Response to comments of Referee #2

We wish to express our appreciation to Referee #2 for the instructive comments as well as the time and effort spent. We have carefully considered all comments that were expressed in the review. In the response below, we address each of these comments. Your comments are italicized, and our responses immediately follow.

The manuscript by Xu et al. is good and the work it describes is very important. In particular because it drives the message that retrievals of ground irradiance using satellite data (in this case from TOMS) have large biases with respect to the in situ measurements by the ground stations. The authors use extensive USDA network of UV-MFRSR's and VIS-MFRSR's. The latter could and should be emphasized more in the paper. They concentrate on UV data to compare results with the retrievals from UV channels of TOMS. Clearly the surface albedo, cloudiness and aerosols pose a great problem for retrievals from satellite data. One could calculate that the amount of signal in UV that carry information about the lower parts of troposphere in TOA radiance is very small, so the sensitivity to the state of lower parts of troposphere is very weak. The presented results do not give us much hope that the situation could be significantly improved. The bottom line conclusion should be that the ground based stations are indispensable in monitoring the state of atmosphere and never will be replaced by space based remote sensing.

We thank the reviewer for appreciation of our work and constructive comments. We totally agree with your conclusion about ground based stations and will add a concluding remark in summary:

"Nonetheless, our comparison indicates that ground-based in situ measurements, like those from the UVMRP network, are indispensible in monitoring the atmospheric states and never replaceable by space-based remote sensing retrievals."

The authors should address the most important epistemological aspect that appears in works like theirs. The comparison between measured and retrieved (constructed) irradiances should include "the placebo test" results. Authors showed that measured irradiances are, say by X% smaller than the retrieved irradiances. The question is whether the given value of X is large or small or rather "good or bad". How "good" is irradiance retrieval from TOMS? We do not know answer after reading this paper,; we only know the value of X. Let suppose we had no TOMS data but instead performed irradiance synthesis from climatologic data only. This would be the placebo test. Suppose we obtain results that differed from ground based measurement by Y%. Comparing X% and Y% values can tell us how good is TOMS data set for what it is being used. If XY one would not need TOMS and that would imply that TOMS is on the level of placebo effect. Could authors show how much (if) X is smaller than Y?

We thank the reviewer for this good suggestion. Most UVMRP sites have UV measurements after 1997. Given our study period of 2000-2004, the available data records are insufficient to construct a meaningful climatology on a daily basis. Since the annual cycle (depending on calendar day of the year) dominates UV variability, correlation coefficients between daily TOMS retrievals and MFRSR measurements are extraordinary high (mostly greater than 0.85). Thus the TOMS retrievals can explain over 70% of the MFRSR daily variance. A question is whether the TOMS retrievals contain any signals other than the annual cycle. To answer this, we construct an annual cycle proxy [needed because of insufficient observational data] as follows:

$$\overline{M}_d = I_d \cdot \frac{1}{N} \sum_{y=2000}^{y=2004} \sum_{d=5/1}^{d=9/30} \frac{M_{d,y}}{I_d},$$

where the summation is done over all available data samples for the study years (*y*) 2004-2004 and julian days (*d*) between May 1 and September 30; *N* is the total number of samples; I_d is the solar intensity depending on the earth orbit (of the middle year 2002) and solar spectra; and *M* denotes for MFRSR measurements. We can obtain the same annual cycle proxy for TOMS retrievals $\overline{T}_d$ from its daily values $T_{d,y}$. Given the TOMS systematic overestimation, a constant reduction is applied to correct the first-order biases in all *T* values at each site. A placebo test can be made to check whether $T_{d,y}$ is better than $\overline{M}_d$ in predicting $M_{d,y}$. One measure for the test is to compare the root-mean-square error (rmse) between $(T_{d,y} - M_{d,y})$ and $(\overline{M}_d - M_{d,y})$, respectively referred to as *ReT* and *RdM*. The TOMS retrievals may contain additional signals (to the annual cycle) if *ReT* is smaller than *RdM*.

