Regarding a potential contribution of the released PFC tracer to the anthropogenic global warming, the release of about 10 kg tracer represents an insignificant contribution. We have added this information in the "Introduction" (P6793, line 27): "Watson et al. (2007) estimated that the present total PFC radiative forcing (RF) is approximately 2×10^{-5} Wm⁻². A PFC release into the atmosphere of 10 kg during a tracer experiment represents a negligible greenhouse contribution (~5 × 10⁻¹¹ Wm⁻²) compared to the combined anthropogenic RF (1.6 Wm⁻²) estimated by the 2007 IPCC (International Panel on Climate Change) fourth assessment report."

4. p.6796: "flow at the ground" Changed in the text.

5. p. 6806: "shown in panels a and b"

Changed in the text. As well, in line 8, "green oval in case a" has been modified to "green oval in Fig. 8a and 8b".

6. Sections 4 and 5: some high concentrations were observed outside the bounds of the modelled tracer plume and conjectures were made to the structure that might have been there in reality. In section 4 it was stated that the tracer plume was underestimated in the NE-SW direction (along the filament) and at a later time "probably dispersed more in the SE direction". In the conclusions it was stated that the "plume simulated by HYSPLIT is somewhat narrower than the one inferred from the measurements". I think here you meant to say that it was shorter in the along-filament direction (rather than narrower which would imply thinner in the across-filament direction).

However, a key aspect of the Lagrangian simulations at 5 hours is that the tracer plume is a simple ellipse which has elongated slightly but has not folded. Even by 25 hours after release the tracer distribution is elongated in one direction only and the structure is very simple. However, it should be noted that the resolution of the advecting wind field is only 1 degree and the longest axis of the tracer plume in Fig.8 is only 0.35 deg and only 2 deg by 25 hours. Although deformation in this relatively coarse windfield is felt it can only result in stretching of a tracer blob into an ellipse. Folding of tracer structures, characteristic of chaotic advection, could only occur once the tracer filament length exceeds the advecting wind field resolution. Therefore, it is likely that any variation in the wind field on scales less than 1 degree, combined with temporal variation, will result in stretching and folding of the tracer distribution on smaller scales than those seen in Figs. 8 and 9. Therefore, the main deficiency of the simulation is that the tracer plume is not sufficiently stretched or distorted, even at 5 hours. This is not the same as concluding that the tracer needs to spread further (as a result of stronger turbulent mixing). The authors should discuss these likely deficiencies in the simulations. Given this discussion, the correspondence in concentration between observation and model is remarkably good at 5 hours.

In the conclusions section, page 6808, the sentence has been modified to "plume simulated by HYSPLIT is somewhat smaller than the one inferred from the measurements", and the corresponding discussion is added in page 6807 (after the second paragraph): "The plume simulation at 5 hours after the release underestimates the along-wind size by some 20% (Fig. 8b) and the cross-wind size by a factor of 1.5-2 roughly (Fig. 8c/d) compared with the size inferred from the locations of the samples taken with the ATS. This is consistent with an overestimation of maximum concentrations in the dispersion simulation compared with the results from PERTRAS for the crossings at 7:10 UTC and 7:55 UTC, implying an underestimation of dispersion and mixing by HYSPLIT. However, at this time the size of the plume is significantly smaller than the grid spacing of the driving wind field data, which might result in an underestimation of the stretching and folding of the plume by local sub-grid scale wind variations. In addition, an underestimation of the turbulent velocity components for that specific meteorological condition cannot be ruled out. As a consequence, preparation and execution of future tracer release experiments would strongly benefit from the use of higher resolution meteorological forecast data, if available."