Interactive comment on "An objective determination of optimal site locations for detecting expected trends in upper-air temperature and total column ozone" by K. Kreher et al.

The reviewers comments are written in italics, the authors' response in bold.

N.R.P. Harris (Referee #2)

This manuscript contains a description of a statistically based approach to identify measurement sites to detect future trends in (a) free tropospheric and stratospheric temperatures and (b) total ozone. The putative future trends are based on CCMVal-2 model runs, and the variability and auto-correlation derived from historic observations are assumed to remain the same in the future. Future measurements are assumed to be stable with all the certainty being included in the instrumental precision. No allowance is made for drifts or offsets. Based on these assumptions, the time to detect a trend is quantified for globally diverse locations (including current measurement sites) and its dependence on measurement frequency and timing is investigated. This study contains a lot of good work and several interesting results. The temperature part is especially convincing and should provide valuable information for network development. The work should be published. However, I think the presentation study could, and should, be significantly improved before publication.

Thank you very much for your summary and feed-back.

General comments

There are two main aspects to this: (i) more context and discussion of the results; and (ii) increased clarity of presentation. Neither of these need much (if any) new calculation, just a re-thinking of the presentation. These aspects are discussed in turn

The authors currently include very little discussion of the results. The rationale for this (p. 1629, lines 20-24) is that "The purpose of the exercise is to show that generating fields....... provide(s) one objective method of selecting optimal sites....". That is fair enough at one level, but I think any interested reader of this study would want to know more about the other factors that could be considered.

It would also help to have a clearer discussion of the assumptions that have been made. For example, what if the model trends are wrong? How well do they reproduce past trends?

Text to address this has been added in the "Discussion and Summary" section.

What is the effect of specific geophysical features (e.g. broadening of the tropics)? Should some sites be chose with this in mind? What additional information is relevant for knowing an atmospheric change is happening? E.g. change in tropopause height or in vertical/latitudinal shape of trends – or, for ozone, trends in the vertical profile. These are factors that should be considered when designing a network of stations as they could provide additional constraints to the results presented here.

The goal of our paper is not to design a network of stations for detecting changes in atmospheric composition. Much of what the reviewer has requested is therefore well beyond the scope of our paper. The goal of our paper is to present one objective method, which in a real-world application would be used together with the multiple additional considerations that the reviewer alludes to, to select the location for measurements of atmospheric temperature and total column ozone.

Some discussion of possible drifts (how well the network stability is known) and other potential instrumental uncertainties should also be included.

Because our paper is not dealing with a specific network, we cannot make any comment on network stability. This study considers an ideal network rather than an existing network. The issue of measurement drift and other potential instrumental uncertainties is therefore tangential to the scope of this paper. Were we to include analyses of the effects of instrumental drift and/or time dependent systematic biases in the measurements (e.g. those resulting from changes in instrumentation), this would add a whole new dimension to the paper that is currently not considered at all. This would be a topic for a follow-up paper. We therefore cannot see how to accommodate this request by the reviewer.

Precision (which, as the authors point out, has a simple relation with sampling frequency) is an important factor but it is not the only one and may not be the most important.

Our paper does not discuss precision at all and this word appears nowhere in the manuscript. We therefore do not understand this point raised by the reviewer.

I would like to emphasise that I am looking for is a fuller discussion of the assumptions, their implications and the broader debate to which this study contributes.

Specific concerns raised by the reviewer are addressed below.

The clarity of the manuscript should also be improved. Most importantly, the logic of the work needs to be made clearer. Some of the critical steps are described very briefly and were easy to skip over.

We made some changes to the manuscript to improve the clarity and comprehensibility.

Some improvements could be made to the figures so that their meaning is more easily understood. Having said that, I think that Figs 8-10 are very clear and make the main points very well. The build-up to them could be improved though. Some specific comments are given below.

These specific comments have been addressed (details below).

My last general comment concerns the section on ozone trend detection. It is weaker than the section on temperature trends, but that is probably inevitable since only the total column is considered and sampling frequency is not discussed. I am ambivalent on whether it stays in. If it does, then it needs strengthening. There was some early work on measurement strategy (station location, sampling frequency) by Tiao et al. (1990) which should be mentioned as a background and complement to the work presented here. (Note that further study by them (not sure if it was published – it is referred to at the end of Section 4) showed that persistence of weather systems weakened their conclusion that "the precision of trend estimates (is) very insensitive to changes in the temporal sampling rate of daily data", and led to a conclusion that measurements needed to be spaced regularly through the month, e.g. on the earliest possible day of each week.)

