We thank the Reviewer for her/his comments.

Minor comments:
It should be noted clearer somewhere, how aerosol and temperature gradients are related and, particularly, what limits this correlation in more complex cases.

We will add additional text as detailed below.

The lidar analysis is based on gradient of aerosol scattering cross section. When aerosols are uplifted after sunrise by convective mixing, they act as "good" tracers of the depth over which mixing occurs. During the day, the level at which air parcels become negatively buoyant corresponds to the first temperature inversion. It often happens that the convection erodes the temperature inversion allowing the buoyant air parcel to be lifted further up. When turbulence weakens in the afternoon, due to decreasing sensible heat flux, the temperature inversion builds up again and the depth over which mixing occurs becomes shallower, but the aerosols can remain aloft, without evident subsidence. In these afternoons and after-sunset conditions, the strongest aerosol gradient corresponds to a 'residual' aerosol layer aloft the actual SML. The residual layer top height corresponds then to a higher temperature inversion. Departure from these dynamics can occur when the residual layer does not entirely disappear during daytime and the lidar-detected layer remains above the actual main temperature inversion. Another case of no matching between aerosol and temperature gradients may occur when a temperature inversion forms in the already developed boundary layer so that aerosols are homogeneously distributed below and above the inversion and no gradient is observed in lidar data.

How well does your algorithm perform in cases where e.g. Saharan dust (or other particle layers) is mixed (entrained) into the BL?

The algorithm depends mainly on the power and SNR of the instrument; also cases where large particles can cause high extinctions limit the algorithm functionality. In cases where the particles do not fully attenuate the lidar signal and are mixed in the BL layer, the algorithm can provide reliably BL detections. The algorithm is configurable using a set of parameters, e.g. height and time constraints give the algorithm the ability to improve its detections where complicated structures of aerosol layers may cause erratic or invalid data from the backscatter profiles.
As an example: on the 15th of April till the end of May 2010 eruptions of Eyjafjallajokull generated several plume events advected over Mace Head. In correspondence to the entrained plume into the Boundary Layer the THT algorithm was applied to the backscatter profiles of the CHM15K with clear detections of the BL height with diluted volcanic aerosols within it (O'Dowd et al, submitted to *Atm Env* "Eyjafjallajokull special issue").

How robust would the algorithm provide BLH to be assimilated into models?
To which extent would it be operational?
What would be the approximate hit/miss ratio and typical data coverage for let's say 6-h intervals – i.e. how often would you have a valid value within each 6 h?

Currently the algorithm is not running in a real-time scenario, if the algorithm were to run 24 / 7 a reliable configuration would be to generate profiles averaged over 5 minutes. The average over 5 minutes will improve the signal to noise ratio. On a standard cloud-free 6-hour period the BLH detections would have a miss rate of 5% . This strongly depends on the settings of thresholds and parameters of the algorithm.

p 578, l 25: typo : : : influences the aerosols and turbulent mixing: : :.

We will make corrections

p 579, l 25: these limitations generally DO exist when different kinds of observations serve for retrieval of dynamically complex quantities (same for tropopause, ozonopause,: : :)

We will change to:
…these limitations exist when applying general algorithms based on the aerosol backscatter vertical gradient.

p 582, l.5: why should the detection of the DRCL with CHM15k perform worse than with CL31 although the aerosol signal is more consistent with the ALS 300, which has the strongest signal and serves as a kind o reference here? Could this be just an artefact of the small sample that is compared for the CL31 vs RS?

Yes the number of CL31 detections of the DRCL was very limited, and was the smallest dataset available from the three instruments. We report here from line 4 to 10 page 582: Despite the lower SNR and *R*-value compared to the ASL300 and CHM15K, the CL31 is more consistent with the RS retrievals of the DRCL with 62.5% of detections closer than 200m to the RS's second temperature inversion. However Table 3 shows the number of samples is significantly lower, the CL31 having samples _60% of CHM15K and ALS300. This means that in 5 cases out 23 the CL31 is getting DRCL detections closer than 200m to the RS's.

Fig 5-7: enlarge the dot-symbols

We will make the change