Author response to minor revisions request from Anonymous Referee #1

We greatly appreciate the effort of both anonymous reviewers to read the revised manuscript and the minor comments provided by one of the reviewers to further improve the manuscript. Below, we address each comment individually. Referee comments are given in *italic*, author response are given in normal font, changes made to the text are given in blue. This document is finalized by a markdown version of the manuscript including all the changes.

Thank you for these revisions. Indeed a lot of work has gone into improving this; I appreciate it. I have one further minor revision request, and the rest are technical corrections.

We have added the following statement in the Acknowledgement section to express our appreciation to both the reviewers: We greatly appreciate the two anonymous reviewers for their critical and constructive comments. Their effort has contributed to major improvements of this manuscript.

Minor Request:

-Thank you for clarifying the SCM. Now that I have a somewhat better understanding of it and it's comparison to WRF-Chem, I have one additional request. I think it's necessary to include an intercomparison of the modeled PBLs in Figure 6. There should be a fifth panel (or perhaps the NO panel can be removed). In this additional panel, you should display the diurnal cycle of the WRF-Chem PBL and the SCM PBL (and ideally the ERA5 PBL too if possible) for the same time frame.

We have included the comparison of WRF-Chem, SCM and ERA5 PBLH's in the manuscript and removed the NO panel. We have made minor adjustments to the result section where we discuss and refer to the (sub-) figure and updated the figure caption. We have updated the text where we discuss the results of PBLH to:
Figure 7d shows a comparison of the SCM, WRF-Chem simulated- as well as the ERA5 reanalysis boundary layer height for the grid resembling the location of Bogota. The SCM is showing a substantially deeper daytime maximum boundary layer with more day-to-day variation compared to WRF-Chem and ERA5 reanalysis data. The SCM also simulates a relatively fast afternoon transition to suppressed nocturnal mixing conditions reflected by a nocturnal inversion layer which agrees well with the ERA5 boundary layer height being shallower than that simulated by WRF-Chem. Interestingly, the SCM simulation results in a better representation of the observed diurnal cycle of urban pollutant mixing ratios, especially regarding the observed early morning maximum CO and NO_x and minimum O_3 concentrations, without requiring the hypothesized enhancement in emissions. This stresses that, besides application of higher-resolution emission inventories and model experiments, the diurnal cycle in boundary layer dynamics (and advection) should be critically evaluated in models such as WRF-Chem which, however, would then also require urban boundary layer structure measurements.

Technical Requests:

-Abstract: The abstract is too long as currently written. I understand that there is no word limit for this journal, but this Abstract could be more concise (ideally <400 words). Some specific suggestions would be to remove Line 9 - 17 starting with sentence "The comparison... mean column". Also, the last 8 lines, Lines 27 - 34 can be shortened significantly.

Thank you for your suggestions to reduce the abstract. We have reduced the abstract from 34 lines to 25 lines especially removing some of the technical model set-up statements. Also in the last 8 lines we have removed the line "Futhermore, this study ... mixing ratios" since it has been implicitly mentioned before when discussing the SCM simulations.

-Abstract: With all of that said, there is one missing item in the Abstract and it is the mention of anthropogenic NOx in comparison to the other sources. It would be helpful to include this near Line 7 (similar to how it is mentioned in the Conclusions near Line 430).

We have added the anthropogenic (but also the soil biogenic to 'close' the budget) and rewritten the sentence to: The model indicates the largest contribution by lightning emissions (1258 Gg N yr^{-1}), even after already significantly reducing the emissions, followed by anthropogenic (933 Gg N yr^{-1}), soil biogenic (187 Gg N yr^{-1}) and biomass burning emissions (104 Gg N yr^{-1}).

-Line 123: Comma between "areas" and "the"

Added ,.

-Line 168: "dependens" --> "depends"

Changed to depends.

-Line 201: I am a bit uneasy that the lower limit filter is 1 pixel (i.e., any location with 2+ pixels are included). It seems from Figure 2, that the limit could be 5 and a large fraction of the area would still be valid. Is there a reason why 1 pixel is the lower limit. For example is a particular city/location excluded if the lower filter limit is 5 instead? If so, please mention this in the text.

In an earlier stage of the study we have looked at the sensitivity to increasing/decreasing this number of available OMI measurements (from ≥2 to ≥4 to ≥6). It indicated that the sensitivity to the frequency distributions (figure 5d and 5e) is low and the mean/median and 90% CI remain relatively the same. However, by applying the 2+ criterion (as in the manuscript) we omit 10% of the total surface area of the domain. By applying the 4+ criterion this would become 20% and by applying the 6+ criterion this would become 34%. This would mainly go at the cost of a part of the Amazon rainforest, but also some of the cities in the mountainous regions. We agree that a 4+ or 6+ criterion would mean a better estimate of the monthly mean NO_2 column. However, we choose to stick with the 2+ criterion because the focus of the study was not to identify an accurate estimate of the monthly mean NO_2 columns in January, but more on their spatial distributions,

comparison with WRF-Chem and link with the four distinct emission regimes which benefits from a larger surface area cover.

Please note that after the interactive discussion stage we have updated Figure 2 from the exact number of available measurements to the % of measurements available during the simulation period (January 2014). So 2 pixels corresponds to ~5%.

-Figure 3. Minor suggestion, but could be helpful to include the phase "AMF change" directly on the top center of the panel a, and the phrase "Column NO2 change" directly on the top center of panel b.

We have updated the figure with the corresponding titles.

-Figure 4b. Thanks for addressing this figure caption, but I don't think it's perfectly clear yet. I see the dark orange color in the center of this panel, yet this color is not in the colorbar (biomass burning is closest but appears to be a light orange). The reason I am confused is that the dark orange, at first glance, seems to represent a combination of biomass burning and anthropogenic. Yet, your explanation makes it seem like the dark orange color is only biomass burning. Either way, dark orange is very hard to differentiate from red. Perhaps a way to remedy this, would be to make the biomass burning a brighter shade of yellow, so that it is clearly different from the other colors.

Thank you for your suggestion to improve the image. The lighter yellow colors are in their turn (to our perception) hard to distinguish from the lighter green colors (which are next to eachother in the Orinoco region). The most optimal combination (to our perception) is to change the biomass burning colors to purple. We have updated the figure, but also updated the (new) figure 8 to keep the colors consistent throughout the manuscript.

