This paper investigates the possible causes of NOx model-measurement mismatch for high-resolution regional modelling. The authors use WRF-Chem, run for 1 year (through 2014) at high (3x3km) resolution, with TNO emissions downscaled using local data to improve the resolution of emissions. They use spectral decomposition of the modelled and observed timeseries to extract the contributions of processes across a range of timescales to the NOx system, finding that the model is good at capturing the long term and synoptic behaviour of the system, but struggles to capture variations at diurnal and intra-diurnal timescales. NO2 measurements at urban and suburban background measurement sites are underestimated (as seen in previous studies too), and they suggest that difficulties in capturing the behaviour of vertical mixing and traffic emissions are most likely to cause the discrepancies between the model and measurements. The sensitivity of simulated concentrations to a top-down correction factor for NOx emissions using traffic data is investigated - based on this analysis they suggest that traffic emissions should be scaled non-linearly (as opposed to the current linear scaling) with traffic counts.

The paper is well written, and the study carried out in a careful and methodic manner. The difficulties of trying to compare ground-level, single point measurements with model predictions averaged over a (much) larger grid cell do limit the analysis that can be carried out in such a study, but I think the authors recognise and work within those limitations well, and the insights gained from this study will be relevant to the ACP readership. I recommend publication, subject to a few minor corrections.

Specific Comments:
1) Comparison of the modelled NO2 concentrations in this study with those in Kuik et al (2016) does suggest that the modification to the chemical mixing routines in WRF-Chem in order to set a minimum vertical mixing value helps reduce the overprediction of NO2 at nighttime. I am, however, uncomfortable with some of the choices made in this routine: using fixed minimum mixing rates, and emissions as a proxy for land use, both seem like clunky fudges to me. CO from shipping emissions would lead to its being activated over the oceans, for example, which I think is counter to the intended influence of this modification. It is good that the authors have noted the use of this modification in this study (and disappointing that if it has been used in previous published studies, that it does not seem to have been documented as being used), as this will add to growing evidence in the literature (c.f. Hu et al, 2013) that this issue exists and needs addressing in a more rigorous and methodical manner.

As changes other than this modification have been made to the model setup between this study and Kuik et al (2016), it would be useful for readers if a figure illustrating the differences in NO, NO2, NOx and O3 concentrations in model simulations with and
without this modification could be added to the supplementary material (as this would
be illustrative, it would only need to be for a couple of days, not for the entire year).

In addition, recent communication from Georg Grell on the wrf-chem-discussions list
indicated that the ACM2 scheme now conducts mixing of chemical species within the
PBL physics routines (improving consistency between the treatment of chemical and
meteorological tracers). Though this is probably outside the scope of this study, have
the authors considered investigating if this scheme improves nighttime pollutant con-
centrations?

2) The locations of measurement stations are not very clear in Figure 2, and it is cur-
rently impossible just using this paper to determine the exact location of individual
named stations within the study. An extra map (either in the main paper, or in sup-
plementary material), with the locations of named measurements stations made much
clearer, would be very useful.

3) In Section 6.2 the authors use wind speed and direction data to select data for
investigating the impact of changing urban traffic emissions on downwind model NOx
predictions. Were there a specific reasons for choosing 2 m/s wind speed and 72 data
pairs as your cut-off points? How dependent on your data selection criteria are the
changes in model bias between simulations - if you adjusted the criteria by +/-10%
would your results change greatly?

Technical Corrections:

1) It might be helpful to expand the abbreviations MAM, JJA, SON, and DJF when they
first appear, to make the paper more accessible to those who are not familiar with these
common abbreviations.

2) page 3, line 7: it would be more correct to use "and" instead of "with" in this sentence.

3) page 8, 2nd paragraph: urRV needs proper formatting twice here.

4) page 16, line 2: the correction factor should be "f", not "F"?

5) page 17, line 19: should it be "wind direction bin", instead of "wind speed bin"?

6) page 34, figure 8: your description misses out the colour (blue?) of the sensitivity
simulation data

7) page 38, table 3: in the table caption, the units of MB and RMSE need correcting
from ug/m3 to K and m/s (respectively).

References:

Hu, X.-M., P. M. Klein, and M. Xue (2013), Evaluation of the updated YSU planetary
boundary layer scheme within WRF for wind resource and air quality assessments, J.

Interactive comment on Atmos. Chem. Phys. Discuss., https://doi.org/10.5194/acp-2017-1037,
2017.