Figure Response.1 compares the relative *ReT* and *RdM* statistics, both normalized by the single average of all $M_{d,y}$ data samples, for each of the 4 wavelengths (305, 311, 324 and 368 nm) at 27 sites under both clear and total sky conditions. The *ReT* is close to or slightly smaller than *RdM* for all wavelengths and sites under clear conditions. On average over 27 sites, the *ReT* minus *RdM* differences for 305, 311, 325, and 368 nm are -0.06, -0.02, 0, and 0, respectively. This indicates that the TOMS retrievals do not have clear advantage in depicting daily fluctuations than that contained in the MFRSR annual cycle. On the other hand, the *ReT* under total-sky conditions is systematically and notably smaller than the *RdM* for all wavelengths at most sites. The averaged *ReT* minus *RdM* differences for the 4 wavelengths are -0.15, -0.10, -0.10, and -0.11 respectively. Thus the TOMS retrievals have strong ability to resolve daily fluctuations in addition to the annual cycle under total-sky conditions, due likely to the overwhelming cloud effects that are detectable by satellites.

Figure Response.2 illustrates the temporal correlations between daily fluctuations of TOMS retrievals $(T_{d,y} - \overline{T}_d)$ and MFRSR measurements $(M_{d,y} - \overline{M}_d)$ for each of the 4 wavelengths (305, 311, 324 and 368 nm) at 27 sites under both clear and total sky conditions. The correlations are systematically much higher at 305 nm than other wavelengths under clear conditions. The reason for this spectral dependence is not known. Nor it is clear why the correlations for 325 and 368 nm at CA02 and TX02 drop to near zero or negative. The averaged correlations over all sites excluding CA02 and TX02 are 0.68, 0.47, 0.33 and 0.34 for the 4 wavelengths. The result implies that, under clear conditions, the TOMS retrievals can capture 46, 22, 11, and 12% of the daily MFRSR variance in addition to the annual cycle, albeit the *ReT* minus *RdM* differences are relatively small. Under total-sky conditions, all correlations are substantially high and exhibit little spectral dependence. One exception is site CO12 at a mountain top (3220 m), where low correlations may result from the TOMS and MFRSR differences in viewing clouds. Excluding site CO12, the averaged correlations for the 4 wavelengths are 0.82, 0.78, 0.77, and 0.79. This indicates that, under total-sky conditions, the TOMS retrievals explain 68, 61, 60, and 62% of the daily MFRSR variance in addition to the annual cycle.

In summary, the TOMS retrievals under total-sky conditions, after correcting for the first-order systematic biases, are much more skillful than the MFRSR mean annual cycle, explaining over 60% of the variance for the additional daily fluctuations. The advantage is largely reduced under clear-sky conditions, but the TOMS retrievals can still capture 11-46% of the variance. Since the ground-based sites are extremely limited over land and none over oceans, the space-based remote sensing retrievals are the only feasible solution to monitor surface UV radiation over the globe. In this regard, the TOMS retrievals are valuable and contain a significant amount of signals for surface UV irradiance variations in space and time.

Xu et al. should expand the cited literature to put more emphasis on the importance of their work in contrast to what was done before. Instead the authors use a lot of space to deal with technical aspects of measurement and retrievals and pepper their manuscript with way too many citation concerning these details. In fact the format of this journal - that requires placing the names and years in brackets - is really very wasteful and it makes reading difficult, disrupting the flow of the major narrative. It would be much better if citations were in the form of less intrusive footnotes. But this is not authors fault; they merely adhere to the journal format requirements.

We thank the reviewer for the suggestion. We will shorten Section 2 by deleting the technical details in lines 111-130 and 145-150 (about a page) of the original manuscript. We also will eliminate Figure 1 and relevant text on the MFRSR calibration from page 10973 line 26 to page 10974 line 7. We also agree with you on the citation format, although we cannot overwrite the Journal's requirement.

The captions under figures should be more elaborate. For example they should state at which wavelength the AOD is shown. In fact, the general rule is that figures should be self sufficient and to make it unnecessary to search the paper for the description of the figure.