-Figure 5. Please write the word "NO2" somewhere on this figure, and in the figure caption. Also would be ideal to display the word "WRF-Chem NO2" directly on the top center of the panel a, and "OMI NO2" on the top center of panel b. And maybe "WRF-Chem - OMI" on panel c.

We have added the corresponding titles and we have added NO_2 to the caption and added the suggested titles to the panels.

-Figures A1 & C1: It's unclear why these are in an Appendix. If possible, it would be nice to include these in the main text when they are referenced. There is certainly room for them if the journal rules allow it.

We have replaced the figures to the main text and adjusted the text/references where needed.

Evaluation of nitrogen oxides sources and sinks and ozone production in Colombia and surrounding areas

Johannes G.M. Barten[1], Laurens N. Ganzeveld[1], Auke J. Visser[1], Rodrigo Jiménez[2], and Maarten C. Krol[1,3]

[1]Wageningen University, Meteorology and Air Quality Section, Wageningen, the Netherlands
[2]Department of Chemical and Environmental Engineering, Air Quality Research Group, Universidad Nacional de Colombia – Bogotá, Colombia
[3]Institute for Marine and Atmospheric Research Utrecht, Utrecht University, Utrecht, the Netherlands

Correspondence: Johannes G.M. Barten (sjoerd.barten@wur.nl)

Abstract. In Colombia, industrialization and a shift towards intensified agriculture have led to increased emissions of air pollutants. However, the baseline state of air quality in Colombia is relatively unknown. In this study we aim to assess the baseline state of air quality in Colombia with a focus on the spatial and temporal variability in emissions and atmospheric burden of nitrogen oxides (NO_x = NO + NO_2) and evaluate surface NO_x, ozone (O_3) and carbon monoxide (CO) mixing ratios. We quantify the magnitude and spatial distribution of the four major NO_x sources (lightning, anthropogenic activities, soil biogenic emissions and biomass burning), by integrating global NO_x emission inventories into the mesoscale meteorology and atmospheric chemistry model WRF-Chem at a similar resolution (~25km) to the EDGAR anthropogenic emission inventory and OMI remote sensing observations. The model indicates the largest contribution by lightning emissions (1258 Gg N yr^{-1}), even after already significantly reducing the emissions, ~~and the lowest contribution by~~ followed by anthropogenic (933 Gg N yr^{-1}), soil biogenic (187 Gg N yr^{-1}) and biomass burning emissions (104 Gg N yr^{-1})~~to total NO_x emissions within the WRF-Chem domain~~. The comparison with ~~in situ measurements is bound to urban areas whereas the use of remote sensing data allows to also evaluate air quality in remote regions. WRF-Chem was set up for a domain centered over Colombia at a similar (~25km) resolution as the OMI observed NO_2 vertical columns as well as the EDGAR anthropogenic emission inventory. However, this apparently poses a challenge regarding comparison with urban air quality measurements. Air mass factors were recalculated based on the vertical distribution of NO_2 within WRF-Chem, with respect to the coarse (1°x1°) a priori profiles because WRF-Chem is expected to better resolve spatial contrasts in NO_2 profiles. The main reason for recalculation is a more consistent satellite-model comparison. The~~ OMI remote sensing observations indicated a mean bias of tropospheric NO_2 columns over the whole domain (WRF-Chem minus OMI) ~~is~~ of 0.02 (90% CI: [-0.43, 0.70])·10^{15} molecules cm^{-2}, which is <5% of the mean column. However, the simulated NO_2 columns are overestimated and underestimated in regions where lightning and biomass burning emissions dominate, respectively. WRF-Chem was unable to capture NO_x and CO urban pollutant mixing ratios, both in timing and magnitude. Yet, WRF-Chem was able to simulate the urban diurnal cycle of O_3 satisfactory but with a systematic overestimation of 10 ppb due to the equally large underestimation of NO mixing ratios and, consequently, titration. This indicates that these city environments are in the NO_x-saturated regime with frequent O_3 titration. We conducted sensitivity experiments with an online meteorology-chemistry single column model (SCM) to evaluate how

WRF-Chem subgrid scale enhanced emissions could explain an improved representation of the observed O_3, CO and NO_x diurnal cycles. Interestingly, the SCM simulation, showing especially a shallower nocturnal inversion layer, results in a better representation of the observed diurnal cycle of urban pollutant mixing ratios without an enhancement in emissions. This stresses that, besides application of higher-resolution emission inventories and model experiments, the diurnal cycle in boundary layer dynamics (and advection) should be critically evaluated in models such as WRF-Chem to assess urban air quality. Overall, ~~the presented approach shows~~ we present a concise method ~~, integrating in situ and remote sensing observations,~~ to quantify air quality in regions with ~~a limited measurement network~~limited surface measurements by integrating in situ and remote sensing observations. This study ~~not only~~ identifies four distinctly different source regions ~~, but also shows the~~ and shows their interannual and seasonal variability ~~of these sources~~ during the last ~~one and a half decade. Furthermore, this study shows that with a critical consideration of advection and (nocturnal) boundary layer mixing, relatively coarse anthropogenic emission inventories can give reasonable results regarding the diurnal cycle of urban pollutant mixing ratios.~~ 1½ decade. It serves as a base to assess scenarios of future air quality in Colombia, or similar regions with ~~distinct~~ contrasting emission regimes, complex terrain and a limited air quality monitoring network~~, as a function of further industrialization and land use changes~~.

1 Introduction

Nitrogen oxides (NO_x = NO + NO_2) are one of the main precursors of lower atmospheric ozone (O_3). Exposure to NO_x has an adverse effect on human health on acute and long-term basis (Panella et al., 2000; Wolfe and Patz, 2002). Likewise, O_3 is toxic to humans (WHO, 2003) and can also reduce agricultural yields (Ashmore and Marshall, 1998). Therefore, accurate monitoring and predictions of surface concentrations of these air pollutants are key. Especially in densely populated regions air pollution has been a major concern and is expected to even have larger impacts in the future due to the continuous urbanization and increasing emissions from for example traffic.

