

**Interactive comment on “Observation of horizontal winds in the middle-atmosphere between 30S and 55N during the northern winter 2009–2010” by P. Baron et al.**

**Anonymous Referee #2**

We would like to thank Referee #2 for his comments and suggestions for improving the manuscript. Below we provide point by point answers to the referee's comments and questions.

The new version of the figures, including a new one (Fig 2) are provided in a supplement file. Note that a mistake has been corrected in Fig 7 (now Fig 8). Larger biases between the ECMWF and SMILES winds in the near meridional direction are seen above 0.1 hPa at the latitude 20N. The difference is consistent with the results in Fig 6 (now Fig 7).

— Page 32475

**Lines 14 and 16:** Is it really "mean" differences you're talking about here (i.e., specifically the arithmetic mean of A-B)? If so, the sign is significant and you should clarify which is greater than the other. Alternatively, is this really the mean of |A-B| or some kind of rms (or just a "rough" summary). If so? then perhaps use some word other than "mean" to describe it to avoid potential for confusion.

We meant absolute value of the mean difference ( $|\langle A-B \rangle|$ ). The text has been changed as :

“In the region between 1–0.05 hPa, **an absolute value of the mean difference**  $< 2 \text{ ms}^{-1}$  is found between SMILES profiles retrieved from different spectroscopic lines and instrumental settings. Good agreement (**absolute value of the mean difference of  $\sim 2 \text{ ms}^{-1}$** ) is also found with the European Centre for Medium-Range Weather Forecasts (ECMWF) analyses in most of the stratosphere except for the zonal winds over the equator (**underestimation of ECMWF zonal winds of 5–10  $\text{ms}^{-1}$**  ).”

**Line 20:** Delete "in the stratosphere" (essentially - though not completely I grant redundant with "stratospheric" two words earlier).

Done

— Page 32477

**Line 8:** Please clarify "meteor radars" for the uninitiated. Is this some use of radar to somehow directly measure wind, or is this using radar to track incoming meteorites and deduce wind from their velocities? (Apologies if this is a dumb question).

The given reference “Jacobi et al. (2009)” provides a description of a meteor radar. We thus do not think that it is necessary to add it in the paper. They summarize the principles as:

“The wind measurement principle is the detection of the Doppler shift of the reflected VHF radio waves from ionised meteor trails.” The trails are advected by the wind and thus their movement is a measure of the wind.

More generally, we tried to provide a reference for each measurement technique that is mentioned in the paper.

— Page 32478

**Line 3:** "for deriving trace gas profiles" is awkward in this context - it's not like SMILES would not have been able to make trace gas profile measurements if it had launched in different month or year. I'd move the trace gas part earlier in the discussion.

The sentence has been changed as follow:

"The instrument has been conceived for measuring trace gas profiles and was launched in September 2009."

**Lines 6-7:** I think this issue needs to be made clearer to the uninitiated reader. Unless you know that the ISS orbit is inclined around 50 degrees, and that SMILES looks ~40 degrees left of the forward direction (neither of which is described in the paper that I could see) it's far from obvious to those unfamiliar with limb-viewing sensors. In addition to helping the reader understand by making these points, a little sketch diagram would go a long way to making this clearer. It's a really nice feature of the orbit and viewing geometry - I'd make it clearer to the reader.

We prefer to not change the introduction since only the general principles are addressed in this section. However we agree that this information is missing in the paper and we have added a new Figure which provides it (see Figure 2 in the supplement file). The figure is introduced in the text as follows (P 32480/line 6):

"As seen in Figure 2, because of the geometry of the instrument field of view in relation to the ISS orbit, the line-of-sight winds during the ascending (descending) portion of an orbit are almost in meridional (zonal) direction."

**Line 14:** "operational winds" -> "operational wind"?

Done

**Line 15:** How about "middle stratosphere"? I only suggest this as it sounds better given that you already have "mid latitude"?

Done

**Line 16:** "...expected to be reliable..." Would be nice to quantify this - better than 5m/s? 10m/s? Any literature one could cite?

By "reliable" we mean that we expect ECMWF to have an error better than the SMILES single retrieval precision of ~10 m/s.

