Anonymous reviewer #1

We would like to thank the reviewer for the careful reading and constructive comments.

Please find below our point-by-point replies. For clarity, the reviewer's comments are displayed in black, our replies in blue, and suggestions for revised text in blue italics.

Main points

1.The comparison contrasts sets of CO2 simulations in lowest 20 m near surface made with emissions emitted either at surface or at more realistic heights. It should be mentioned that the observations are often made at higher elevations than 20 m, using either small towers (40-100 m) or tall towers (200-300 m tall). For modeling such observation sites, the conclusions presented in this study can serve more as a warning, rather than ready to use estimate of emission height-related bias.

This is a very valid point. For observations from tall towers the effect will likely be smaller, though not negligible. We added the following sentence at the end of Section 3.3

The impact on observations from tall tower networks measuring CO_2 some 100 m to 300 m above the surface (Bakwin et al., 1995; Andrews et al., 2014) will likely be somewhat smaller than suggested by the numbers above, especially in winter when the atmosphere is less well-mixed.

and the following sentence in the conclusions section:

Since measurements of CO_2 are often taken from towers some 100 m to 300 m above the surface (Bakwin et al., 1995), the impact on actual ground-based observations will likely be somewhat smaller.

2.Lagrangian plume models (eg STILT, FLEXPART) are often used in backward, adjoint mode for inverse modeling, and some are used in studies cited here (Page 3 Line 8). In that setting they have to assume emissions are mixed quickly in surface layer of nonzero thickness. It can be as thick as diurnally varying PBL height (Lin et al., 2003) or assigned a constant value (Ganshin et al., 2012). This is done to minimize sampling errors in estimating adjoint tracer concentration near surface, which is made by counting particles in the surface layer. In case of using relatively thick layer, the assumption may reverse the effect of neglecting CO2 emission height, towards having more errors from surface emissions rather than from elevated stacks.

Thank you for pointing this out. In fact, in FLEXPART the particle residence times can be written out for multiple vertical levels, which would offer the possibility to account for emissions at the surface and at higher altitudes separately. To our knowledge, this possibility has not yet been explored in the context of emissions from power plants and industrial sources, though. We added the following sentences:

In Lagrangian models such as STILT (Lin et al. 2003) or FLEXPART (Stohl et al. 2005), which are often used in backward, adjoint mode for inverse modelling, particles are typically sampled over a fixed vertical depth above the surface or relative to the height of planetary boundary layer to derive source-sensitivities. Similar to the release of emissions at the surface in Eulerian models, this ignores the potentially different sensitivities to emissions from elevated sources.

Technical corrections:

Page 2 Line 11 Add period after CO2 and before "Top-down".

There was already a period.

Page 3 Line 8 Add year to Lauvaux et al.

Year added.

Page 3 Line 15 It is worth noting earlier references to air quality modeling, such as SMOKE-CMAQ modeling system (eg Houyoux et al, 2002), to emphasize that the problem had long been recognized and addressed. For CO2 modelling audience it is also useful to mention that in air quality modeling effort is made to account for plume rise height of biomass burning emissions (eg Achtemeier et al, 2010).

Thank you. We have added these references and modified the paragraph to the following:

In the air quality modeling community, the importance of vertically distributing emissions has been recognized much earlier (e.g., Houyoux et al. 2002) and is now well established, especially for species such as SO_2 that are primarily emitted from power plants and industrial sources (Bieser et al., 2011; Mailler et al., 15 2013; Karamchandani et al., 2014; Guevara et al., 2014). Accounting for plume rise has also been demonstrated to be critical for biomass burning emissions (Achtemeier et al., 2011).

Page 4 line 6 Written as "COSMO is the first NWP model worldwide" - it appears that similar effort with ASUCA model (Shimokawabe et al., 2010) was done in about same time, suggest checking, rephrasing.

Right, efforts for GPU acceleration of NWP/climate models such as for ASUCA or for the CAM-SE model have indeed been acknowledged in the publication of Fuhrer et al. (2014). Nevertheless, we maintain that COSMO is the first NWP model run on GPUs in an operational weather forecasting context. We slightly changed the wording to " *COSMO is the first operational NWP model worldwide*".

Page 7 line 16 Suggest revising “In order to prevent re-heating,” as “In order to avoid re-heating,”

Done

Page 12 Line 10 Need to add year to Bagley et al.

Done