Major concerns

As indicated above, the major concern with this contribution is that some of the conclusions are presented as universal truth while they are not backed by the material presented. These conclusions are all related to the choice of a Tikhonov-Phillips (TPh) regularization that would be more appropriate than OE, e.g.:

– T-Ph regularization it is more suitable for long-lived gases with a pronounced trend
– the profiles retrieved by the OE method are less reliable, showing distorted profile shapes when compared to T-Ph results

Although the T-Ph approach and its tuning are discussed in length in the manuscript, important information is missing regarding the OEM retrievals. No details are given (or even summarized) on the choice of the many OEM parameters, on the resulting information content (DFS), on the error budgets. And what about the efforts deployed and the indicators used to validate the definition of the OEM a priori states for the three targets? A paper by the same authors listed in the references (Polyakov at al. , 2018) might provide these details, but unfortunately this work is not freely available to me.

If we miss this kind of information, we cannot be convinced that T-Ph is necessarily a better option than OEM. It might well be that non-optimum OEM results are compared with carefully tuned T-Ph products. Then it is potentially comparing apples and pears in Figures 2 to 4.

Really, Polyakov et al. (2018) did not analyze the dependence of solution result and parameters on apriori matrix. They used matrices most suitable for a retrieval solution at that moment. We included the description of matrices considered in the text. And, yes, we agree, it is not correct to compare the retrievals with finely optimized T-Ph parameters with that with more general apriori OE matrices. Thus, we cannot make a conclusion on the T-Ph approach advantage. We made changes in a discussion of the figures and the table as well as in the conclusions. But we presume, T-Ph method must be a best solution, therefore, we analyzed it in the manuscript. Optimization of OE matrices can be an interesting problem but it is out of a scope of the manuscript. We included in the text the information about approach used by Polyakov et al. (2018).

In earlier studies, Polyakov et al. (2018) used the OE method (Rodgers, 2000, p. 65) for solving the inverse problem, i.e. apriori information was given in the form of a mean profile and a covariance matrix with 0.05 relative uncertainty at all altitudes and a correlation coefficient of 5 km built on the basis of the WACCM dataset for the period of 1980-2020.

Regarding these figures, it is also important to keep in mind that total columns are retrieved, or a single piece of information (DFS is merely larger than 1). Then the shape of the retrieved profiles is not very relevant, especially because the computed error bars affecting each single (and meaningless taken alone) mixing ratio are so large that they define a broad range of possible solutions with an identical total column. A profile with a more realistic form could likely be drawn without exceeding the resulting uncertainty ranges.

A shape of retrieved profiles also made a sense for total columns due the dependence of absorption on temperature and pressure that change with the altitude.

Another reason given by the authors for selecting T-Ph against OEM regularization is the lack of supporting geophysical information needed to build a covariance matrix: ”it should be taken into account that the OE approach requires the use of the covariance matrices for describing the variability of the target gases profiles, preferably the real covariance matrices, that are unavailable for the considered freons”. However, 16 years of ACE profiles are available, they could be used (and consistently extrapolated down to SPB altitude) to construct it, including for the extra diagonal elements. MIPAS data products are also available to the community (see Chirkov et al., 2016). Then at least two multi-year data sets with global coverage.

These data indeed have a high scientific value, but they cannot be used for constructing covariance matrices for several reasons: 1) The data of limb or solar occultation usually do not allow information on atmospheric gases below 6-7-8 km, whereas the target gases are mainly located in the troposphere. The use of extrapolation to construct covariance matrices to altitudes with the largest content of a gas makes it meaningless to use the measured profiles. 2) The horizontal resolution of these measurements is of hundreds (typically 300-500) km, thus satellite measurements are in a poor agreement with local vertical profiles, for which the covariance matrix is needed. 3) The construction of local matrices for St. Petersburg is complicated by a small number of ACE-FTS measurements for this location and a short period of MIPAS measurements. We indicated these reasons in the text.