The text has been changed as follows:

"At mid-latitude in the middle-stratosphere below 1 hPa, ECMWF winds are expected to provide a wind representation with an error smaller than that of a SMILES single retrieval (~5-10 m/s), and, thus, to provide a good .... "

This sentence is still "vague" but we do not think, because of a lack of observations, that we can provide any number for ECMWF errors from the middle-stratosphere to the lower mesosphere, especially during the same winter as SMILES operations. Actually as mentioned on line 18, one of the conclusion of this paper is to better define the reliability of ECMWF by providing such numbers.

In Section 3 (p 32483), we mention Baldwin and Gray (2005) who focused on tropical winds but the rocket measurements used for the altitudes relevant to the SMILES data stopped in 1983. They found that the correlations between ERA-40 and the rocket measurements strongly decrease from 0.9 to 0.3. between 10 hPa and 1 hPa. They concluded that ERA-40 should be used with caution above 3-2 hPa.

At mid-latitudes, Rufenacht et al. (2012) has presented a comparison in the same altitude range as SMILES using new observations by a ground-based micro-wave radiometer during Winter 2011. No quantitative difference are given, but the conclusion states:  
“Comparing the 11 month long series of zonal wind data over Bern to the ECMWF operational analysis reveals very good statistical agreement”

**Line 18:** "...results are uncertain..." Again, it would be nice to be a little more quantitative here.

See previous answer and the reference Baldwin and Gray (2005) has been added in the text:

“On the other hand, ECMWF results are uncertain in the mid- and high stratosphere of the tropical regions (Baldwin and Gray, 2005). “

**Lines 24/25:** Not clear whether mean ECMWF differences are discussed in section 3 or 4 (sorry to be picky).

Line 25, the sentence has been changed in the text: “In Sect 4., results are illustrated ...”

— Page 32479

**Line 20:** "induced" -> "induces". Elsewhere you use the present tense.

Done

**Lines 19-21:** You quote the typical Doppler shift and frequency resolution, but I think you should also mention the typical line width (though it probably gets wider lower down). I can imagine shifts are harder to spot in wider lines.

Indeed, the width of the spectral line is 40-50 MHz at 10 hPa (30 km) and decreases to ~1 MHz at 0.2 hPa (60 km). Above, the line width roughly remains constant because of the line Doppler broadening. At 10 hPa, the amplitude of the frequency-shift signature becomes weak and it is difficult to distinguish it from small features in the spectral baseline (amplitude is <10 mK for 10 m/s wind). It is the reason why a large bias is found on the retrieved winds below 30 km. On the other hand, at high altitudes, the line intensity strongly decreases because of the decrease of the molecular concentration. Retrievals from the HCl line go higher than that from the O3 line because of the slower decrease of HCl concentration and also because of the presence of 3 relatively strong HCl lines within 30 MHz. Note that between 0.01-0.003 hPa, the O3 concentration strongly increases during night time (O3 secondary peak) which gives enough signal for wind retrieval.

The information has been added:

“... 1.2 MHz and to the full-width-half-maximum of the spectral line that decreases from ~50 MHz at 10 hPa to ~1 MHz at 0.2 hPa.”

and the text has been changed in page 32480/line 20:

“In the mesosphere, the steep decrease of the O<sub>3</sub> concentration is responsible of the loss of the retrieval sensitivity from the O<sub>3</sub> line and the HCl line, though weaker in the stratosphere, becomes more suitable because the HCl VMR profile has its maximum in the lower stratosphere and mesosphere, and the HCl spectrum is composed of three relatively strong spectral lines within a narrow range of 30 MHz.”

— Page 32480

**Lines 1-3:** Presumably there is a latitude dependence also? I presume that is considered in composition retrievals - might be good to mention that.

Yes, the latitude dependence is taken into account. The text has been changed as follows:

“Profiles of geophysical parameters are retrieved from the inversion of the set of spectra that are measured during a single vertical scan of the atmospheric limb. To simplify and fasten the trace gas and wind retrieval calculations, the same line-of-sight velocity is removed from all spectra composing a vertical scan. The velocity which changes from one scan to another, is chosen at the middle of the scan. After the retrieval calculation, the derived wind profile is corrected from the altitude variation of the line-of-sight velocity ( $\sim 0.8$  m/s/km). “

**Line 4:** "Because of the geometry..." As stated above, you haven't actually told us what that geometry is. Again, either here or in the earlier discussion, more details on the orbit and geometry and a figure would be really nice.

A figure has been added (see Figure 2 in the supplement file), and the text has been changed as follows:

“As seen in Figure~2, the geometry of instrument ....”