Anthropogenic NO_x is produced in combustion processes and is an indicator of industrial activity and transportation as well as other anthropogenic activities like biomass burning and agricultural activities. Anthropogenic sources add up to ~70% (~50% industrial activity/transportation, ~20% biomass burning) of the total global annual NO_x emissions (Lamarque et al., 2010). In addition to anthropogenic sources, natural sources contribute to total nitrogen budgets. NO emissions from soils add up to ~12-20% of the global NO_x emissions on a yearly basis (Bradshaw et al., 2000; Ganzeveld et al., 2002a; Jaeglé et al., 2005; Vinken et al., 2014). Lightning emissions are estimated to attribute on average 10-18% to the global yearly NO_x emissions (Pickering et al., 2016). In the tropics (35°N - 35°S), anthropogenic activities (7.81 Tg N yr^{-1}), biomass burning (8.28 Tg N yr^{-1}), soil emissions (5.44 Tg N yr^{-1}) and lightning discharges (6.33 Tg N yr^{-1}) all contribute an approximately equal fraction to the total NO_x emission budget (Bond et al., 2002). A modeling study in the tropics must therefore provide accurate estimates of all these source categories.

In Colombia, where economy is thriving after a period of civil war (Vargas et al., 2015), further industrialization and intensified agriculture have already resulted in- and are expected to further increase- NO_x emissions (Ganzeveld et al., 2010). Previously, Grajales and Baquero-Bernal (2014) aimed to assess the air quality of Colombia with a relatively coarse (2.5°x2.0°) 3D global

model (GEOS-Chem), whereas other studies focused mostly on air pollution of other compounds in cities using local emission inventories (Zárate et al., 2007; Kumar et al., 2016; González et al., 2018). Currently, there is a lack of understanding of the baseline state of air quality in Colombia on regional scale. Following from this, an application of inventories of the different sources of NO_x (and other pollutants) and covering both Colombia and its surrounding, upwind, areas can give valuable information about the current state of air quality in Colombia. This is also essential to determine how air quality might change in the future, e.g., due to further urbanization and land use changes such as the conversion to oil palm (Vargas et al., 2015).

Up until now, Colombia does not have an air quality monitoring network covering the entire country. Current measurement sites are mainly located in or close to the major cities. The rural areas, which are now undergoing rapid land use changes, do not have air quality stations nearby. This makes air quality monitoring for the whole country a challenging task. The use of satellite data, to observe species like NO_2 and formaldehyde (CH_2O), is a valuable tool to fill the gaps and evaluate air quality in remote regions (Bailey et al., 2006; Kim et al., 2009; Webley et al., 2012). However, satellite retrievals in the tropics are often limited by the presence of clouds.

During the last decades, computational advances have increased the possibility to conduct more detailed meteorology and air quality studies (Bauer et al., 2015). The recognition of the effects of chemical composition of the atmosphere on meteorology have stimulated the development of online coupled meteorology/chemistry models (Baklanov et al., 2014). Nowadays, these models can be run for a large range of temporal and spatial scales. Not only the models, but also global emission inventories have considerably improved in spatial resolution during the last decades (González et al., 2018). Even though they may not provide enough spatial detail and heterogeneity for local scale (< 1 km) studies, e.g. to compare with in situ observations, they have provided essential information regarding emissions for regional scale (∼20 km) studies (Saide et al., 2012; Ghude et al., 2013). In this study, rather than using high resolution urban emission inventories (e.g. González et al., 2018), we will demonstrate the importance of boundary layer mixing and advection in the comparison of simulated and observed in situ measurements.

The primary objective of this study is to assess the current baseline state of air quality in Colombia, diagnosed with a focus on NO_x, using global emission inventories in a regional atmospheric chemistry model resolving the atmospheric chemistry and meteorology at a resolution comparable to that of the emission inventories as well as the remote sensing observations. Furthermore, we evaluate surface NO_x, O_3 and CO mixing ratios in urban regions. We are aware that concerns about air quality in Colombia are generally not limited to smog photochemistry mainly involving O_3-NO_x-VOC chemistry. Actually high concentrations of particulate matter might pose the largest risk to public health in many Colombian urban areas (Kumar et al., 2016). However, in this study we focus on NO_x as an insightful metric to assess the spatial and temporal patterns in air quality in this region given its role in O_3 photochemistry as well as the availability of remote sensing observations to be integrated with a bottom-up model analysis. In this study we use the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) (Grell et al., 2005). The model outcomes will be compared to in situ measurements and satellite retrievals to address the performance of the model both at the surface and integrated over the troposphere. This evaluation of surface and total column —using a highly resolving coupled meteorology-air quality model including the identification of different NO_x sources— seeks to fill the gaps between local scale (González et al., 2018; Zárate et al., 2007; Kumar et al.,

2016) and larger scale studies (Grajales and Baquero-Bernal, 2014). This study also includes an evaluation of the interannual and seasonal variability of air pollution for the different source regions during the last one and a half decade. This analysis is not only useful to address the representativeness of the performed simulation and to identify the baseline state of air quality in Colombia but also justifying potential use of the modeling system such as WRF-Chem to assess future changes in air quality using future anthropogenic emission and land use change scenarios (e.g. Ganzeveld et al., 2010).

2 WRF-Chem & its emission inventories

2.1 Model: WRF-Chem

In this study we use WRF-Chem (Grell et al., 2005) version 3.7.1. WRF is a non-hydrostatic mesoscale numerical weather prediction model used for operational and research purposes. Figure 1 shows the WRF-Chem domain including cities and regions that we refer to in this research.

The simulation was set up for one domain with a spatial resolution of 27 km centered at 4.89 °N, 71.07 °W. The entire domain consists of 100 grid points in both the North-South and the East-West direction with 60 vertical levels —in a sigma coordinate system— up to 50 hPa. The simulation length is one month (also given technical constrains on conducting much longer integrations with WRF-Chem), with a spin-up time of 24h, covering the whole month of January 2014. Selection of this study period is motivated by the fact that January is the dry season in Colombia where loss of remote sensing data due to presence of clouds is minimized (see Sect. 3.1.1). In addition, in ~~Appendix ??~~ Sect. 5 we show how the selected study period can be put into context of the baseline state of air quality in Colombia using the interannual and seasonal variability in emission sources inferred from the remote sensing observations. ~~A more detailed analysis on this is presented in Sect. 5.~~

The European Centre for Medium-range Weather Forecasting (ECMWF) ERA-Interim product provides us with the meteorological initial and boundary conditions. The chemical initial and boundary conditions are constrained with the Copernicus Atmosphere Monitoring Service (CAMS) near-real-time dataset. The boundary conditions are updated every six hours on a spatial resolution of 0.4° (∼44 km) with 60 vertical model levels. For January 2014, boundary conditions of O_3, NO_x, CO, SO_2 and CH_2O are available. For tropospheric chemistry, the Carbon-Bond Mechanism version Z (CBM-Z) chemical scheme (Gery et al., 1989; Zaveri and Peters, 1999) is used here because it has been successfully implemented and tested in similar studies (Gupta and Mohan, 2015). Additional parametrization schemes used in this research are listed in Table 1.

