1 General comments :

The paper is well-written so it is easy to read and it presents a useful development for an algorithm for an operational model of a line source. The description of the algorithm is the strongest part of the paper. The description of the comparison with diffusion tube data and with other models begs many questions. Hopefully the questions can be answered by some additions to the text based on existing material, although some questions reveal weaknesses in the validation and model comparisons.

2 Specific comments

I agree with the comments of the other referee on: the abstract, Romberg integration, Model error and Comparisons to the HV formulation. Section 1.

-Please, refer to answers given to referee 1.

The time resolution of the background concentration and meteorological data is not given, is it hourly?

-Yes background concentration and meteorological data are hourly. We now mention it in the text.

The measurement height, location with respect to the kerb and hence the classification (kerbside, roadside etc) of the diffusion tubes should be given.

-Measurements were performed at a 2 meter height. The passive diffusion tubes are located at various distances from the roads, ranging from about 10 m to 100 m. Because of the large number of passive diffusion tubes used in this study, it is not feasible to specify the individual locations.

There is no mention in the paper of treatment of non-road emissions and emissions from roads other than those modelled explicitly. These should not be neglected.

- In the modeled area, roadway traffic is the major contributor to NO2 concentrations. Other non-road sources of NOx are included indirectly via the background concentrations of NO and NO2. For roadway emissions, only major roads are treated explicitly and emissions from minor surface roads are neglected and we, mention that emissions from those non-modeled roads could be a source of uncertainty (page 16, line 3-4).

Section 2.2

The choice of 89 is arbitrary and would be expected to lead to over-estimates by the HV formulation. P3354 line 23 says this is the case but Figure 3 shows that HV underestimates compared with Polyphemus when winds are close to parallel to the road.

-Yes indeed, 89 degree was an arbitrary value and it turns out that HV model underestimate the reference. However, with 89.9 degree the HV model would have diverged and would then give a higher value than the reference.

Section 2.3
The nature of the "correction functions" and their effect should be described as they are the key difference between HV & Polyphemus.

-The Polyphemus Gaussian model, including the correction functions, was fully described in the article titled:
Briant, R., Korsakissok, I., and Seigneur, C.: An improved line source model for air pollutant dispersion from roadway traffi c, Atmos. Environ., 45, 4099–4107, 2011. 3, 6, 11, 13

Therefore, this work focuses on the evaluation of the Gaussian model and only improvements made on the original model formulation (e.g., the use of a Romberg integration to account for the road width) are described in detail here.

Section 3.1
COPERT 3 would be expected to under-estimate NOX and NO2 emissions from certain vehicle classes: NOX emissions from diesel cars under urban driving conditions do not appear to have declined substantially up to and including Euro 5. There is limited evidence to suggest that this same pattern may occur for motorway driving conditions; and NOX emissions from HGV vehicles equipped with SCR reduction are much higher than expected when driving at low speeds.

-We thank the reviewer for pointing out this possible source of underestimation of the mobile source emissions and added it to our discussion of possible reasons that the NOx emissions are underestimated.

There is no mention of whether diffusion tubes were co-located to assess accuracy and co-located with automatic monitors to calculate a local bias-adjustment factor. Properly verified and bias-adjusted diffusion tube results would be useful for comparing annual averages. 1 month measurements from diffusion tubes are not going to provide very good data (absolute values) for model validation, but are useful in showing spatial variation.

-Measurements were validated with a comparison with chemiluminescence measurements. We now mention it in the text. However, the bias of the passive diffusion tube with respect to the chemiluminescent method may depend on the NO2 level and, therefore, may not be representative of the bias of passive diffusion tubes at other locations (i.e., measuring higher or lower NO2 concentrations). We agree with the reviewer that the main advantage of a large network of passive diffusion tubes is to capture the spatial variation of NO2 concentrations.

Here we see the spatial variation (Figure 2) and we see that the model is not capturing the range of values, see comment under section 3.2.

I don't understand why a primary NO2 value of 10% was considered for the base case as it was obviously an under-estimate.

-The 10% NO2 was considered in the base case because it is the default value used by the ADMS model in its standard regulatory applications.

Section 3.2
The "rural" and "urban" options of the HV and Polyphemus models are not explained. It's puzzling that the "urban" option does not apply in the suburbs of Paris.

-Rural and urban options refer to Pasquill stability classes. now specify this point in the text. The rural option is justified because of the relatively low population density in the modeled suburban area (on the order of 2000 inhabitants per km2) compared to downtown Paris (20 000 inhabitants per km2). Neverthells, the "urban" option was also used and results are given in supplementary materials.

