Response to the reviewer 2#.

Three different tests have been carried out to analyze the effect of the variation of the calibration coefficients with time in the Microtops performance: a) using the original factory calibration coefficients throughout the entire time period; b) changing the calibrations coefficients in a stepwise way; and c) using a linear interpolation of the calibration coefficients as it was explained in Sec. 3.2.

The test has been studied through the relative deviation against the Brewer measurements. In this section, only comments for airmasses lower than 3 are considered because the RDEV behaviour for large airmasses will be analyzed in Section 4.3.

a) The use of the original calibration over the entire period yields different effects depending on the Microtops Channel (Figure 1 a, b and c).

The deviation for the three channels shows a daily cycle, (i.e. the RDEV is larger for low airmasses), and increases with time since the calibration.

Channel I shows a slight underestimation of the Brewer ozone. A limited RDEV variation which increases for low airmasses ($m < 1.2$) is observed. The RDEV does not exceed 4% for the measurements done until 2004, but it reaches 8% for the Lampedusa campaigns, in 2008-2011. Excessive RDEV and notable daily cycle are observed ten years after the calibrations.

Channels II and III are more sensitive to the calibration deterioration. Excessive differences and large daily cycle appear already in 2001, earlier than those obtained for Channel I. Channel II and Channel III overestimate and underestimate respectively the Brewer ozone for the whole airmass range. The largest RDEV is observed at airmass 1, with values reaching -20% and 18% for Channel I and Channel II respectively, for measurements performed until 2004. Successive measurements show larger RDEV, up to -50% and 40% for Channel I and Channel II, respectively.

b) The step by step calibration has been applied to the whole data set. Therefore, the original calibration is applied to the 2001 campaign; the 2002 calibration is used from the period 2002-2009; and the 2010 calibration is used for the 2011 measurements (Figure 2 a, b, and c).

The Microtops measurements show RDEV < ±2, 4 and 5% for Channel I, II and III respectively during the two years closest to the calibration. The deviation increases and

the daily cycle progressively appears with time since the calibration, as it was observed also when the original calibration was applied (test a). These effects are observed in 2008 and 2009, six years after the 2002 calibration. This time interval is shorter than that observed in test a. Apart from the calibration, that is probably due also to the degradation of the filters caused by the use and transport of the instrument.
Also in this case the channels II and III are more sensitive to the time passed after the last calibration.

c) The linear interpolation of the calibration coefficients between consecutive calibrations is the best option to monitor the Microtops performance over long time periods, even if they are far from each other in time. With this approach the performance of Microtops measurements remains stable with respect to the Brewers, and the daily cycle observed at low airmasses is reduced. For this reason, the data set of the entire period 2001-2011 was analyzed using this approach (Figure 3 a, b and c).

The different sensitivity to the calibration observed for the three Microtops channels induces differences between Channel I and Channel II retrievals, which are larger at low airmasses. The use of the original calibration induces differences larger than 30 DU for 2001 (4 year after the calibration) increasing with time (Figure 7a). If the stepwise calibration is used, these differences are less than 10 DU during the two year after the calibration slightly increasing up to 20 DU for airmasses lower than 1.2 (Figure 7b). The linear interpolation of the calibration coefficients (Figure 7c) reduces the differences between channels, with a maximum value of 10 DU, and inhibits its diurnal cycle.