2.2 Emission inventories

Anthropogenic emissions are described by the Emission Database for Global Atmospheric Research (EDGAR) dataset for greenhouse gases (Janssens-Maenhout et al., 2017) and Non-Methane Volatile Organic Compounds (NMVOCs) (Huang et al., 2017). Emission estimates are gridded on a 0.1°x0.1° resolution. EDGAR emissions are monthly estimates implying constant emissions over the whole simulation. In this study we use the EDGAR-HTAP emission inventory updated for 2010. EDGAR-HTAP uses nationally reported emissions combined with regional scientific inventories. For this research we assumed that

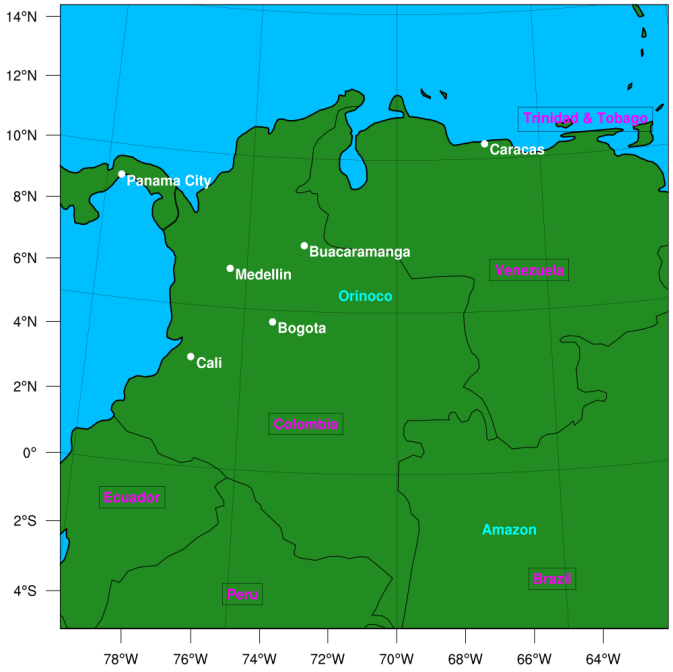

Figure 1. WRF-Chem domain including countries (pink), major cities (white) and regions (blue).

Table 1. WRF-Chem physical and chemical parametrization schemes.

WRF-Chem option	Configuration
Physical parameterizations	
Microphysics	Morrison 2-moment (Morrison et al., 2009)
Long wave radiation	RRTM (Mlawer et al., 1997)
Short wave radiation	Dudhia (Dudhia, 1989)
Surface layer	Monin-Obukhov (Janić, 2001)
Land surface	Noah (Chen and Dudhia, 2001)
Boundary layer	YSU (Hong et al., 2006)
Cumulus	Grell 3D (Grell and Freitas, 2013)
Lightning option	P&R neutral buoyancy (Price and Rind, 1992)
Chemical options	
Gas-phase	CBM-Z (Gery et al., 1989; Zaveri and Peters, 1999)
Photolysis	F-TUV (Tie et al., 2003)
Lightning chemistry	Single-mode vertical distribution (Ott et al., 2010)

95% of the total anthropogenic emission of NO_x is emitted as NO and 5% as NO_2 (Carslaw, 2005). VOC (Volatile Organic Compounds) speciation is according to Archer-Nicholls et al. (2014). In densely populated urban areas, the anthropogenic emissions are dominated by vehicular emissions (Dodman, 2009). These emissions have a clear diurnal and weekly variation in contrast to emissions from the industry sector (Streets et al., 2003). Zárate et al. (2007) estimated traffic emission factors for Bogotá using in situ measurements and inverse modeling techniques. To account for this diurnal and weekly variation we multiply the EDGAR emissions with the hourly and daily emission factors presented by Zárate et al. (2007).

The Global Fire Emissions Database version 4 (GFEDv4) dataset (Randerson et al., 2015) provides us with the biomass burning emissions. GFED is available on a spatial resolution of 0.25°x0.25°, approximately the same size as the WRF-Chem grid cells. Biomass burning NO_x emissions are assumed to be completely in the form of NO.

Natural emissions of VOCs from terrestrial ecosystems are considered in this study using the Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGANv2.1) (Guenther et al., 2012). Biogenic emissions are updated on-line using the WRF-Chem simulated surface temperature, soil moisture, leaf area index and photosyntetically active radiation. MEGAN also provides estimates of soil biogenic NO emissions.

The lightning-NO_x parametrization scheme (Price and Rind, 1992), embedded in WRF-Chem, is used to account for NO_x emissions by lightning. For this study we used an IC:CG (intracloud:cloud-to-ground) ratio of 2:1 constant over the whole domain with a flashrate factor of 0.1. Per lightning flash (both for IC and CG strikes), it is assumed that 250 moles of NO are emitted (Miyazaki et al., 2014). It has to be noted that in an initial simulation, using standard WRF-Chem settings (flashrate factor = 1.0 & 500 moles of NO per strike), resulted in a significant overestimation of the lightning emissions (see Sect. 4.1) (Bradshaw et al., 2000; Miyazaki et al., 2014; Murray, 2016). The settings we used resulted in a twentyfold decrease of lighting emissions compared to standard WRF-Chem settings.

3 Observations of atmospheric composition

3.1 Satellite retrievals

Observational data on the distribution of NO_2 is retrieved from the Ozone Monitoring Instrument (OMI) onboard the National Aeronautics and Space Administration (NASA) Aura satellite (Levelt et al., 2006). OMI measures, among other pollutants, NO_2 column densities (Boersma et al., 2007) with daily, global coverage. The pixel size of 24x13 km^2 may be coarse for particular applications, such as assessing urban pollution, but is suitable to assess contrasts in regional-scale air quality with apparent contrasting emission regimes. In addition, the resolution of the OMI observations is also comparable to the resolution of the anthropogenic emission inventory.