The RMSE is very large compared with the mean and what the scatter plots show is the model not capturing the monitored range in concentration, that is presumably due to variations in receptor siting and the range of traffic flow.

-Section 3.6 gives some explanation of those results (background concentration, underestimation of emission...). The location of passive diffusion tubes regarding to the road is taken into account by models with the distance from receptors to the corresponding road sections; therefore, the fact that the models underestimate the actual variability in the measurements results from averaging processes in the modeling (e.g., emissions, meteorology) and simplifying some physical processes (e.g., flat terrain assumption).

No mention has been made of whether canyon effects and other local effects have been considered. These are well known to be important when modelling near road concentrations.

- Most Gaussian models are assume a flat domain (although some models allow some corrections for impacts on simple terrain features such as impacts on elevated terrain) and, indeed, the models used here were applied with this flat terrain assumption. This assumption may result in model uncertainties. However, the suburban setting of this modeling domain minimizes the influence of major features such as street canyons. Such street canyons could indeed be a source of local effects in other areas (e.g., downtown Paris) that should be taken into account by adding appropriate parameterizations. We added a short discussion of this point in the manuscript.

Section 3.3

No reason is given for using a smaller domain for the comparison with ADMS-Urban, nor how the subset of 62 receptors compared with the whole set of 242 (proportion of kerbside and roadside sites).

-The ADMS-urban application cannot be conducted on the entire road network at once because of limitations in the number of road sections treated and several simulations needed to be conducted. For this study, the same year-specific meteorological inputs were used for all three models (ADMS-urban, Polyphemus and HV) and, therefore, the comparison was limited to a subdomain for which ADMS-urban could be applied in a single model simulation.

The study mixes up differences in meteorological processing, chemistry and fundamental dispersion algorithms with the line source representation. The effect of chemistry scheme could have been excluded by using the NOX/NO2 post-processing for ADMS-Urban. There is a comment on NOX concentrations but this needs to be backed up with tables and figures. It would be useful to run a baseline comparison between ADMS-Urban and Polyphemus modelling a line source perpendicular to the wind direction, and then a parametric study varying the wind direction so that the effect of the line source formulation alone could be identified.

-We agree with the reviewer that the effect of the NO2/NO chemistry can be excluded as the cause of the difference between ADMS-urban and Polyphemus because the same trend was obtained with NOx concentrations and we have rewritten the text to clarify this point.

As the meteorological data actually used was not spatially varying it would have been much better to use hourly data from a synoptic station. Similarly, it would have been much better to use hourly monitored rural and urban background concentrations rather than modelled value. These steps would have reduced uncertainty in the model inputs.

- Hourly measurements of meteorology were not available. Therefore, the use of modeled values was necessary. Although air quality hourly data could be obtained, the use of modeled air quality data was consistent with the use of modeled meteorological data.

Section 3.5

This section describes major differences between HV & Polyphemus and ADMSUrban. The former use broad categories of stability and only use wind speed in postprocessing results. This is a major difference in the fundamental dispersion algorithms and should have been identified separately as shown above.

-In equation 7 the wind velocity is just a coefficient in front of the equation, which does not depend on the orientation of each road section, the emission rates and dispersion coefficients. We simply use this fact to improve computational time; there is no change in the dispersion algorithm.

Although ADMS-Urban was run on a different PC, as computer run time is given as a key, determining factor at a couple of points the run times and computer specifications should have been reported, so the reader has the information.

-The ADMS simulation was ran not only on a different computer, but also for additional receptors that were not used in this study. The additional receptors significantly increase the computational time and make a comparison between the computational times of ADMS and Polyphemus and HV unfeasible.

Section 3.6

If you are going to compare modelled and values from a dedicated monitoring campaign, it is not sufficient to pick up a bias adjustment factor from the literature. A local bias adjustment factor or factors should have been determined as part of the campaign.

-As mentioned above, a comparison with chemiluminescence measurements was made to ensure that there was reasonable agreement between the two methods. However, this comparison is representative of a single location and may not be applicable to every single passive diffusion tube monitoring location.

Section 4

I recommend deleting this late, unsubstantiated comment on the performance of Polaris3D, that has only been mentioned once previously in the paper. With a resolution of 5km it is not “fit for purpose” as a model kerbside and roadside concentrations so I don’t understand why that comparison was made. Your point surely is that you intend to embed the new line source algorithm in the Polaris3D model which has a horizontal resolution of 5km.

-We have deleted this part of the conclusion as suggested by the reviewer.

Tables 1-3.
The mean values should be labelled as “monthly mean values of NO_2”.

-We made the change.

Figures 4 & 5
Should there be some data points?

-There are no hourly measurements of NO_2 concentrations, therefore, only modeled values could be compared here.