Comments on the GMD Discussion Paper by J. Soares et al -
**"Refinement of a model for evaluating the population exposure in an urban area"**

24th May 2014

# General Comments

It is good to see further refinement of detailed exposure models such as EXPAND, which are a critical element of air pollution research. This paper presents useful insights into the location of most of the exposure, and some insight into the sources responsible.

Since the revised model can handle population subgroups, it would be interesting to see the exposure results summarised by age groupings, given the differing sensitivity of various age groups to air pollution.

Overall the paper presents the material in a credible way, by including caveats as appropriate ( e.g. the discussion about the fact that traffic congestion was not accounted for). In a few places some critical issues were not adequately covered (see *Specific Comments* section below).

It's not clear why the two regions studied (Helsinki Metropolitan Area, and the City of Helsinki) were analysed with different years. The paper would have presented a much clearer narrative by sticking with a single year, or by using both years together (i.e. 24 months of simulation), which would have allowed more rigorous comparisons of the city area with the broader metropolitan area.

# Specific Comments

**Page 2341, line 17 (section 2.2.2)**

I note that a detailed model was used for computing shipping emissions, but were the shipping emissions treated as coming from ground level (sea level), or were they treated as stacks with a specific height and exhaust temperature & velocity? Ship exhaust stacks can be ~40m above sea level, with hot exhausts that give rise to significant buoyancy – this can significantly affect the predicted concentrations at ground level in populated areas.

**Page 2342, lines 17-20 (section 2.3)**

This brief section is the only part of the paper that mentions meteorology, which is a critical input to air pollution modelling. For transparency and reproducibility, it is important for a modelling paper to describe or summarise all the key input datasets, including meteorological data.

I suggest a few key parameters be presented, such as example temperatures, wind speeds and mixing heights. This information can also be very useful for presenting exposure results in a weather context, for example, exposures can be summarised by month of year, or by temperature strata.

**Page 2342, line 21 (section 2.3)**

PM2.5 was treated as a tracer contaminant in this study, however it is well known that PM2.5 has a strong secondary component (both organic and inorganic aerosols). This is confirmed by the data in Table 2, which show that the concentrations in the shipping affected areas are not that much different from the urban background level – the urban background PM2.5 being strongly affected by secondary production of aerosols from various sources, with the amount of secondary production depending on meteorology. There are also other factors that affect PM2.5 -

including sea salt and wind-blown dust, which do not appear to have been modelled in this study. These sources also depend on meteorology, and geographic factors such as proximity to the sea.

Therefore it is likely that there would be some meteorological conditions (and locations) for which the predicted local contribution to PM2.5 would vary significantly from the true contribution, due to the lack of PM2.5 chemistry and treatment of natural sources. Given that population exposures were computed for each hour, the temporal variability in PM2.5 is important.

Consequently, some assessment of model performance at shorter time scales than annual (Table 2) should be presented. I note (page 2350) that a verification study is to be presented in "Aarnio et al. (2014)" , but this paper is still in preparation and is not accessible. There is a summary of findings in terms of an index of agreement and bias, but a graphical representation would be better in the current paper, e.g. a quantile-quantile plot, or a scatter plot. As part of this, some discussion of likely causes of disagreement between modelled and measured values (e.g. sources that are not modelled, or absence of chemistry, etc.) should be provided.

**Page 2345, line 4 (section 2.5)**

The definition of the infiltration factor ($F_{inf}$) appears to be a steady-state definition. In reality, there is differential equation that describes the dynamics of how outdoor air enters a building (accounting for air exchange rate, and other factors such as chemical transformation and destruction within the building).

This study has computed hourly PM2.5 values and used these to derive exposure estimates in various microenvironments, with different infiltration factors. The use of steady-state infiltration factors probably means that the true indoor variation in concentrations may have been overestimated, since the outdoor PM2.5 that enters a building will take some time to do so, resulting in a smoothing out of the variation in the outdoor signal.

The paper should note the impact of using steady state infiltration factors on estimated hourly exposure values.

**Page 2349, line 13 (section 3.1)**

This text (under the Results & Discussion section) refers to small-scale wood combustion, and notes that this is a significant source, but then indicates that dispersion modelling was not possible due to a lack of knowledge about the spatial distribution of the source. It is standard practice in regional airshed modelling to spatially allocate the distribution of domestic emissions to a surrogate such as population density, in the absence of any more detailed information. This would have been more realistic than simply neglecting the source altogether.

A related question arises which is about all other domestic and small business emissions. In typical urban environments these sources can be significant, not only in terms of total emissions but also because they often emit pollutants near to ground level (unlike large industry which typically uses tall stacks). In policy terms, such sources are even more important, because (compared with industry and motor vehicles) they tend to be poorly regulated, and can assume a greater relative importance over time as vehicle sources become more tightly controlled. An overview of what is known about domestic and small business emissions in this region would be useful.

# Technical Corrections

**Page 2366, line 12 (ABSTRACT)**

The revised model can also be used for evaluating intake fractions for various pollutants, source categories and population subgroups"

I suggest replacing "evaluating" with "estimating", since this is a computer simulation of intake fractions.

**Page 2336, line 17 (ABSTRACT)**

"The population exposure originated from the long range transported background concentrations was responsible for…"

This appears to be incorrect grammar. Should this have said "The population exposure originating from long range transported background concentrations .." ?

**Page 2337, lines 3-5 (INTRODUCTION)**

Since the urban population spends typically 80–95% of their time indoors (Hänninen et al., 2005; Schweizer et al., 2007), the exposure to outdoor particles is dominated by exposure in indoor environments.

Delete the word "outdoor" here.

**Page 2337, lines 5-6 (INTRODUCTION)**

The most simplistic approaches ignore indoor and outdoor conditions.

Presumably this should read "ignore the differences between indoor and outdoor air".

**Page 2347, line 6 (section 2.6)**

The units presented for *Ei* are incorrect. **μg $m^{-3}$ $s^{-1}$** should be **μg $m^{-3}$ s**

**Page 2368, Figure 2**

In the figure caption, the greek letter "**μ**" should be used instead of "**u**" to represent the prefix micro (for micrograms per cubic metre). Also, if these are annual averages, this should be clearly stated.

**Page 2373, Figure 6**

Population exposure is presented here in units of concentration x persons "(μg$m^{-3}$ x no. People)". However a true exposure metric must include the duration of exposure (i.e. a time element). I expect what is being presented here is annual average concentration x population density, which has the time unit removed by computing the annual average concentration. This needs to be clarified, and perhaps some appropriate terminology introduced, e.g. "population exposure per year."