In this research we use the Quality Assurance for Essential Climate Variables (QA4ECV) NO_2 data product (Boersma et al., 2018). The measured slant columns —the tilted path directly from sun through the atmosphere to surface back to satellite— are converted to vertical columns using Air Mass Factors (AMFs) [-] by

$$VCD = \frac{SCD}{AMF}, \tag{1}$$

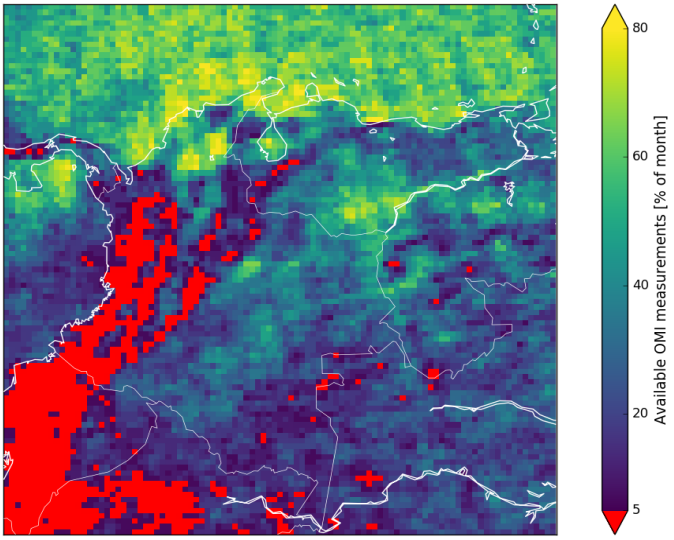

Figure 2. Spatial distribution of the available OMI measurements in January 2014 after filtering has been applied.

where VCD and SCD are the Vertical Column Density and the Slant Column Density [molecules cm^{-2}], respectively. The AMFs define the relation between slant column and the vertical column above a pixel based on external information on e.g. surface albedo, scattering, clouds and the vertical distribution of NO_2 (Boersma et al., 2011). The vertical distributions of NO_2 in the QA4ECV product, which are used to calculate the AMFs, are simulated by the TM5-MP global chemistry transport model at a resolution of 1°x1° (Williams et al., 2017).

3.1.1 Filtering

We follow the data filtering recommendations by the QA4ECV consortium. Presence of clouds (cloud radiance fraction > 0.5) led to omission of 63% of OMI NO_2 data. Figure 2 shows the amount of OMI data per WRF-Chem grid cell after filtering the observations of January 2014. Especially above mountainous regions, where we also find the main urban areas of Bogotá and Medellín, there is a lack of available data due to the continuous presence of clouds. This limits the quality of the measurements and which increases the uncertainty in the averaged tropospheric NO_2 column (Boersma et al., 2018). On average 9 data points per grid cell are available for this specific domain in January 2014, but with a large spatial heterogeneity. Some areas have >20 data points and other only two valid observations in this month.

3.1.2 AMF recalculation

The AMF ~~dependens~~ depends on assumptions of the state of the atmosphere and surface (e.g. surface albedo, cloud fraction, vertical distribution of NO_2) at the specific moment and location of a satellite observation (Lorente et al., 2017). This vertical sensitivity is described by an averaging kernel, which describes the relationship between the true column and the estimated, or retrieved column (Boersma et al., 2016). High-resolution models such as WRF-Chem are expected to better represent spatial

gradients in NO_2 profiles compared to coarse-scale global models such as GEOS-Chem or TM5-MP. Consqeuently, we can expect WRF-Chem to better resolve strong enhancements in tropospheric NO_2 VCDs in densely populated areas. Using grid sizes comparable to the size of such large urban areas is a major advantage of this procedure (Krotkov et al., 2017). The application of the averaging kernel is shown to reduce systematic representativeness errors for a satellite-model comparison (Boersma et al., 2016). We can recalculate the AMF based on the a priori concentration profile x_a (from the TM5-MP model) and the concentration profile in the high-resolution model x_m, in this study WRF-Chem (Boersma et al., 2016):

$$M'(x_m) = M(x_a)\frac{\sum_{l=1}^{L} A_l x_{m,l}}{\sum_{l=1}^{L} x_{m,l}}, \quad (2)$$

where $M(x_a)$ is the tropospheric AMF used in the retrieval, A_l are the elements of the averaging kernel for each l^{th} vertical layer and $M'(x_m)$ is the recalculated AMF. In a next step, the new VCDs can be calculated by dividing the SCDs (retrieved by the satellite) with the recalculated AMFs (Eq. (1)).

Fig. 3 shows the difference in AMFs and the subsequent effect on the tropospheric NO_2 columns for the WRF-Chem domain. On average, we find a mean decrease in AMF of 0.05 with a standard deviation of 0.15. Regarding inferred changes in the VCD due to this recalculation of AMF, we find a mean increase in the VCD of $0.02 \cdot 10^{15}$ molecules cm^{-2} ($\sim$3% of the average VCD) and a standard deviation of $0.07 \cdot 10^{15}$ molecules cm^{-2}. Above cities (e.g. Caracas, Bogotá, Medellín), we find mostly a decrese in AMF (Fig. 3a), and, consequently an increase in VCD (Fig. 3b). This indicates that there is more NO_2 present near the surface in WRF-Chem compared to TM5. This is consistent with our expectation that WRF-Chem better captures the sub-1°x1° processes that are not resolved by TM5, such as the localized urban emissions. Furthermore, we find increases in VCD ($0.5 \cdot 10^{15}$ molecules cm^{-2}) above the Amazon region. Decreases in VCD (Fig. 3b) are found mostly across the border from Colombia to Venezuela, better known as the Orinoco region (Fig. 1). This reflects a higher abundance of NO_2 higher up in the troposphere from lightning sources, which are potentially underestimated by the global TM5 model (Silvern et al., 2018), combined with less NO_2 near the surface. We also find two isolated hot spots of decreases in VCD in southern Venezuela which correlate well with topography within the WRF-Chem domain which is less well resolved in the coarser resolution of TM5. Despite locally significant changes in VCDs, a domain average of $0.6 \cdot 10^{15}$ molecules cm^{-2} indicates that the difference in the NO_2 a priori profiles of TM5 compared to those in WRF-Chem does not lead to domain-wide significant changes in VCDs.