OE approach, with the apriori information given in a form of normal distribution of the target gas profile, requires the use of the covariance matrices for describing the variability of the target gases profiles, preferably the real covariance matrices, that are unavailable for halocarbons considered. Although there are datasets of available satellite measurements (ACE-FTS and MIPAS) of halocarbons, they cannot be used for constructing covariance matrices for several reasons: 1) The data of limb or solar occultation usually do not provide information on atmospheric gases below 7–8km, whereas the target gases are mainly located in the troposphere. The use of extrapolation to construct covariance matrices to altitudes with the largest content of a gas makes it meaningless to use the measured profiles. 2) The horizontal resolution of these measurements is of hundreds (typically 300–500) km, thus satellite measurements are in a poor agreement with local vertical profiles, for which the covariance matrix is needed. 3) The construction of local matrices for St. Petersburg is complicated due to a small number of ACE-FTS measurements for this location and a short period of MIPAS measurements.

Second, OEM regularization could make use of ad hoc parameters such as to determine the a priori states. This might sound as artificial, but it is not that different than setting up the alpha for T-Ph. It is true however that it could be trickier, because several parameters are available: the per-layer a priori covariance, the type of inter-layer correlation (Gaussian, exponential), the correlation length.

Yes, indeed, but this choice is difficult to approve. This approach was used by Polyakov et al 2018.

Note by the way that for long-lived and well-mixed gases, it could be relevant to set up a length of several kilometers in order to avoid oscillations in the retrieved profiles.

Polyakov et al. (2018) used a value of 5 km for length.

Finally, it should be noted that the gases involved in the present study do not present "pronounced trends". CFC-11 and CFC-12 will have only changed by about 5% after 10 years. Even if it is larger for HCFC-22 (25% over 2009-2019), this remains less than the changes observed for some other FTIR products over a single year. For example, ethane presents total columns varying on average by more than a factor two over a season, and still OEM regularization is fully and successfully applicable.

We used an incorrect word. We should use "reliable trends" instead of "pronounced trends". But it is more important to point out that trend estimates are one of the aims of the measurements analysis. The "dragging in" the retrieval to the mean (apriori) profile for measurements with low information content (DFS<1.5) certainly influence the trend estimate.
We remove the phrase and change it to:
In the current study, we used the Tikhonov–Phillips (T–Ph) approach to improve the retrieval strategy as in contrast to the optimal estimation (OE) method it does not "pull" the solution to the apriori profile for measurements with low information content. Moreover, the T-Ph approach allows more stable results to be derived than the OE method (e.g. Senten et al., 2012).

Addressing the concerns detailed above could in my opinion be done in two ways. First, provide some information on how the OEM regularization was optimized, and supply elements allowing the reader to evaluate and compare on more solid grounds the data products, in terms of uncertainty, daily scatter, DFS: : :
We pointed the matrices of Polyakov et al (2018) parameters, that allow reader estimation valuable of the comparisons.

Or to significantly reduce section 2.3 and to give focus on the valuable derived time series and the comparisons with other data sets.
And we shifted the focus by removing a statement on advantage of T-Ph regularization.

In any case, one should not be left with the message that T-Ph regularization is necessarily the good option for the retrievals of CFCs or HCFCs. Such conclusions cannot be reached with a single site study, especially when involving challenging conditions. It remains to be demonstrated that it would also be the case for other FTIR stations. So adding "for our site", "in our case" at some selected places in the text could be appropriate to temper the argumentation and make it more specific.
We agree, this is true only for our site. We did not suggest using the strategy developed directly at other sites. Strategy may be used as a part of optimization and development of the unified retrieval strategy for all sites, for example, as a first guess.
We corrected the text defining that the technique developed can be used as a basis for developing the retrieval strategies at other sites, i.e. in the conclusion:
The retrieval techniques suggested may be used in further development of unified strategies for halocarbons retrieval at the sites of the NDACC observational network that use the Bruker IFS125HR spectrometers for IR solar spectra measurements.

