30 September 2013. Review of Staufer et al: Trajectory matching of ozonesondes and MOZAIC measurements in the UTLS - Part 2. .... global ozonesonde network

SYNOPSIS OF PAPER

Using a method of trajectory mapping developed in an earlier paper (Staufer et al., AMTD, Part 1, 2013), comparisons are made among 28 ground stations measuring ozone and reporting to WOUDC, and ozone from any MOZAIC flight that sampled air linked to a sounding either before (via backward traveling air parcel) or after (forward trajectory) the instantaneous sonde reading was made. *The purpose of the comparison is to use an independent set of ozone measurements, supplied by regularly calibrated spectrometers [on MOZAIC] to evaluate consistency in time, vertical distribution, etc, of the ozonesonde data, with an objective of 'quality assurance' of the latter for trends analysis.* Ozone measurements from soundings (a Brewer-Mast or ECC sonde) and aircraft uv photometer are averaged in 50 hPa segments. Tropical to high-latitude data are examined; numbers of samples vary widely depending on regional coverage of balloon launches and aircraft flights. In Europe, where sonde coverage is most spatially and temporally dense and where MOZAIC aircraft are based, there are thousands of sonde-flight coincidences; in more remote regions, especially in the southern hemisphere there may be fewer than 100 matches. The sonde-MOZAIC agreement averages 5-15% although it is worse than that range for older Brewer-Mast design sondes than for ECC sondes. In other published studies, ECC sondes of different type (manufacture) or sensing solution composition have been found to have systematic offsets one to another within the lower atmospheric segment of ozone profiles. These influences were not generally apparent in the MOZAIC comparisons. Ie when an instrument change was made, agreement between sonde and aircraft did not vary.

OVERALL RECOMMENDATION & COMMENTS

The study is of great interest to a large community for two reasons. (1) It is more comprehensive than prior sonde-MOZAIC comparisons; (2) By using trajectories to match air parcels sampled by aircraft and sondes, it attempts to go beyond historical comparisons of MOZAIC profiles and nearby sonde station data (cf Thouret et al, JGR, 1998). (3) Results in Figures and Tables are good. The paper should eventually be published, with the authors addressing the questions raised here.

There are several overall deficiencies in the paper, however. In its present form it fails to achieve the goal of demonstrating that the sondes over all the stations examined, are suitable for trends analysis. Here are the main reasons the paper falls short.

First, the Abstract does not convey clearly the 'framing' and motivation of the study. The reasons for conducting the comparisons are not given. For example, a conclusion is

reached about trends being derivable from the ECC sondes but no statement is given about the reliability of the MOZAIC instrument. Statements like 'close agreeement' are not defined. Presumably the values referred to are better than the 20% discrepancy quoted for Brewer-Mast instruments. This is but one example of vagueness that does not make for a convincing conclusion. The less than clear Abstract reflects deficient organization in the paper. A series of scientific questions is not addressed but rather a recitation of statistics that are not prioritized is given.

Second, analysis of the vertical sonde-MOZAIC results does not place the results in context of previous work. This detracts from the value of the study because no interpretation is supplied. For example, ozone comparisons of the sonde measurement with coincident ground-based and satellite measurements constitute an important part of establishing quality and reliability of the sonde data at any given station (cf Comment below on P 7102, Lines 10-18). There are no references to such studies for the stations used although literature exists on this topic, in a string of Ozone Assessments (the latest one dated 2010) and references therein, if not in more recent literature. Few or no details are given about some of the technical issues to which the authors refer, eg impact of sonde sensing solution in the ECC sonde measurement, high background currents, instrument type. In lab studies and in evaluations of sonde data at some of the stations included in the present study, there are definite biases in the sonde measurement. Examples of station bias applicable to the tropical sites used here, for example, are given in Thompson et al., 2007; 2012. However, none of the quality assurance and technical results, that are referenced in Smit et al in the papers (2007, 2011) or in Deshler et al, JGR, 2008) are applied to interpretation of the results shown. An obvious question for the reader is: why are some stations closer to MOZAIC results than others? Are these offsets what one would expect based on biases displayed by the sonde technique used at the station? That is at the heart of the data quality assurance goal of the paper.

Third and finally, the conclusion about "trends" is not supported by the evidence presented in the paper. Yes, interesting and valuable comparisons are given and the research methods are good, but without interpretation other than the figures given, there is not enough information to conclude that the sondes are 'ready for trends analysis,' as implied by the last sentence in the Abstract. Indeed, as mentioned below, data consistency at some stations appear to be better than at others where speculative discussion is given (useful) but no conclusions can be drawn. That message in itself is a more suitable, Abstract-worthy conclusion than what is written in the present Abstract.