**Lines 13-14:** "should have been" -> "are"? "Should have been" makes it sounds like they aren't and you don't understand why.

Done

**Line 17:** While it's good to have the citation to Merino et al., a few more words on what "measurement sensitivity" means would be good. (I'm guessing it's the sum of the averaging kernel rows).

The sentence has been changed to:

“Figure 4 (left panel) shows the vertical resolution of the retrieved wind profiles and the measurement responses, i.e., the sum of the averaging kernel rows (Merino et al., 2002) that indicates the altitudes of good measurement sensitivity. Good sensitivity (defined as where the measurement response ranges from 0.9 to 1.1) is found from 25 to 70 km”

**Line 18:** The sentence beginning "Considering altitudes..." is a bit clumsily worded. I suggest changing to "Good sensitivity (defined as where the measurement response ranges from 0.9 to 1.1) is found ..."

We followed the Referee suggestion and revised the sentence.

**Line 23-24:** I presume the difference in HCl and O3 reflects their different abundance profiles (with O3 peaking in the lower stratosphere and HCl being maximum in the upper stratosphere and lower mesosphere). It might be nice to add a few words to that effect (or if my guess is wrong, to give whatever the reason is).

See answer of the comment related to page 32479/lines 19-21.

— Page 32481

**Lines 1-2:** This discussion is unclearly described.. Was the non-linearity not considered at all in Baron et al., (2011), or was some different value chosen? If the latter, then why choose 20% here when another value was used before?

The version of the level-1b data (calibrated spectra, version 5) available at the time of the analysis by Baron et al. (2011) assumed a linear relation between the intensities and the the instrument output. A correction for a non-linear instrumental response has been added on the level-1b data (version 7) used in this analysis.

The sentence has been changed as: “ which include a correction for non-linearity in the receiver which was not applied in the previous analysis (Ochiai et al., 2012a). An error of 20% ...”

**Line 8:** "lower limit of accurate retrieval" is awkwardly phrased. Also, any way to be more quantitative, what value of accuracy did you consider as being the worst tolerable?

As stated in the next line of the text the accuracy for a good retrieval is considered to be 5 m/s.

The sentence has been rephrased as follows:

“The accuracy of the retrieval at lower altitudes is set by systematic effects on the O3 line retrieval ...”

**Line 16:** Your description of chi-squared is at odds with the typical definition (and probably not what you actually did). Statistically, chi-squared is essentially defined as the sum of the squares, as you describe, but importantly it is divided by the estimated error in the radiances. Also, most people in the community typically (but, strictly speaking, incorrectly) divide by the number of measurements (or some similar factor). I’m guessing that’s what you actually did here. If so, you should reword the description. If not, then I’d call it something other than chi-squared.

The Referee is right. The correct definition of the chi2 we are using is given in Baron et al, (2011, Eq 2). The text has been changed as follows:

“... from the sum of the squares root of the spectral fit residual weighted by the corresponding measurement errors and divided by the total number of measurement (Eq 2 in Baron et al, 2011). For regularising the inversion, an a priori knowledge of the winds is included in chi2 (see Appendix A).”

Also, I’m unclear about the "before the retrieval" chi-squared. You need to be clear about what starting guess you’re using for the state vector here, as that is what largely determines this value. I expect it’s something based on a "dead reckoning" tangent point altitude, but it should be described. I’m not clear what "disturbances in the field of view" are? Is this things like obstructions (I know the ISS solar arrays

get in the way from time to time), or is it something else? As to "[disturbances in] the pointing", won't the retrieval take them out, in which case why does the "before" chi-squared matter?

The description of the initialisation of the retrieval process has been improved in the Appendix A following the Referee recommendations. We believe that it is not necessary to provide details on the retrieval computation in the section 2.3 since we simply want to describe the different data rejection tests that we have applied.

Disturbance in the FOV mostly means ISS solar panel or any astronomical objects like the Moon. Disturbance of tangent heights means some imperfect attitudes data because data from the ISS guidance, navigation, and control system which does not represent correctly the attitudes at the position of the JEM module due to the deformation of the ISS. Most of the bad scans are filtered out before or after the first retrieval steps. However some are still present at the wind retrievals.