Second order issues and minor comments
Page 2
I think the one-sentence description (starting line 42) of the impact of halogenated source gases on the formation of the ozone hole is a bit oversimplified and may deserve an additional statement informing about the roles of the stratospheric reservoirs of chlorine and of the heterogeneous reactions in chlorine activation. The gas phase chemistry does not explain the massive polar ozone depletion, and the photolysis of CFCs does not happen in the Polar Regions since there is no high-energy UV photon available there. Another option is to identify and include a good reference.
We corrected the text:
Although the major content of these gases is concentrated in the troposphere, in the equatorial region the global circulation moves them out into the lower and middle stratosphere and transports to high–latitude regions. In the stratosphere, CFCs are photochemically decomposed to chlorinated free radicals (Cl, ClO) that are deactivated into chlorine reservoirs HCl, ClONO2,

and HOCl (WMO, 1985, Chapter 3). In polar regions, the heterogeneous reactions on the surfaces of polar stratospheric clouds and cold sulfate aerosols convert inert reservoir molecules into active forms that photolyze producing free radicals and cause the chemical ozone depletion in spring through catalytic cycles resulting up to the appearance of so–called ozone holes (Solomon et al., 2014).

Line 45: the Montreal Protocol does not act directly on the halocarbons emissions. Instead, the production of the relevant gases are limited and then banned. I suggest replacing “emission” by “production” on line 45.
Done
As the result of the Montreal Protocol and its amendments and adjustments that restricted the production of chlorofluorocarbons …..

Line 61: this sentence is somewhat misleading: the phase-out (100%) of the CFCs was decided in 1992, for a complete implementation by the end of 1995 (Copenhagen Amendment). In 1989, only a reduction was enforced by the initial treaty.
Corrected
The Montreal protocol from 1987, which came into force in 1989, limited production and consumption of CFCs. Later on, in 1992 in Copenhagen and in 1995 in Vienna, phase-out of CFCs was stated by the end of 1995 in developed countries and by the 55 end of 2010 in developing countries.

Line 62: Brown et al. 2011 is a good reference, but more recent trends have been published by the ACE team, considering now 16 years of measurement and improved versions of the data. I strongly suggest considering here and each time it is relevant these updated results (see Bernath et al., 2020), also for the trend comparisons.
We corrected the text.
ACE-FTS satellite measurements in last 16 years (Bernath et al., 2020) illustrated the success of the Montreal Protocol by estimated decreasing trends in CFC-11 (-0.53 % per year) and CFC-12 (-0.61 % per year) abundancies and a slowing rate of increase in HCFC-22 abundancies (1.8 % per year).

Lines 63-66: it also depends on the evolution of the bromine and nitrogen stratospheric loadings!
We corrected the text.
The time needed for recovery of the ozone layer **among other factors** depends on the sustainability of the reduction in the concentration of CFC-11, CFC-12 and other halocarbons in the atmosphere.

Line 74: to my knowledge, the HFCs are targeted by the Kigali Agreement, not the HCFCs. HCFCs regulation is under the earlier amendments or adjustments. Please check and amend if needed.
Yes, Corrected
Under the Montreal Amendment (1997) all countries must gradually phase down HCFCs.
In September 2007 it was decided to accelerate the phase-out of HCFCs. Developed countries had been reducing their consumption of HCFCs and completely phasing them out by 2020. Developing countries agreed to start their phase out process in 2013 and are now following a stepwise reduction until the complete phase-out of HCFCs by 2030.

Line 83: do you mean down to the Earth surface?

Of course, down *, corrected*

Line 86: it might be true that the publications on the subject were rather episodic, but not the measurements which were continuously performed and exploited at some of the NDACC sites. For example, halocarbon FTIR time series have been systematically included in the successive editions of the WMO assessments on ozone depletion.
We took this remark in account, added:
Time series of CFC-11, CFC-12, and HCFC-22 TCs above Jungfraujoch station, Switzerland, are presented in periodic WMO reports on Scientific Assessment of Ozone Depletion (e.g. WMO, 2018).

Line 89: note that “freon” is a registered trademark.
We changed the term «freon» by «halocarbon».