Other Papers referred to above:

A. M. Thompson, J. C. Witte, H. G. J. Smit, S. J. Oltmans, B. J. Johnson, V. W. J. H. Kirchhoff, F. J. Schmidlin, Southern Hemisphere Additional Ozonesondes (SHADOZ) 1998-2004 tropical ozone climatology.

3. Instrumentation, Station Variability, Evaluation with Simulated Flight Profiles, *J. Geophys. Res.*, 112, D03304, doi: 10.1029/ 2005JD007042, 2007.

A. M. Thompson, S. K. Miller, S. Tilmes, D. W. Kollonige, J. C. Witte, S. J. Oltmans, B. J. Johnson, M. Fujiwara, F. J. Schmidlin, G. J. R. Coetzee, N. Komala, M. Maata, M. bt Mohamad, J. Nguyo, C. Mutai, S-Y. Ogino, F. Raimundo Da Silva, N. M. Paes Leme, F. Posny, R. Scheele, Henry B. Selkirk, M. Shiotani, R. Stübi, G. Levrat, B. Calpini, V. Thouret, H. Tsuruta, J. Valverde Canossa, H. Vömel, S. Yonemura, J. Andrés Diaz, N, T. Tan Thanh, H. T. Thuy Ha, Southern Hemisphere Additional Ozonesondes (SHADOZ) ozone climatology (2005-2009): Tropospheric and tropical tropopause layer (TTL) profiles with comparisons to OMI-based ozone products. *J. Geophys. Res*., **117,** D23301, doi: 10.1029/2010JD016911, 2012.

SPECIFIC COMMENTS:

Page 7102. Line 6, after Smit et al, 2007; should add the first reference above, Thompson et al., 2007

Page 7102. Lines 10-18. The authors refer to some of the types of studies mentioned above that are needed to give perspective, ie 'space-borne, ground-based or other airborne...' but with no references or discussion of what those reports imply for the data used in the present study. This should be amplified and clarified.

Page 7102. The statement on line 17 that quality is most important in the UTLS is not unreasonable. However, it is somewhat misleading. The biases among techniques shown in the various JOSIE studies appear throughout the sonde profiles, including the troposphere (Smit et al and Deshler et al, references). Thus, biases *do* matter in the interpretation of the MOZAIC comparisons, that are mostly below 12 km.

Page 7102. Lines 25 following. So what changed in 1998? Sonde instrument or something with the MOZAIC instruments or protocol? The reader cannot discern what changes have taken place over time and why. Pointing out that these uncertainties exist is a strength of the paper. However, the interpretation and conclusions are less definite as a result. FOR EXAMPLE - later in the paper, (page 7122, lines 15-25), the authors argue that problems with the sondes are indeed not likely to be the sole reason for discrepancies with the MOZAIC. *This is an important result of the paper and should be included in the Abstract. Is not the "conclusion" for SI2N readers that <u>both</u> sondes <u>and</u> MOZAIC need further examination before either one is 'robust' and reliable enough for trends?* Reviewer recommends highlighting this finding in Abstract and Conclusion.

Page 7103. The second paragraph needs some correcting or clarification. Recommend deleting the sentence about 'led to confusion...'. Better wording. 'Led some groups to change their technique. (Remove word 'Additionally'). NOAA sites... [this is ok, but NOAA also changed instrument and technique at Pacific tropical stations it operates:

Samoa, Hilo, Fiji, San Cristobal]. The latter stations, because they are not near MOZAIC routes, are not included in the present paper.

Lines 25-26. Recommend stating ... "After 2004 mainly ES sondes were launched but the 1.0% solution strength was retained (see meta-data at WOUDC). Recently the ozonesonde data user community has been addressing how to account for changes in radiosonde instrumentation that have accompanied ozonesonde changes at a number of stations in the past 5 years or so [Stauffer et al., 2013*]. The radiosonde changes propagate to each ozone measurement, but mostly above 100 hPa; newer radiosondes mostly affect ozone data after 2009. Thus, these influences are neglected here."

*R. M. Stauffer, G. A. Morris, A. M. Thompson, E. Joseph, G. J. R. Coetzee, Propagation of radiosonde pressure sensor errors to ozonesonde measurements, *Atmos. Meas. Tech. Disc.,* **6**, 1-40, 2013.