3.1.3 Comparison of OMI and WRF-Chem

In WRF-Chem we calculate the tropospheric NO_2 column by integrating from the surface to the tropopause, determined to be approximately the 50^{th} model level ($\sim$90 hPa, $\sim$17km). This level is determined based on the average temperature profile (from surface to 50 hPa) of the complete simulation. Furthermore, to assess the daily differences in total NO_2 columns from OMI and WRF-Chem we need to co-sample their data points. For Colombia, OMI passes around 17-19 UTC (1:00 PM local time). Grid points with none or only one measurement after filtering (see Fig. 2) will be completely discarded. In this way we

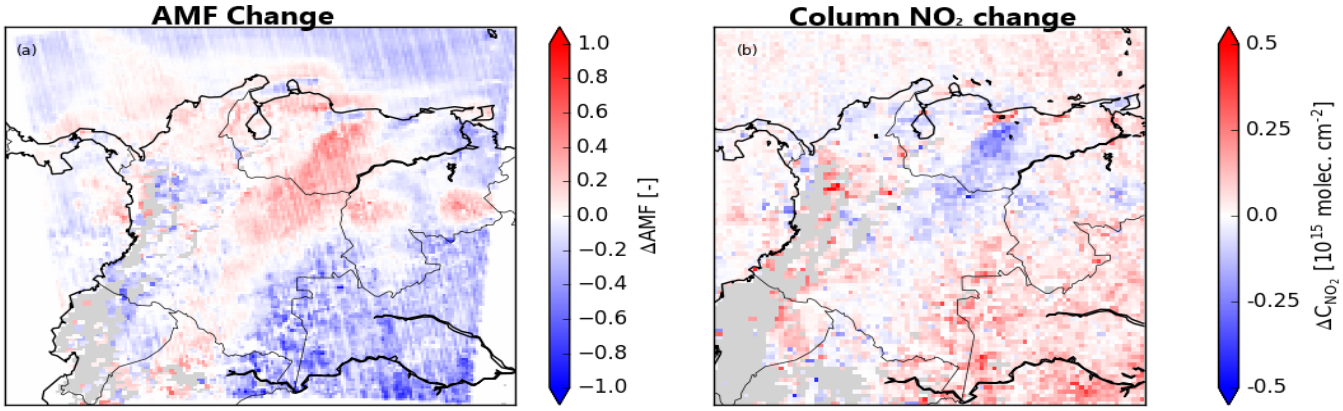

Figure 3. Spatial distribution of (a) the AMF difference (recalculated minus QA4ECV standard product, ΔAMF [-]) in January 2014 based on the WRF-Chem simulation and (b) the subsequent effect on the NO_2 column difference (recalculated minus QA4ECV standard product, ΔC_{NO_2} [10^{15} molecules cm^{-2}]) on the WRF-Chem grid.

aim to get a reliable comparison between WRF-Chem and OMI, which enables us to determine systematic biases in the regions dominated by different emission sources.

3.2 In situ data

To further evaluate the model, observational data from air quality monitoring stations in Colombia are used. These include 30 stations confined to four cities in Colombia: Bogotá, Bucaramanga, Cali and Medellín (see Fig. 1). Observational data consists of 1-hourly averaged CO, NO, NO_2 and O_3 concentrations. The complete availability and locations of the station within the WRF-Chem domain can be found in Table A1. The data and information on the measurements is publicly available at http://sisaire.ideam.gov.co/ideam-sisaire-web/ (last access: March 2020). In this paper we only show the results of Bogotá also because the comparisons for the other cities show similar results. Even on the still coarse resolution of the current WRF-Chem simulations, we expect that the evaluation of the temporal variability in simulated and observed concentrations indicates how well the model captures some of the key drivers of atmospheric pollution.

4 Results

4.1 Nitrogen emission budgets and distribution

First of all, we identify the major sources of NO_x within the domain of this study. The anthropogenic and biomass burning emissions are prescribed using their inventories whereas soil NO and lightning NO_x emissions are explicitly simulated in WRF-Chem. Some large cities contribute dominantly to the total NO_x emissions (Fig. 4a). Total emissions are in the order of 200-1000 kg N km^{-2} $month^{-1}$ for the Colombian cities. However, largest NO_x emissions, according to the EDGAR inventory, are found in and around Caracas, Venezuela. All these emissions can be attributed to anthropogenic emissions as reflected by a

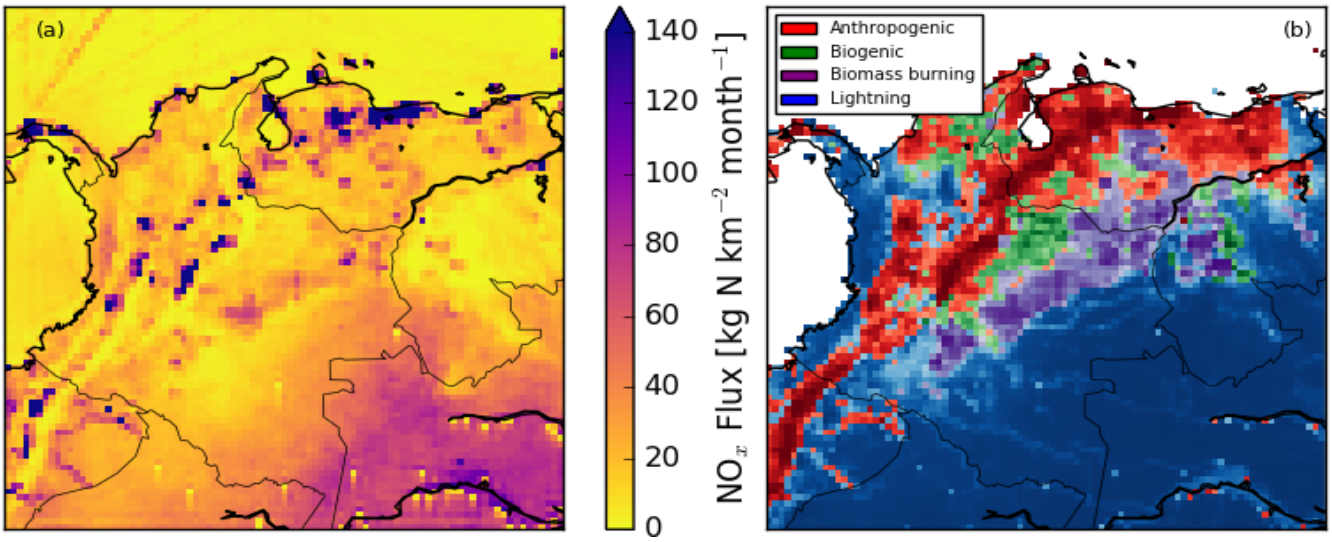

Figure 4. Spatial distribution of (a) the total NO_x flux [kg N km^{-2} month^{-1}] in January 2014 and (b) the dominant emission source per grid cell over land. In (b), the saturation of the color indicates the % of the total NO_x flux coming from the dominant emission source going up to 100% for the darkest colors.