We agree and will elaborate the captions for Figs. 1, 2, 3, 5, 6.as follows:

Figure 1. The relative differences ($1-I_{aero}/I_{lamp}$) of 3-minute clear-sky spectral irradiances for 325 nm between the UVMRP lamp calibration (I_{lamp}) and AERONET transferring calibration (I_{aero}) at the collocated UVMRP research site in Greenbelt, Maryland.

Figure 2. Scaling factors for the re-convolution of the TOMS spectral irradiances at (**a**) 305 nm, (**b**) 310 nm, (**c**) 324 nm, and (**d**) 380 nm to the UVMRP measurements at in Bondville, IL (IL02) in July, 2003 in terms of the total column ozone for the range of 200-500 DU and three solar zenith angles (0,

30, 60 degrees). The scaling factors were normalized by the corresponding values of 300 DU ozone.

Figure 3. The relative mean (**a**) and standard deviation (**b**) of biases (*left axis*)of the 2000-2004 summer (May-September) noontime surface UV irradiances at 305, 311, 325, and 368 nm under clear-sky conditions using the TRT (TOMS reflectivity threshold) cloud screening method during 2000-2004. Also shown as the top curves (*right axis*) are total optical depth (aerosol plus cloud) at 368 nm (**a**) and total number of clear-sky days (**b**).

Figure 5. The temporal correlations of the TOMS irradiance biases from the UVMRP measurements with in-situ optical depths at 368 nm observed by UVMRP at each monitoring site for 305, 311, 325, and 368 nm.

Figure 6. Same as Fig. 2 except using the L&A (Long and Ackerman [2000]) cloud screening method.

The results are presented chiefly in tables for 27 UVMRP stations. While the bar plots give a full picture of the results of comparisons, they are not easy to interpret and draw any conclusions. It remains a mystery to this reviewer what is the reason for the particular order the stations are listed. Perhaps station could be ordered according to geographic latitude or averages AOD value.

We thank the reviewer for the suggestion. The 27 UVMRP sites are first put in order on the basis of the 6 climatic regions they are located (see section 5.4 regional dependence). These regions, now named in the table as the Northwest-Rocky Mountain, Southwest, Central Plains, Midwest, Northeast, and Southeast, are arranged from north to south and then from west to east. Following your suggestion, we will rearrange the order of the sites within each climatic region according their AOD values from smallest to largest. This will be done in Table 1 and Figures 3, 5, 6, 7. As such, we will add the following text in page 10973, line 11 after "Table 1 lists their site specifications, including brief descriptions."

"These sites are listed in the order of 6 climatic regions (Northwest-Rocky Mountain, Southwest, Central Plains, Midwest, Northeast, and Southeast) arranged from north to south and then from west to east. Within each region, the sites are listed in the order of increasing AOD values."

Perhaps, an addition of plots showing mean and standard deviation as functions of AOD, cloudiness, albedo, latitude, longitude would be a better form of presentation unless there is no meaningful correlation. But authors show Fig. 4 where correlation is rather low with respect to SO2 abundance.

We agree and will add a figure to show standard deviations at 368 nm in correspondence with AOD values, which are strongly correlated with a coefficient of 0.75. The correlations with latitude and longitude are very weak, while credible cloudiness and albedo data are lack.

I suggest shortening the paper, reducing amount of technical citations concerning how instruments were calibrated, etc and try to emphasize the differences between TOMS and UV-MFRSR's and do a budget of errors that could attribute the errors to instruments, methods and factors related to the state of atmosphere as well as help to explain the reasons for the discrepancies.

We agree and will shorten the paper. See the responses above.