Page 7104. The first paragraph is about CF then ends with reference to background current and SOPs. ?? Is it out of place or unfinished? The idea is picked up again in P 7121.

Lines 10-19. Information on background currents is given. Are background currents available to the authors? If so, how were they used? This paragraph does not relate to the rest of analysis in the way it is written. In a Table there is reference to background current 'increasing' or 'decreasing'. What does that mean for the analyses?

Page 7104. Like NILU, NASA Goddard has also collected campaign data that is archived in WOUDC. A number of the North American stations from which statistics are taken in this paper have additional data due to the Intensive Ozonesonde Network Study (IONS) experiments (eg http://croc.gsfc.nasa.gov/seacions). It would be appropriate to acknowledge it here by adding to line 25: "The Intensive Ozonesonde Network Study (IONS) experiments (Tarasick et al., 2010; Thompson et al., 2011) over North America have operated in 2004, 2006, 2008 and 2013. This has agumented regular launches at Boulder, Hunstville and Wallops as well as most of the Canadian stations listed in Table 1. The data are archived at NASA/Langley and WOUDC."

D. W. Tarasick, J. J. Jin, V. E. Fioletov, G. Liu, A. M. Thompson, S. J. Oltmans, J. Liu, C. E. Sioris, X. Liu, O. R. Cooper, T. Dann, V. Thouret, An ozone climatology for INTEX and ARCTAS from IONS ozonesondes, *J. Geophys. Res.,***115**, D20301, doi: 10.1029/2009JD012918, 2010.

A. M. Thompson, S. J. Oltmans, D. W. Tarasick, P. Von der Gathen, H. G. J. Smit, J. C. Witte, Strategic ozone sounding networks: Review of design and accomplishments, *Atmos. Environ.,* doi:10.1016/j. atmosenv.2010.05.002, **45,** 2145-2163, 2011.

Page 7105. (Line 1 ...) Thanks for pointing out somewhat confusing state of archives.

Pages 7105-7106. MOZAIC observations. Here is where details on quality assurance,

calibration over time, etc, are lacking. More references and facts are required. How does the reader know whether or not this 'standard reference' is constant over time? Actually the authors seem to doubt that it is. (See comments above).

Pages 7106-7108 on methods, etc. Good.

Pages 7108-7114. These are the best part of the paper because there are a lot of statistics; Further, there is a lot of information on the history of the European data and instrument variations from which the authors can convincingly make meaningful deductions.

Page 7114 (first para). The literature concerning the ES vs SP differences has been based heavily on JOSIE and other tests since 2000. Maybe these instrument issues would not have affected changes in 1997.

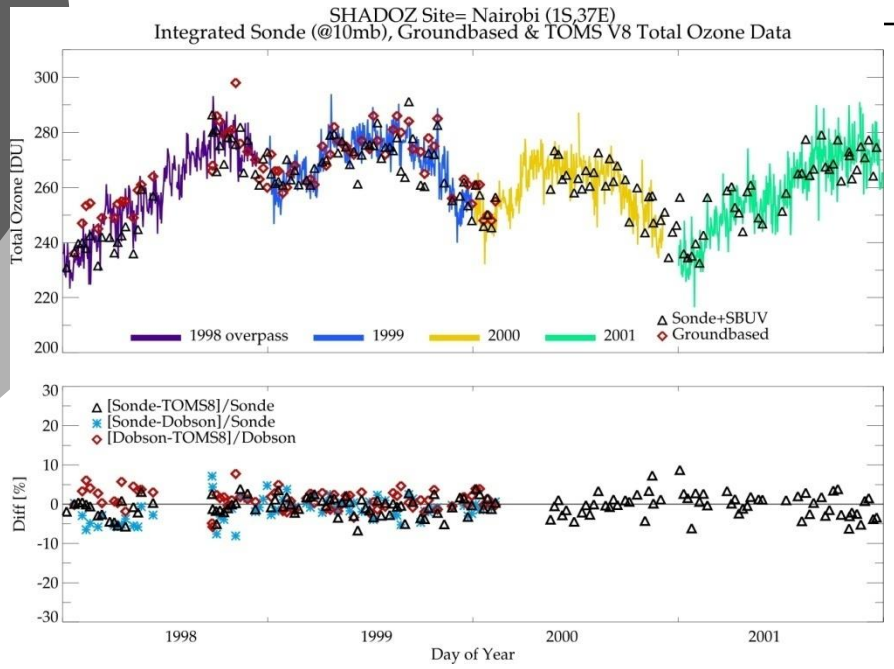
Pages 7114-7119. Compared to the discussions of mid-latitude European sondes, these sections, although important to report, suffer from (1) fewer statistics; (2) long-trajectories that imply considerable uncertainty in matches; (3) instrument changes that are not easy to interpret. Drawing conclusions from these sections can be misleading. Recommend in the **Summary & Conclusions** that a distinction is made between stations with a lot of history and literature (mostly in Europe) and those where the analysis methods, in addition to instrument issues, only leads to uncertain conclusions.

