

## ***Interactive comment on “Multi-annual changes of NO<sub>x</sub> emissions in megacity regions: nonlinear trend analysis of satellite measurement based estimates” by I. B. Konovalov et al.***

**Anonymous Referee #3**

Received and published: 18 June 2010

The aim of this paper is the derivation of NO<sub>x</sub> emission trends over European and Middle-East megacities using (1) the 1996-2008 record of summertime satellite observations of tropospheric NO<sub>2</sub> columns retrieved from the GOME and SCIAMACHY sounders, (2) the CHIMERE chemistry transport model run at a resolution of 1°×1° over Europe, and the assumption of a linear relationship between NO<sub>2</sub> columns and NO<sub>x</sub> emissions in the model, (3) a perceptron-type method for estimating the trends, deciding on their linearity or not, and estimating their statistical significance. This is a nice study, which constitutes an improvement over past work conducted by the same group, and a further application of neural network-based statistical methods in atmospheric physics. Nevertheless, despite the effort spent by the authors to describe their meth-

C4266

ods thoroughly and to explain the reasons that dictated several choices, the manuscript is often uneasy to follow, especially Section 3.2, where the estimation of a non-linear trend is described. The probabilistic neural network approach for estimating the trends being a major building block of this study, I think that an additional effort is needed to amend the readability of this section. The manuscript can be accepted for publication in the *Atmospheric Chemistry and Physics Journal* only after the following points are adequately addressed and elucidated.

1. In Equation 5, the symbols are not explained, the difference between  $w$  and  $\hat{w}$  is not provided. Further, the summation starts at 1 and runs over the total number of neurons, whereas in the next page (p.10941, l.19)  $N$  could also be equal to zero. To my understanding, when  $N = 0$  the trend is reduced to the linear one, correct?
2. The errors  $\epsilon_i$  are assumed to satisfy the normal distribution. Could you specify what are the initial values assumed for these errors? How the sampling is impacted by accounting for the uncertainty of the convolution scale  $s_C$  (p.10942, l. 11-15)?
3. Do you fit a parametric distribution to the sample of  $x_e$  obtained by the Monte Carlo method described in lines 9-17 of page 10942?
4. The authors choose to work with a level of significance of 0.683. How would the results be impacted if a higher level of significance (0.90 or 0.95) is assumed?
5. The method for the evaluation of the statistical significance level for the non-linear trend is not easy to track. It is not clear to me when the non-linear trend differs in a statistically significant way from the linear one. I would say that if the values of the linear trend lie within the area defined by the 68.3 significance level applied on the non-linear trend values distribution, then the difference between linear and

C4267

non-linear trends is not statistically significant. A short discussion on this very important definition should be included in the manuscript.

6. In page 10941, a way to determine the number of neurons  $N$  is presented. To my understanding, neurons should be removed as long as the leave-one-out error remains constant or decreases, but not when it increases. If this is true, then please state it clearly in the manuscript. Further, a network with more weights might be prone to overfitting and one with less weights might be inadequate to model the trend function. How are you sure that over- or under-fitting does not occur in this case ?
7. I believe that the article would benefit from a schematic picture, including the different steps necessary in order to derive the trends. To avoid lengthening the manuscript and for the sake of continuity, I would suggest that an appendix or supplement is included in the revised version of the manuscript, with emphasis on the technical aspects.
8. Please explain what is the meaning of the uncertainty intervals shown on p. 10944, l.20-25, and Fig.9.

Finally, find below a list of some of the typos and misspellings found during the reading.

- p. 10931, l. 15 : years is misspelled
- p. 10938, l.4 : missing 'that'
- p. 10938, l.14 : replace 'was' by 'were'
- p. 10938, l. 20 : remove 'the'
- p. 10939, Eq.5 : 'w' on the left hand side should be boldface, also leave a bigger space after the comma

C4268

- p. 10941, l.11 : remove 's' from 'algorithm'
- p. 10941, l.24 : replace 'the' by 'for'
- p. 10946, l.3 : read 'we evaluate whether'
- p. 10948, l. 24 : 'simple and transparent', not very convincing statement
- p. 10949, l. 14 : 'regular', do you mean significant?
- p. 10951, l. 26 : Put 's' in agglomeration, correct 'available'
- p. 10952, l. 11 : remove 'source'
- p. 10952, l. 30-31 : check names and title
- p. 10956, l. 32 : correct 'Tarrasn'

---

Interactive comment on Atmos. Chem. Phys. Discuss., 10, 10925, 2010.