Lines 93-94: this is not correct, Prignon et al. (2019) proposed an approach for the determination of total *and* partial columns, for the 1988-2017 time period (three decades), not only for 1999-2018.
You are right, the retrieval algorithm allows deriving not only total column of the target gases but partial columns in troposphere and stratosphere as well. This is a misprint, surely, 1988-2017.
Corrected.
Prignon et al. (2019) proposed a technique for estimating two partial columns and TCs of HCFC-22 at the Jungfraujoch mountain station and corresponding time series of HCFC-22 TCs for 1988-2017 along with the trend analysis for various time periods.

Lines 106-108: I somewhat disagree with the arguments on the situation as to the absorptions and spectral signatures (quoted as low and smoothed). In my opinion, this discussion as presented misses the fact that the situation is very different for the three target gases: CFC-11 is likely the more difficult with a broad feature perturbed by strong water vapor lines; CFC-12 has a stronger and more isolated signature peaking at more than 10% and HCFC-22 presents a narrower feature quite free of interferences (HWHM probably on the order of ~0.05 cm-1), resulting in the possibility to select a less wider micro-window. Of course, the spectral scenes will also be influenced by the latitude and altitude of the station.
Yes, you are right. HCFC-22 is quite narrower. The text is corrected.
Finally, the CFC-12 and at a larger extent the CFC-11 bands have smoothed spectral dependency of absorption that requires to use wide micro-windows for retrieving their abundancies, 2 cm-1 for CFC-12 and not smaller than 30 cm-11 for CFC-11.

Line 126: I guess QHN is another description of channeling? Note the relevant discussion paper on AMTD by Blumenstock et al. (2020). Harmonizing the designations could be helpful.
We changed everywhere the term on commonly used, i.e.
Since this filter was plane–parallel, a parasitic interference arose in it, leading to the appearance of an effect of the optical resonance (“channeling”), see (Blumenstock et al., 2020).
….
The period of the channeling is caused by the material and the thickness of the filter, and in the spectral region 800-900 cm^{-1} it is close to 1.1 cm^{-1}, while the amplitude of the channeling varies from zero to a few percent value depending on random filter positioning parameters.
etc

Line 163: the discussion on the spectral transmission function is interesting and original. If I understand correctly, the selection of the relevant parameters is conducted such as to limit the CFC-11 intraday variability. In the end, do you see any correlation between the water vapor and the CFC-11 total columns?

We added the information on water vapor continuum in the text and the figure in Appendix:

Water vapor continuum makes a significant contribution to radiation attenuation by the atmospheric water vapor (Mlaweret al., 2012). Our calculations have shown that radiation absorption by water vapor continuum in the considered spectral region under conditions of the St. Petersburg site can significantly exceed 50%. For a 30 cm−1 window, the selectivity of 205 continual uptake is sufficient to influence the spectra processing results. To calculate the water vapor continuum, we used a free–distributed computer code (MT_CKD, 2017) and the daily profiles of water vapor independently derived from the FTIR measurements (Virolainen et al., 2017). In the first approximation, the contribution of the water vapor continuum to absorption is proportional to the water vapor partial pressure squared, and it can be detected only in a very humid atmosphere. We estimated the contribution of water vapor continuum numerically by analyzing spectra with and without it considering on the most humid days in 2018: July 29, August 2 and 9. The neglecting continuum absorption in these days led to overestimates of the CFC-11 TCs by an average of 2.9 %. Although this value is less than the measurement error (see Table 4), it systematically depends on the water vapor content. Therefore, it is necessary to consider the water vapor continuum. In Fig. A1 of Appendix A, we presented an example of transmission due to absorption by water vapor continuum

Lines 236-239: the statements regarding the Prignon et al (2019) paper are not correct. These authors also used a T-Ph regularization with alpha=9, minimizing the smoothing and measurement errors as per Steck (2002). Please amend your text accordingly.

Yes, they used T-Ph approach with alfa parameter equals to 9, minimizing the smoothing error and retrieval errors. We corrected the error in the text.