G.C. Tiao, G.C. Reinsel, D. Xu, et al., Effects of autocorrelation and temporal sampling schemes on estimates of trend and spatial correlation. J.G.R. 95(D12), 20507-20517 (1990).

This has been addressed in the text and a discussion of the work by Tiao et al. (1990) and Weatherhead et al. (2000) has been included.

At the same time, the final part of this work (1631, 15-24) needs to be described more fully – it feels like an express train at the moment.

We agree with the criticism and have extended the discussion in the manuscript.

Specific comments

Abstract: The current version of the abstract does not exactly entice the casual reader further in to the paper. The problem is in the first paragraph (which reads like an earlyish draft) as the second one is fine. I am loath to make too many comments as it should basically be rewritten, starting with the aims of the paper and then the three stages should be described (probably without explicitly saying there are three stages) along with some of the assumptions. The aim should be to make it shorter. The start of the summary is clear and would be a good place to start.

The abstract has been extensively modified taking the reviewers comments into account.

Section 2.1: The core of the temperature analysis, in some ways, is the second paragraph. I think it need a bit more substance – how good are the NCEP fields?

An assessment of the quality of the NCEP CFSR temperature fields is beyond the scope of this paper. We refer the reviewer to the outcomes of the SPARC Reanalysis Intercomparison Project which is undertaking a detailed assessment of the quality of the NCEP CFSR temperature fields. For the purposes of our study, the key characteristics that the temperature fields need to capture are:

1) **The amplitude of day-to-day variability,**
2) **The amplitude of the diurnal cycle in temperature,**
3) **The structure of the seasonal cycle in temperature,**

at each of the sites considered. We have not conducted an assessment of the quality of the reanalyses in regard to these three characteristics.

; state explictly that it is temperature only

We have inserted the word 'temperature' after NCEP CFSR to make it clear that it is the temperature data from the NCEP CFSR database that is being used to assess the effects of sampling frequency on the random uncertainty on monthly mean temperatures.

and what spatial resolution the measurements are?

The spatial resolution of the temperature fields is 0.5°×0.5° and this is now stated in the manuscript.

How well do they correspond to an in situ radiosonde profile? I assume that many of the radiosondes going in are 12 hourly – if so what is the main influence on the intervening 6 hour temperatures and is the diurnal cycle reproduced well? I guess I worry about the use of the word 'true'.

The reviewer is correct in stating that the reanalyses should not be considered as 'the truth'. The extent to which the 4-dimensional variational assimilation captures the:

1) **The amplitude of day-to-day variability,**
2) **The amplitude of the diurnal cycle in temperature,**
3) **The structure of the seasonal cycle in temperature,**

is what matters for the purpose of our study. However, as stated above, we have not conducted an assessment of the quality of the reanalyses in regard to these three characteristics.

Also, a bit more detail on the Monte-Carlo method would be useful

More detail has been added and the section has been rewritten to clarify the method further.

– should the material at the beginning of Section 2.2 be moved earlier?

We believe it is easier to follow the discussion if the material at the beginning of Section 2.2 stays where it currently is.

Figs 2 and 3: Figure 2 does not show that much which could not be stated in the text

We agree. We have removed Figure 2 and stated the information in the text.

, while Figure 3 has 12 panels in which the most obvious difference is between sampling frequencies rather than sampling time. And not much is said about the differences resulting from changing sampling times in the text. I wonder whether it would make more sense to cut the number of panels in Fig 3 to, say, 6 which show different sampling frequencies and to state in the text that changing the time of day does not have a significant influence on the uncertainty on the monthly mean. The same format could plausibly be used in Figure 2

We have implemented the suggestions and cut the panels in the new Figure 2 and in Figure 3 back to 8 panels only using the sampling frequencies and not the change in time of day. To be consistent, we have also applied this to Figures 4, 5 and 7.

Figure 5 would also be easier to interpret if the x-axis represented different sampling frequencies, i.e. the different sampling strategies were removed.

We agree and have done this (see comment above).

Figs 6 and 7 and accompanying text. It would help the reader if the connection between Figs 6 & 7 could be made clearer. For example, the model trend values for the two cases shown in Fig 7 should be mentioned in the text and/or the Figure caption. This would help give some meaning to equation 1.

The values for the projected trends at the 2 sites have been added to the manuscript.

Fig 7 is presumably meant to build up to Figs 8-10, but I personally find these latter figures much easier to grasp the meaning of (though the assumed sampling frequency should be mentioned in the caption). The main use of the figure it to show the effect of sampling frequency, so again the number of scenarios shown could be reduced.

We agree and have simplified the figure by only using the sampling frequency.