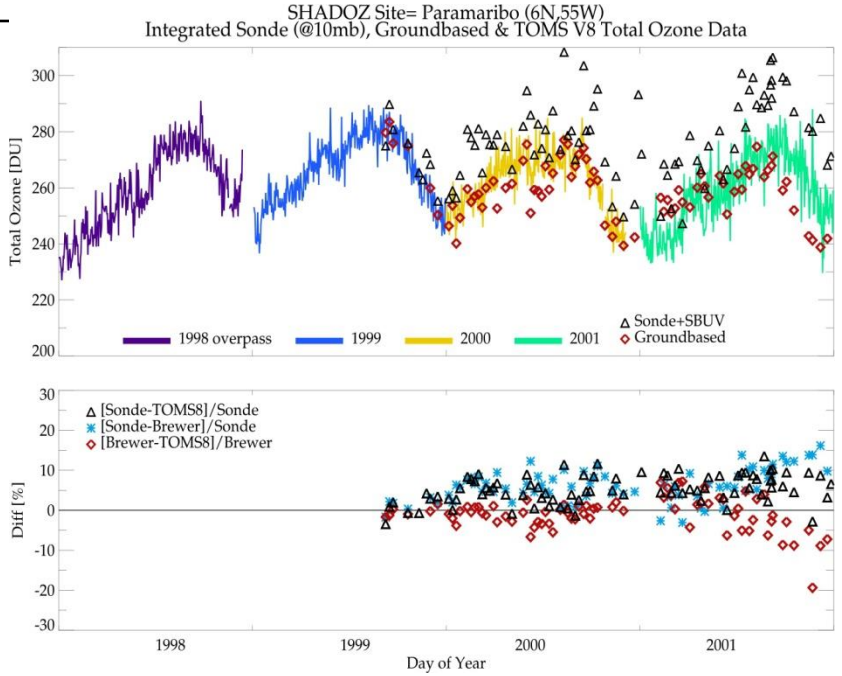
Examples... (A) very well stated on page 7115, lines 5-10. What to make of the Boulder comparisons?

(B) page 7116, lines 10-15. The discussion of Japanese data really underscores limitations of technique, with > 50 hr trajectories!

[C] page 7119 - about the tropical stations, ie Paramaribo. In SHADOZ analysis (eg Thompson et al., 2007) Paramaribo alone of a dozen or so sites, did not agree well with TOMS overpass nor with the co-located Brewer. A positive bias compared to all other tropical stations in the stratospheric segment of the profile was striking (see Figures at below taken from SHADOZ soundings and submitted for publication in 2011.) In 2012, as presented at an ozonesonde workshop (part of the O3DQA WMO/SI2N activity), a re-processing of Paramaribo was carried out by KNMI and total ozone offsets with satellite and spectrometer were nearly eliminated (also attached, as published). Evidently the data used in Staufer et al (downloaded in 2010) reflect high-Paramaribo ozone archives. The authors should inquire of KNMI about these data, when and how re-processing was carried out. To the Reviewer's knowledge, details of the data change have not been published, making it difficult to draw conclusions about this very valuable dataset.

Sonde-TOMS Comparisons (to 2001)