Prignon et al. (2019) reported the higher values of DFS (DFS = 1.97) caused by T-Ph regularisation 285 with parameter alpha = 9 and a low atmospheric water vapor content above the mountain (3580 m a.s.l) site Jungfraujoch.

Line 268: regarding the smoothing error, Prignon et al. (2019) have indeed evaluated it to be small (see Table 1 in their paper).

Thank you, we gave a reference to Prignon et al. (2019), although this estimate was done for the OE method.

Prignon et al. (2019) showed that the smoothing error for HCFC-22 is rather small (0.3 %).

Line 297: is the mean molar fraction (or MMF) another name for the dry air mole fraction (often denoted xTARGET, see e.g., section 2.4 in Barthlott et al., 2015), as used by the NDACC and TCCON communities? If yes, it might be good here also to harmonize the designations.

We changed it on the "Xgas" type and added:

Xgas are calculated by dividing the gas total column by the dry air column.

Table 4 reports much higher daily SD for HCFC-22 than for other gases. Is this a filter 3 effect (before 02/2016)? Or is this just because intraday variability for an unregulated gas might be affected by polluted episodes (excursion above the baseline in in situ surface time series)? Also and with the exception of HCFC-22, the random errors quoted on line 18 are significantly larger than the intraday SD on line 4. Shouldn't they be commensurate?

This question is discussed in lines 300-308. This is due to a large contribution of the uncertainty of the temperature profile for CFC-11 & CFC-12 and the predominance of the contribution of spectral noise for HCFC-22 to the random error.

Trends on Table 5: it is more and more clear that the uncertainties affecting trends are often underestimated because the methods used do not account for the autocorrelation present in the time series (see e.g., Santer et al., 2000). This is particularly critical and becomes problematic for species with small rates of change. And one can certainly expect significant auto-correlation for long-lived gases. Did you account for auto-correlation in your uncertainty estimates reported in Table 5? They appear rather small in both cases (using the Gardiner or Timofeev methods) and your comment on lines 354-355 puzzles me.

We corrected the text:

Gardiner et al. (2008) model the intra-annual variability in terms of a Fourier series, Timofeev et al. (2020) use monthly mean values of the considered period to describe a seasonal cycle. In both methods, trends were estimated by subtracting the seasonal variability from initial time series. In first method, we considered periodicities of 4 month and larger, in second method, monthly mean values accounted for periodicities from 1 month. The estimation of the width of the confidence interval of the trend value for the Gardiner's approach was carried out using the Bootstrap method, for the Timofeev's method, it was calculated on the basis of a theoretical statistical approach. It is worth mentioning that we did not take into account the autocorrelation that could be presented in long-lived Xgas time series. Santer et al. (2000) demonstrated that neglecting of the autocorrelation in time series can affect the trends estimates and underestimate uncertainties, however due to substantially irregular FTIR measurements it was difficult to estimate it.

Figure 8: the seasonal modulations for the various data sets are presented, and it is immediately obvious that surface measurements are much more flat than the others. But do you think that a direct comparison is meaningful? Unlike remote-sensing measurements, surface sampling is unaffected by atmospheric dynamics (tropopause height changes, pressure variation: : :) and the include very high frequency measurements, leading to very robust averages. It might be useful to alert the readers of these specific and different situations.

We added into the manuscript:

There are some fundamental differences between local surface and remote sensing measurements (satellite and ground-based FTIR). First, surface measurements are performed with a high regularity and frequency, resulting in stable averages. Secondly, they are unaffected 480 by variations in pressure and tropopause height. And finally, the surface data used were obtained in close to background conditions. Therefore, Figure 8 demonstrates the low seasonal variability of the GVMR – within tenths of a per cent for CFC-11 and CFC-12 and within 0:7% for HCFC-22.

Typos

Typos were not systematically searched for. I just spotted a few ones. See below. And the list of references was not thoroughly checked.
Line 32: replace FS1125HR by IFS125HR
Line 120: replace FS1125HR by IFS125HR
Line 316: replace HCFH-22 by HCFC-22
Line 319: replace Analisys by Analysis

Done.