Atmos. Chem. Phys. Discuss., 9, C7819–C7821, 2009
www.atmos-chem-phys-discuss.net/9/C7819/2009/

Atmospheric Chemistry and Physics Discussions

Interactive comment on "Background ozone over Canada and the United States" *by* E. Chan and R. J. Vet

Anonymous Referee #2

Received and published: 3 December 2009

This is an extensive study of ozone in North America. The analysis could reveal useful information on the ozone patterns and distributions. However I have significant concerns about the presentation and about the key assumption on defining "background". First, there is so much material here that it is difficult for the authors (and reader) to really understand the results and implications. In many places key details of the analysis are glossed over or not clearly presented. Regarding the key assumption on background, the authors' method to define ozone distributions by trajectory clusters is informative, BUT the cleanest cluster is not the "background". For all of the above reasons, I would say that the paper needs extensive rewriting prior to acceptance.

Specific comments below: Pg 21114, line 4: this definition of background is the same as what EPA defines as the policy relevant background. Pg xxx15, line 8: This sentence

is not quite right. The "controllable" does not change so long local emissions don't change.

Lines 10-20: Background is not directly observable. All of these studies reported on measured O3 and inferred what was happening to background. Pg xxx18, line 8: 925 hPa. What about mtn sites?

I am confused how you connect the 4 trajectories per day on GMT time, with the one 6 hr average O3 on LT.

Line 22: "The cleanest clusters are assumed to represent background air".. This is not correct. Consider a simple case for a site with just two wind directions. One direction has very low "background" due to few natural or foreign emissions. The other direction has high background due to high natural and foreign sources. Your method defines background as simplest the lowest cluster and ignores the contribution from the other natural and foreign sources. The cleanest cluster does not give the background distribution and it doesn't matter whether you use the 50th, 90th or 95th percentile.

Pg xxx19, line 10-15: I don't understand. Are you doing the PCA analysis for each season separately? Does this mean the geographical groupings change by season? Pg 19-20: Again, don't understand how you link the 4 GMT traj per day with the one 6-hour ozone value. Are you including all trajectories in clustering even though the ozone data is only 6 hrs per day? Pg xxx21: The organization is very confusing. You must start with the results of the PCA analysis as this is key for the rest of the analysis. If reader doesn't understand the PCA, they will not understand rest of paper. This organization is partly why it took me so long to get thru your paper. What are all the numbers here? What does it mean where you state "the mixing ratios ranged from 31 to 38 in spring". The mixing ratios of what? Very confusing. Pg xxx22: What regions are PC1, PC5, etc?. Table 1b is impossible to figure out. This is all very poorly presented. Line 24, What is the significance of this statement "The PCA regions were ordered by the percentage of the total variance. . ." Did you use this ordering? As I

ACPD

9, C7819–C7821, 2009

understand the PCA analysis, you grouped sites by their 6-hour daily ozone values. In other words sites that have similarities in the time series of the 6-hour values would get put into the same PC groups. This is a very interesting analysis and could be the subject of one paper on its own. In fact I would recommend this as there is too much in this paper and it is not well explained. That there are daily variations between sites that are well correlated on a regional basis is very interesting. That said, I am very surprised that there was not a single outlier. That is the PCAs didn't put a single site into a wrong category. Seems quite remarkable. Pg xxx23: Figure 2 is all trajectories or just the "background" cluster? I spent a lot of time reading and rereading sections to understand what was done. I ran out of steam at this point. The text needs substantial rewriting.

Table 1: I cannot figure out what is being presented. What are the two numbers in each column? Are all sites in each grouping averaged into one value or do the numbers in the table somehow give information on variations between sites. Why are there multiple PCs for the same region? Don't the PC regions change for each season? For the diurnal values are these the min/max O3 mixing ratios or the amount of daily variation (max minus min)? I assume the diurnal variations are based on hourly data, but everything else uses 6 hour averages?

Table 2: This table needs to give information on the regions.

Interactive comment on Atmos. Chem. Phys. Discuss., 9, 21111, 2009.