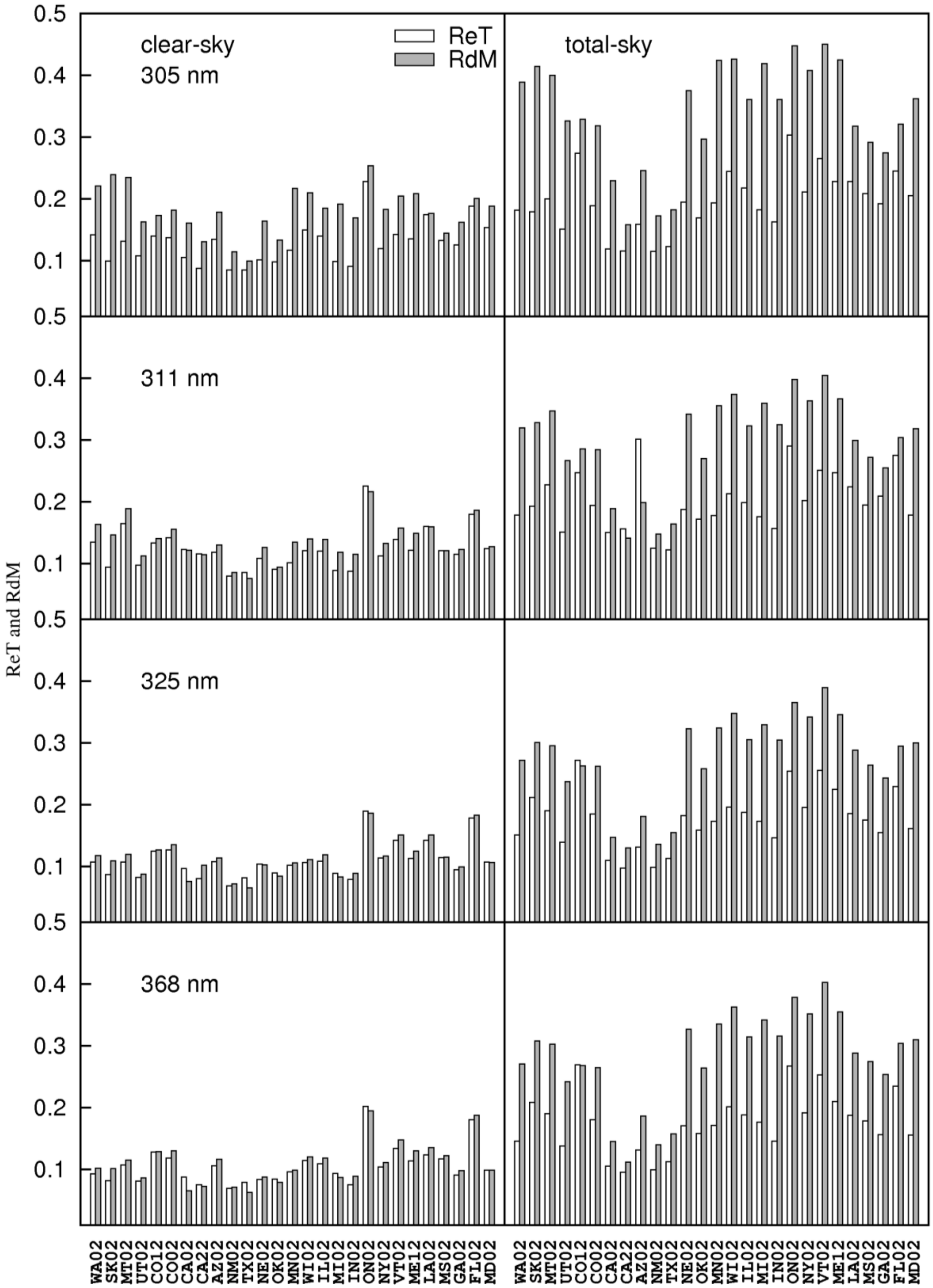

Figure Response 1. The relative root mean square errors of TOMS (white bar) and standard deviations of UVMRP (grey bar) noontime surface UV irradiances at 305, 311, 325, and 368 nm during 2000-2004 summer (May-September) under clear-sky conditions (left panel) and total-sky conditions (right-panel).