$\sim$100% contribution of anthropogenic emissions to the total emissions shown in Fig. 4b. Another major source of NO_x is found in the south-east of the domain with values ranging up to 100 kg N km^{-2} month^{-1}. In this region, with land cover dominated by rainforest, large convective systems are present generating thunderstorms with associated lightning NO_x emissions. They appear to be the most important emissions in this region (Fig. 4b) also because anthropogenic and biomass burning emissions are mostly absent (with some exceptions near rivers). Biomass burning and soil biogenic emissions seem to be the most prominent sources of NO_x across the Colombian-Venezuelan border (Fig. 4b), in the Orinoco region, in our model study. This region is dominated by savanna type grasslands which emit a relatively high amount of soil NO_x but also have a high probability of catching fire. NO_x emissions in these regions are up to one or two orders of magnitude smaller compared to anthropogenic emissions, but, on the other hand, cover a larger area.

Lightning NO_x emissions is the most dominant emissions source over land in 63% of all grid cells, followed by anthropogenic- (22%), biomass burning- (9%) and biogenic (6%) emissions (Fig. 4b). Since we use four different emission inventories, all with their own estimates and uncertainties, the distinct contrasts in the spatial distribution of emission sources will be key to determine spatially heterogeneous biases in satellite retrievals compared to WRF-Chem. From budget calculations integrating over the whole domain, and using these January emissions to infer a NO_x emission budget expressed per year (see Table 2) we also find that lightning NO_x is the largest source in absolute terms, followed by anthropogenic, biogenic and biomass burning respectively.

Table 2. Total NO_x emissions in the WRF-Chem domain per source category using January 2014 emissions to infer yearly total NO_x emissions [Gg N yr^{-1}]. The percentage surface area over land where the emission source is dominant is also indicated.

NO_x source category	Emission [Gg N yr^{-1}]	Surface area [% of total land surface area]
Lightning	1258	63%
Anthropogenic	933	22%
Biogenic	187	6%
Biomass burning	104	4%
Total	2482	

4.2 WRF-Chem & OMI comparison

To assess whether WRF-Chem is able to reproduce filtered and recalculated NO_2 VCDs satisfactorily we check for the spatial and frequency distributions for both WRF-Chem and OMI (see Fig. 5). For WRF-Chem we find a wide range of column densities (Fig. 5a). We find very low VCDs ($\sim 0.3 \cdot 10^{15}$ molecules cm^{-2}) over the Caribbean sea and across the eastern border of Colombia into Venezuela. High VCDs in WRF-Chem are simulated above the major Colombian cities and the northeastern part of the domain ($\sim 5 \cdot 10^{15}$ molecules cm^{-2}) while the highest VCDs are simulated above the city of Caracas with values up to $8 \cdot 10^{15}$ molecules cm^{-2}. Similar to WRF-Chem, we find the lowest VCDs over the Caribbean sea in OMI (Fig. 5b). Also, we find the highest VCDs above major cities —most pronounced for Caracas and Medellín— but the magnitude of the OMI observed VCD ($\sim 2.5 \cdot 10^{15}$ molecules cm^{-2}) is much smaller compared to WRF-Chem. In OMI we find low VCDs above the Amazon rainforest.

The large WRF-Chem VCDs ($\sim 5 \cdot 10^{15}$ molecules cm^{-2}) we find in the northeastern part of the domain (Fig. 5c) seem to reflect mostly the role of the imposed CAMS boundary conditions. High mixing ratios of NO_2 are found in the lowest CAMS model layers advected into the WRF-Chem domain by the prevailing easterlies. High VCDs are not seen in the OMI retrievals where we only find a small plume coming from Trinidad & Tobago transported westward. The model overestimation of the VCD above Caracas might be due to an overestimation of anthropogenic emissions but this is not supported by a systematic major overestimation above cities (for example, we find no overestimation above Panama City or Bucaramanga). However, the EDGAR emission inventories are based on the year of 2010, when Venezuelan economy was still at its maximum (Wang and Li, 2016). After 2010, Venezuelan economy and oil production have declined strongly (Wang and Li, 2016) and therefore also emissions of pollutants have likely also decreased substantially. Lastly, we find a systematic overestimation in the WRF-Chem simulated VCD above the Amazon rainforest. Even though the model overestimation is small in absolute terms ($\sim 0.5 \cdot 10^{15}$ molecules cm^{-2}) it is quite substantial relative to the background mixing ratios. In this region, soil NO_x release is small, anthropogenic activities are hardly present and there are no known sources of biomass burning during January 2014. Also the role of advection from outside the model domain by the prevailing easterly winds seems to be limited indicated by lower VCDs at the eastern most grid cells. Consequently, overestimation in the simulated VCDs is most likely due to the simulated major

influence of lightning NO_x emissions in this region (Fig. 4b), even though they have already been significantly reduced relative to the standard settings (see Sect. 2.2). However, we have to take into account that the OMI retrievals used for this comparison are near the detection limit and reflect those conditions when cloud formation, and therefore lightning production, is less active resulting in very low VCDs. In contrast, the co-sampled WRF-Chem columns might reflect simulated cloud cover resulting in production of NO by lightning. Nonetheless, the question whether lightning production was actually present or that it could not be picked up by OMI, being less sensitive to the presence of NO_2 below clouds, remains unanswered.

Remarkably, we find a region with systematic underestimations ranging from the center of Colombia to the northeastern border with Venezuela (the Orinoco region). In this region, there is no presence of major cities and lightning NO_x emissions are small. The discrepancy we find might be due to missing agricultural- or biomass burning emissions. Localized enhancements observed in OMI ($\sim 2.5 \cdot 10^{15}$ molecules cm^{-2}) might also be caused by biomass burning emissions since enhanced soil NO_x are expected to result in a more homogeneous enhancement of VCDs over a larger area with smaller intensities. We find that this intensity of biomass burning is not picked up by the WRF-Chem simulation using the GFED biomass burning inventory.