Note that the problem related to the SMILES attitude estimation is explained in the following reference which has been added in the list of references:

*Ochiai, S. et al.: Tangent height accuracy of superconducting Submillimeter-Wave Limb-Emission Sounder (SMILES) on International Space Station (ISS), pp. 1290–1293, 10.1109/IGARSS.2012.6350824, <http://igarss12.org/>, 2012c.*

The Appendix has been modified as follows:  
(page 32494, line 19)

“Firstly the profiles of atmospheric constituent relevant for the spectral region, temperature and spectrum tangent-height are retrieved disregarding any spectral shifts due to the wind which has no significant impact on the results. Then the wind is retrieved on a 5 km vertical grid by initialising the inversion using the parameters retrieved in the first step. The vertical sampling of the retrieval grid is consistent with the information content of the spectra. Note that in the initialisation step, baseline parameters, i.e, additional radiance offsets and slope used to improve the measurement fit, are set to 0. Since, a small baseline is usually retrieved for good SMILES scans, the initial chi<sup>2</sup> (Baron et al, Eq 2) has a relatively small value between 20--40.”

The link to the Appendix has been added to the text p 32481 (117):

“... in the field-of-view or in the pointing (see Appendix A).”

**Line 27 (to line 1 of next page):** I'd put "from 25km to 16km" in parentheses.

Done

Also, why did you not use something like the ERA (or MERRA) dataset that would, presumably, not be subject to the resolution discontinuity that you talk about? If it's just that this is what you had to hand then that's fine, but you should probably say so.

The Referee guess is right. The choice of the operational data was because they were available when

we did this analysis. However the main differences between the ECMWF and SMILES are persistent over the full period indicating that the versions of ECMWF we used has no impact on the conclusions of this paper (see also the comment 8 (p32480/9) from Referee 1 and the additional figure 14).

— Page 32482

**Line 6-7:** Do you do anything about the smearing of the wind measurement along the line of sight (I think you mentioned 500km earlier)? That's a significant number of ECMWF grid points. Some kind of weighted average along that length (though not at a fixed height I guess - tricky) might improve some of your comparisons. Again, if you didn't do this, a quick mention of that, and why it can be neglected is fine by me.

We did not do any smearing of the ECMWF data. Without smearing, we expect a small increase for the standard-deviation of the difference SMILES - ECMWF, especially for the zonal-wind at high latitudes where winds have large variations. This effect is mentioned in pages 32486 (15) and 32487 (18).

The bias due to this effect should be small when looking at averaged data and will not change the results presented in this analysis.

**Line 10:** Delete the comma after "retrieval"?

Done

**Line 16:** Delete "with a direction"

Done

**Line 19:** At first, I was going to ask you to roughly quantify how close to "zero" this is by using ECMWF. However, I see later (lines 15-16 of the next page) that you not only do that, but correct for the small remaining wind ECMWF says you have too. Is there anyway to reword this discussion to encompass that point from this early stage?

The following sentence has been added in line 17:

"... is predominantly zonal. Since the meridional wind is not truly zero, a similar zero-wind computed with the paired ECMWF data is added to the observed zero-wind."

The sentence page 21483, line 15 "In order ..." has been removed.

— Page 32483

**Line 20:** Delete "data since" and "are"? Done

— Page 32484

**Lines 20-24:** Is there anything you can say about why the SMILES standard deviation is so much more than ECMWF? Is this real? It seems that it's more than can be accounted for simply by the SMILES precision? Is that correct? Some discussion would help.

Since the current section deals only with the inter-comparison of SMILES winds we prefer to keep the text as it is and do not mention the ECMWF data here. The comparison with ECMWF data is addressed in the next section (Sect. 3.2).

Note that it is expected that the SMILES standard deviation (which include both the winds variability and the measurement random errors) is significantly larger than that of the ECMWF meridional wind in the tropics (see next sub-section about SMILES and ECMWF difference, p32486, line 5-8). In the mesosphere the ECMWF standard-deviation is expected to be smaller than the actual wind variability because of an under estimation of the waves and tides (see answer to Ref 1 comment: 14. p.32486).

— Page 32485

**Line 4:** "overestimation" is unclear I'm afraid. Precision, resolution and accuracy are unfortunate words. Qualitatively, English makes it sounds like more is better, but quantitatively, more is worse. So "overestimation" is jarring. I tend to use words like "better/worse" or "optimistic/pessimistic". Also I don't understand why the large variability of the mesospheric O3 line intensity would lead to an "overestimation" (i.e., pessimistic estimate?) of the precision. Did you overestimate the O3 intensity variability in your error budget? Sorry to not be following this fully.

