

## ***Interactive comment on “The first estimates of global nucleation mode aerosol concentrations based on satellite measurements” by M. Kulmala et al.***

### **Anonymous Referee #1**

Received and published: 23 August 2011

#### General comment:

It is nice to see that satellite products are used in the studies of the ultrafine particles, even though the results are not so good. The authors state that the derived proxies are able to predict the concentration of nucleation mode particles over the continents but the results do not support this conclusion. Reasoning for that will be presented in the Specific comments section of this review.

After some changes the manuscript could be suitable for publication in ACP. Not for the goodness of the results but in order to encourage developing more suitable satellite products for the analysis of fine particles. My suggestion is that the main conclusion

C8097

should be reformulated such that it is not possible to get adequate estimation results for nucleation mode particles with current satellite products.

#### Specific comments:

Page 18830, lines 5-9: How reliable is the estimating method of [ORG]? If the method by Paasonen et al. (2010) is used, some kind of estimate of the additional variation caused by the calculation of [ORG] should be given.

Page 18832, lines 20-24 and Page 18834, lines 20-29: Removing [SO<sub>2</sub>] and [ORG] from the proxies only because data is not available is disturbing. Several studies have shown that SO<sub>2</sub> is an important factor on the new particle formation and growth. In a clean environment like Hyytiälä it might be possible to set [SO<sub>2</sub>] constant but then [ORG] should be included in the proxy due to the emissions from the boreal forest environment. In more polluted environments removing [SO<sub>2</sub>] from the proxy would most probably lead to flawed results.

- What is the correlation between UV\*SO<sub>2</sub> and N<sub>nuc</sub>?

- Do you think that the results made with Hyytiälä data can be extrapolated into global environment? Especially over polluted areas?

Page 18834, lines 4-19: This section gives many explanations why the data is scattered, but would it be possible to eliminate some of the variation with a simple parameterization?

Figure 1: I would like to see this plot such as the axes would be scaled to the same range.

Figures 1-2 and referring text:

The coefficient of determination (R<sup>2</sup>) is a measure of how well the regression line represents the data. The analysis has been made with linear least squares fit and the model contains intercept term. Thus, by definition, R<sup>2</sup> can be calculated as square

C8098

of the Pearson correlation between the observations and the predictions given by the estimated function, which I assume to be  $r$ -value reported in the captions of Figures 1-2. In Fig 1.  $R^2$  values for the fits are less than 0.3, which indicates that the proxies are able to explain less than 30% of the total variation of  $N_{nuc}$ . In Fig 2 the situation is even worse; the explanation capability of the proxies is around 6% and the lack of explanation capability can also be seen straight from the scatter plots. These plots are text book examples of a situation where there is no correlation between the two variables. Drawing a line through a random sample and claiming that there exists significant correlation is bad statistics and in some cases even deceptive.

In the light of these results I would say that the statement in the end of the abstract: "The global pattern of nucleation mode particle number concentration predicted by satellite data using our proxies compares fairly well with both observations and global model simulations." is far too optimistic.

Page 18835, lines 6-9: I did not find the number of data points in Figure 2 but it seems that the  $N$  of observations is so large that the test for the significance of the correlation coefficient is meaningless. Since the test for significance is highly dependent of number of observations, even weak correlations may seem to be statistically significant.

Page 18835, lines 9-11: I agree with this sentence: "We may conclude that while our solution to replace CS with AOD is necessary in order to apply the proxies to a global scale using satellite data, it is clearly not the ideal one." The Section 3.1 should end on that.

Page 18836, lines 9-15 and Figure 4: Does the AOD have a seasonal cycle or is the seasonality shown in the plot caused only by UV?

Page 18837, lines 21-23: If the proxy fits poorly to Hyytiälä data, how does it justify the use of it to global environment? In addition, how is it possible to use only data measured in Hyytiälä to justify the use of the derived proxies outside of boreal forest region?

C8099

Page 18838, lines 9-19: This section is the main result of the study and the Concluding remarks section should be written accordingly.

Minor comments:

Add basic statistics from the measured parameters used in the proxy construction; e.g. mean, median variance,  $N$  of observations, averaging period

Page 18825, line 5, Raes et al., 2010 missing from the reference list

Page 18826, line 11, change Merkikanto to Merikanto

Page 18829, line 7, Sipilä et al., 2010 missing from the reference list

Page 18836, line 23, Suni et al., 2008 missing from the reference list

---

Interactive comment on Atmos. Chem. Phys. Discuss., 11, 18823, 2011.

C8100