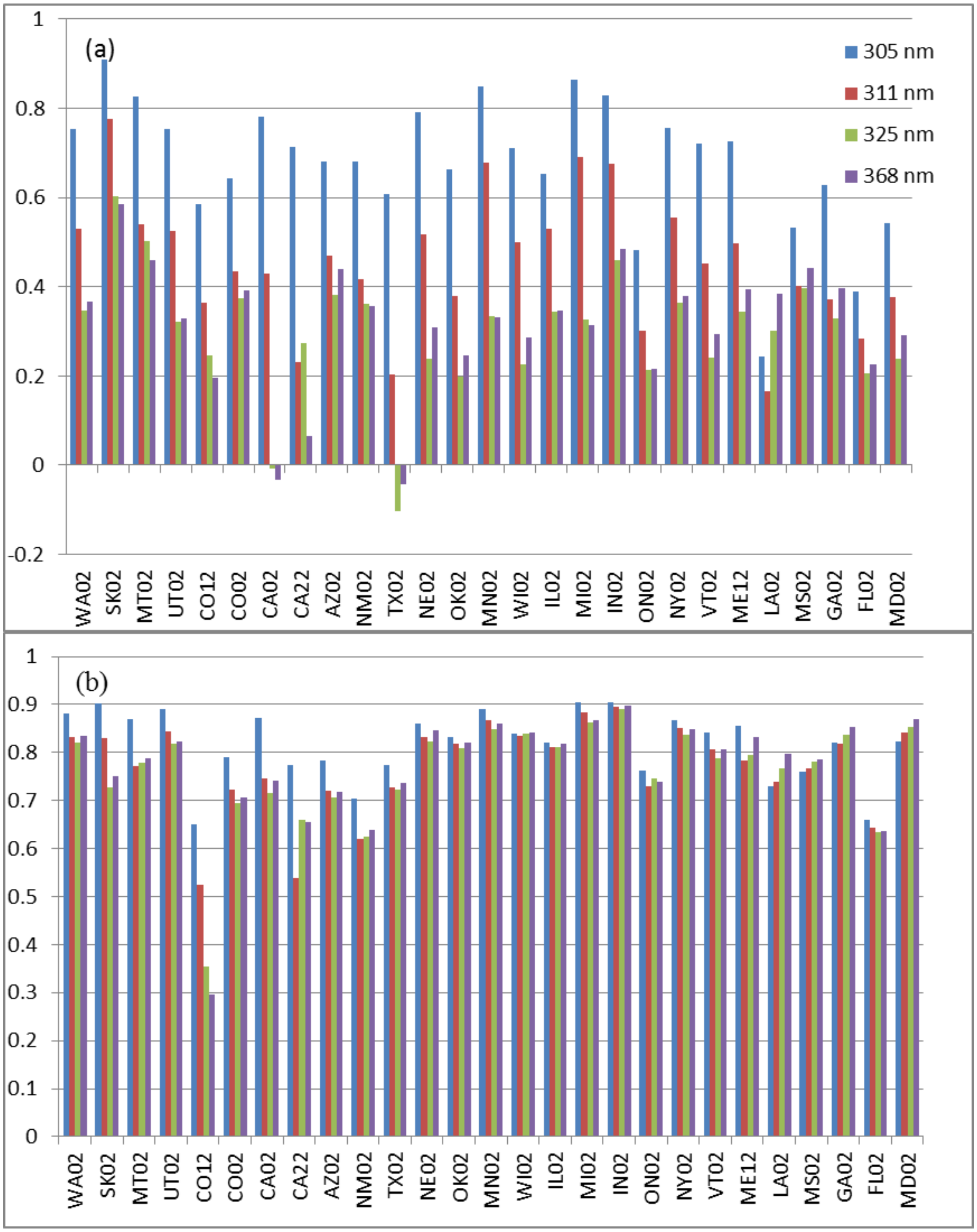

Figure Response 2. The temporal correlations of the perturbations of TOMS at 305, 311, 325, and 368 nm during 2000-2004 summer (May-September) under clear-sky conditions (a) and total-sky conditions (b).