We agree with the Referee about the fact that the sentence is unclear and we changed the text as follows:

“The standard- deviation ( $16 \text{ ms}^{-1}$ ) is smaller than the standard-deviation expected from the theoretical errors (root-sum-square of the errors is  $24 \text{ ms}^{-1}$ ) indicating a pessimistic estimate of the theoretical error. The actual mean precision of the winds derived from O3 is likely between  $12\text{--}15 \text{ ms}^{-1}$ . At 70 km (0.05 hPa), ozone has strong diurnal and latitudinal variations (Kasai et al., 2013) which induces a large variation of the O3 line intensity of more than an order of magnitude, and therefore, of the wind retrieval precision.”

**Line 23:** Having "estimate" and "underestimated" in the same sentence is awkward. Again, using words like better / worse would be clearer, and avoid that near repetition.

The sentences have been changed as follows:

“The actual precision is significantly worse than the theoretical computation because of a residual error after the zero-wind correction. An additional noise of  $5 \text{ ms}^{-1}$  likely arises from the fluctuation in the spectrometer frequency errors.”

— Page 32486

**Line 14:** "less" -> "fewer": Done

**Line 19:** "scaled" - surely you mean "offset" don't you? If not, then I've clearly misunderstood the zero wind issue. Please clarify.

The Referee is right. The sentence has been changed in text:

“ ... should be taken with caution since the mean ECMWF tropical meridional wind is added to the zero-wind.”



— Page 32487

**Line 14:** "< 2ms-1" is a little bit confusing. Could we possibly say "> -2ms-1"? Sorry, I must sound really fussy. On the other hand, it is more consistent with the way you describe the other biases later in the same paragraph.

The changed “< 2ms-1” by “between +/- 2 ms-1”

— Page 32489

**Line 24-28:** See the discussion lower down on figure 2. In fact, I'd suggest you make this a separate figure. I'd show N2O in a polar projection with combined wind vectors overlaid.

See our answer to the comment on figure 2.

— Page 32490

**Line 11:** Actually this is quite an important one. You quote this as a 2-sigma error. Have all the errors up to this point been 2-sigma also? If so, you should make that much clearer earlier on. Why 2-sigma? I can perhaps understand for accuracy-related errors, as that gives you a ~95%-confidence range. However, for precision-related errors, most people quote one-sigma. If this is the first time you've used 2-sigma then why?

In the text all the errors are 1-sigma error but the plot shows error bars corresponding to 2-sigma error (simply for better visualization of the small error bars). Following the comments of the Referee and to avoid any confusion, we changed the plot to show 1-sig error bars.

— Page 32491

**Line 14:** Add "in" between "variation" and "amplitude": Done

— Page 32492

**Line 9:** Please give a citation for the equatorial symmetry of the QBO and SAO? Has that been shown from some kind of observations? Is it a model-based or theory-based assertion? How symmetric is symmetric (i.e., what is a typical wind difference betwpolar projection een corresponding north/south latitudes)?

The statement in the paper is wrong.

In the stratosphere the SAO signal in the zonal winds is not symmetric about the Equator and it is the reason of the asymmetry seen in SMILES data (also seen in past observations). As for the QBO, old and sparse balloon observations depict an Equatorial symmetry of the QBO zonal winds but newer analysis slightly modified this point of view. In particular, Hamilton et al (2004) reported some differences that can exceed 10% in the QBO zonal-wind amplitude at 10 hPa for westerly prevailing winds.

We changed the text as follow:

“Unlike the QBO signal in the zonal wind which is nearly symmetric about the equator (Hamilton et al., 2004), rockets and HRDI observations have reported an asymmetry in the upper-stratospheric SAO with a stronger amplitude in the southern tropics (Hirota, 1980). This is consistent with the SMILES observations.”

— Page 32494

**Line 16:** Any citation other than a web page for the SMILES research chain?

The NICT processing chain is introduced in Baron et al, 2011 (already cited in the paper). However this paper is not up-to date and the information given in the web page should be preferred.

— Figure 1

What does the color of the bar signify (e.g., why is SMILES in red - is it simply because it's what you're talking about). A better way of quantifying the resolution would be nice. Perhaps having more ticks on the x-axis. Alternatively, having thin horizontal lines at selected heights in the "bars" with little legends quoting the precision above/beneath.