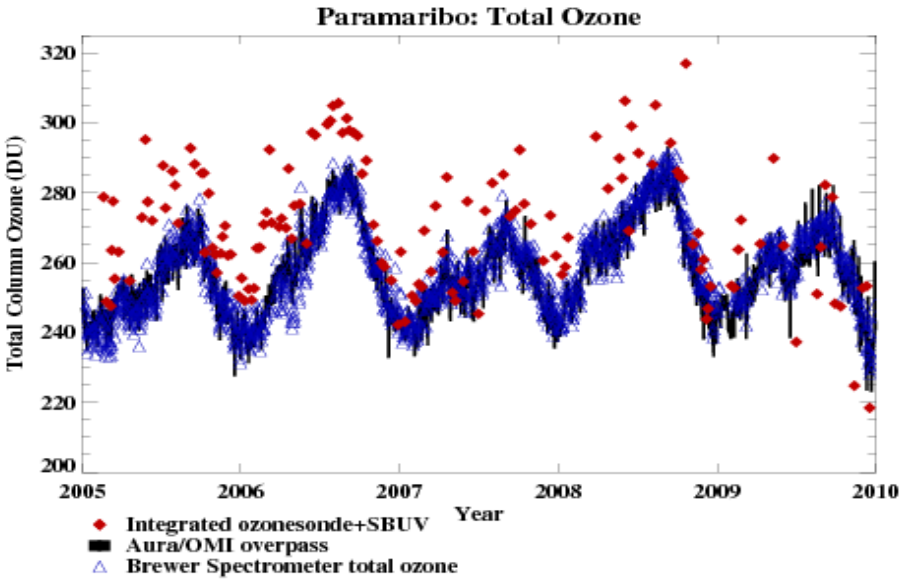


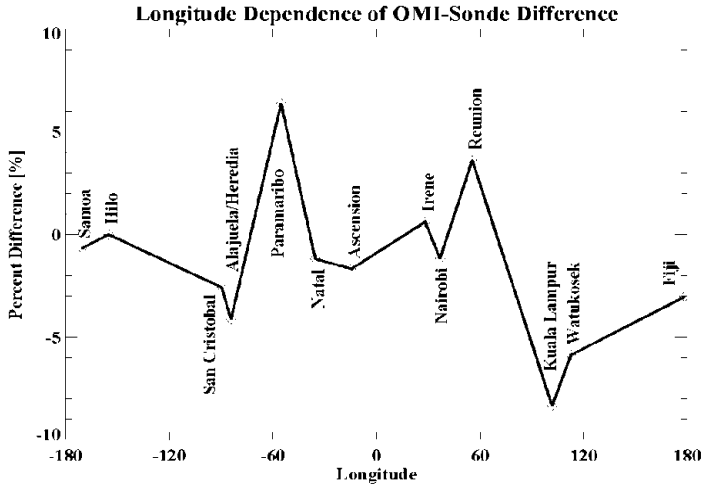
Nairobi - Excellent TOMS-total ozonesonde-Dobson agreement

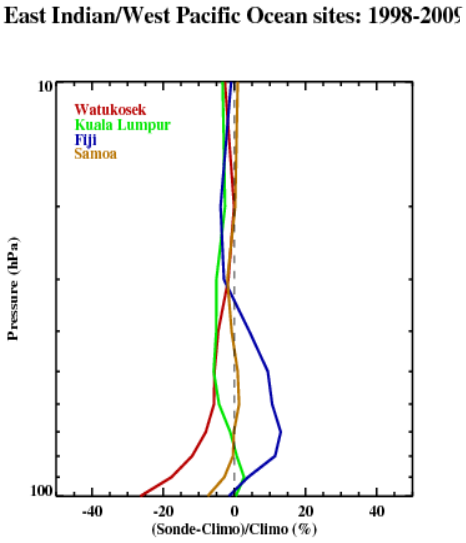
Paramaribo higher sonde than TOMS, drift ?

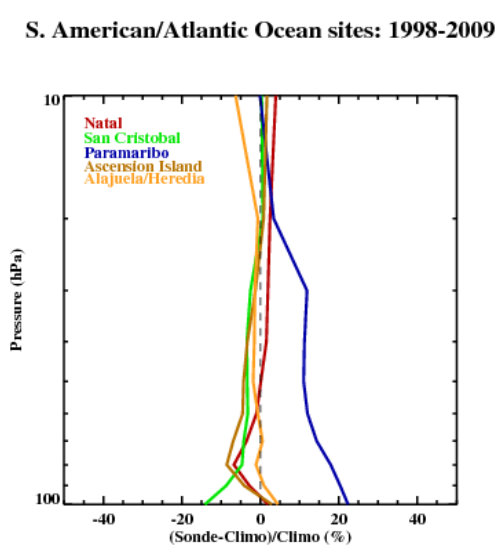
(Thompson et al., 2003)

Fall 2011- Thompson et al, JGR – submitted. Very high bias, lower right









Re-Processing of Paramaribo Data Improves Total O3, some of Profile. As published, JGR, 2012