The difference between the clear and dark blue is explained in the caption: dark blue is for theoretical errors while thin blue is for actual errors). The red color for SMILES has been removed (it was to highlight the subject of the paper). Dashed vertical and horizontal lines have been added in order to better indicate the vertical and horizontal ranges.

— Figure 2

I actually don't like this figure. I don't think splitting the two components of the wind field is good if you're then going to show them with arrows - the arrows are not actually pointing in the direction of the wind (unless you're lucky and the other component is zero)! Also, does the length of the arrow mean the same wind speed at all latitudes? Further, why have N2O here at all, you don't talk about it till way later. I'd make a separate figure (on a polar projection) for the stratospheric sudden warming discussion. In any case, what do the white values for N2O mean? As it is, the arrows are essentially invisible in the low latitudes and then very cluttered and hard to interpret (not least because they are split into two components and thus the direction is really not meaningful) in the higher latitudes.

Figure 2 is now Figure 3

The purpose of this plot is to show examples of single retrieval line-of-sight winds and the observation geometry along the orbits (it is complementary to the new Fig 2). We modified the caption to clearly indicate that the arrows indicate LOS winds. Splitting the ascending and descending orbits (and thus meridional and zonal observations) allows us to avoid the confusion between LOS winds and real winds. The relation between the length of the arrows and the LOS wind amplitude is the same for all latitudes (this statement has also been added in the figures). Since we do not combine the different LOS to create wind vector in the analysis, we think it is better to not show such data in the plot and to only show the LOS winds. Note that meridional and zonal observations have 6 hours difference.

The N2O background helps to interpret the large LOS wind variation by showing the position of the Vortex. It is used for the discussion on the SSW in the last section.

— Figure 3

Figure 3 is now Figure 4

There is a lot of wasted white space on this figure. I'd delete the redundant y-axes and make the plots closer. Also it looks like it's been compressed along the x-direction (the fonts are very tall and thin). Why have the "W-" prefixes for each entry in the legend? You've not used this notation elsewhere.

'W' is for indicating that we show results for a wind profile retrieved from a given line and not the retrieval of the constituent profile itself.

— **Figure 4**

Figure 4 is now Figure 5 (see supplement file)

I'd have an extra color and use it to differentiate the AOS-1 and AOS-2 cases for the O3 band A measurement. Also, what are the error bars? (Your e-bar term?)

Done. The error bars correspond to the 2-sig line of sight retrieval errors. As mentioned previously the bars will be change to represent the 1-sig. Compared to the original figure, the ECMWF winds has been subtracted from the retrieved zero-wind to be consistent with the analysis.

— **Figure 5**

Figure 5 is now Figure 6 (see supplement file)

The tall-thin nature of these plots makes them hard to take in. Also, I'd decrease the number of vertical dotted lines if possible. Why is the red line dashed? (Simply fo clarity?)

The plots have been increased and the red dashed-line has been changed to a solid-line as for the other lines.

— **Figure 7**

Figure 7 is now Figure 8

A lot of wasted space, delete redundant axes (both x and y) and color bar (one horizontal color bar under each column) and tighten up.

Figure has been changed (see supplement file).

— **Figure 8**

As figure 7, also include the +/-10 degree caveat in the caption, as in Figure 7.

Figure 8 is now Figure 9

Figure has been changed (see supplement file).

— **Figure 9**

Again, a lot of wasted space, getting rid of it would give larger and thus more readable plots. Also is this true zonal wind (i.e. you've done a vector sum of the two components and then taken the zonal projection) or does your +/-10 degree caveat apply.

Figure 9 is now Figure 10

Figure has been changed (see supplement file).

— **Figure 10**

Again what do the error bars mean. Also, reduce the number of y-ticks and labels in the lower plot to avoid clutter. It's going to be hard to read in single column format as it is now.

Figure 10 is now Figure 11

Figure has been changed (see supplement file).

— **Figure 11**

The font for the color bar is tiny compared to that elsewhere in the figure. Also, the y-axis label has got ever so slightly cut off.

Figure 11 is now Figure 12

Figure has been changed (see supplement file).

— **Figure 12**

Again much wasted space that could be used to make bigger plots. Also again, what are the error bars

Figure 12 is now Figure 13

Figure has been changed (see supplement file).