Figure 5d shows, for both WRF-Chem and OMI, the frequency distribution in the NO_2 VCD. We find that both model simulated and observed VCDs show similar distributions, peaking at approximately the same VCD. However, WRF-Chem shows more outliers especially regarding the simulation of high NO_2 VCDs. The 90% confidence interval of the WRF-Chem simulated VCDs is $(0.33,1.33) \cdot 10^{15}$ molecules cm^{-2} while for OMI the 90% confidence interval is $(0.32,1.06) \cdot 10^{15}$ molecules cm^{-2}, with medians of $0.59 \cdot 10^{15}$ and $0.56 \cdot 10^{15}$ molecules cm^{-2}, respectively.

We find the median and mean of the absolute overestimation by WRF-Chem to be $0.02 \cdot 10^{15}$ and $0.09 \cdot 10^{15}$ molecules cm^{-2} respectively (Fig. 5e). The 90% confidence interval equals $(-0.43,0.70) \cdot 10^{15}$ molecules cm^{-2}. The distribution is approximately Gaussian with a standard deviation of $0.53 \cdot 10^{15}$ molecules cm^{-2} but somewhat left-skewed indicating an overestimation by WRF-Chem. This confirms the finding that WRF-Chem is able to produce on average good estimations for vertical NO_2 columns above Colombia. However, over- and underestimations can be significant, e.g. larger than the $\sim$10% uncertainty in monthly averaged OMI VCD over polluted regions (Boersma et al., 2018), due to numerous factors in both the model setup and the characteristics of the retrievals.

4.3 Surface mixing ratios

~~We retrieve January averaged diurnal cycles in O_3, NO_x and CO surface mixing ratios (Fig. 7) by removing the significant spread in observed surface mixing ratios and by averaging the day-to-day variation in both the model and observations. We also compare the model~~ Figure 6 shows the model simulated and observed temporal evolution in NO_x and O_3 over the whole simulation period~~(see Fig. 6, Appendix **??**).~~ . WRF-Chem represents the lower limit (generally mid-day) of observed NO_x mixing ratios, but is unable to simulate the observed maxima (generally morning rush-hour) up to 200 ppb for particular events. Regarding O_3, WRF-Chem resembles the upper limit of observed mixing ratios ($\sim$40 ppb) during daytime but is unable to reproduce the observed low (<5 ppb) nighttime mixing ratios.

We retrieve January averaged diurnal cycles in NO_x, CO and O_3 surface mixing ratios (Fig. 7a-c) by removing the significant spread in observed surface mixing ratios and by averaging the day-to-day variation in both the model and observations. Regard-

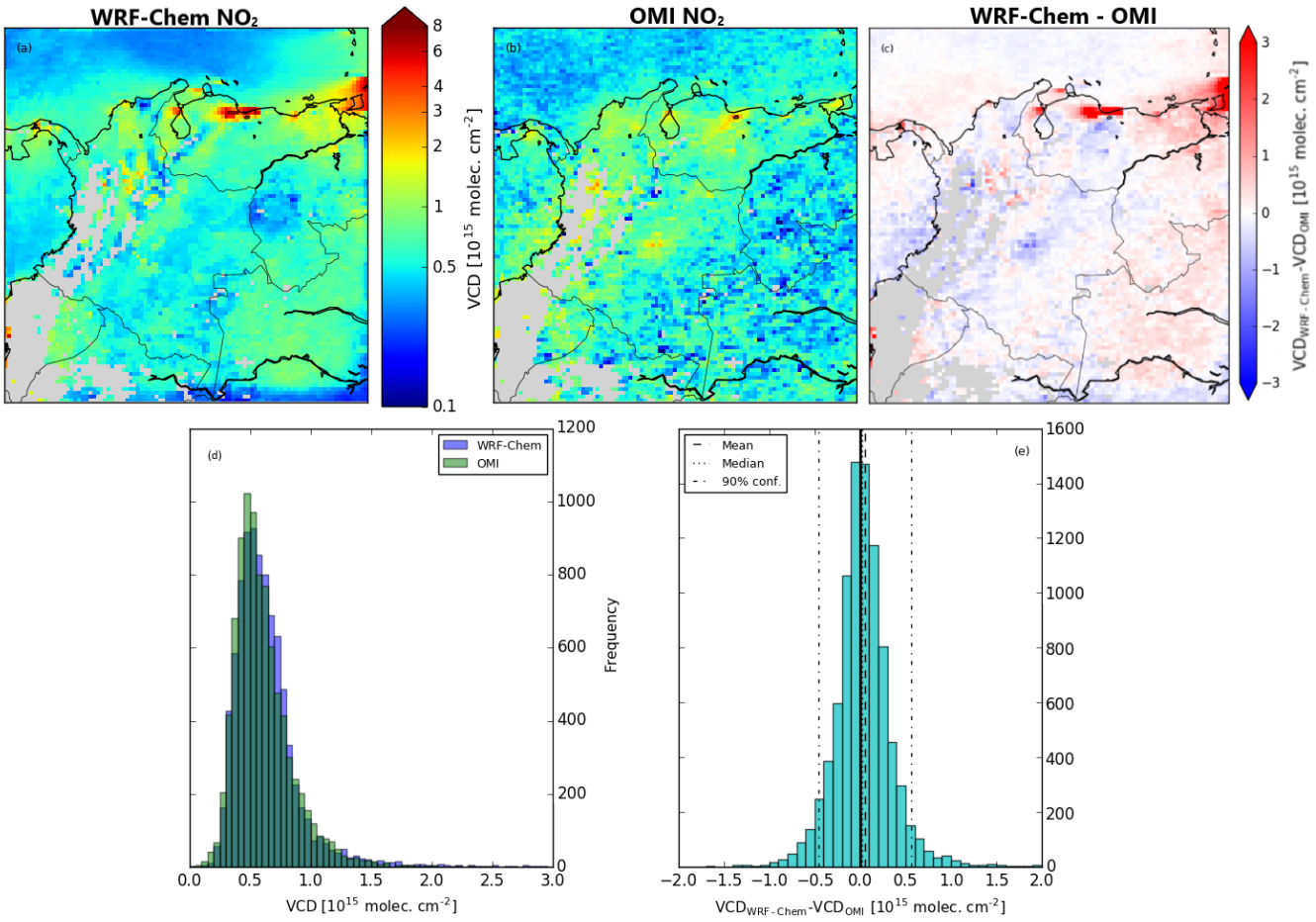

Figure 5. Spatial distribution of averaged co-sampled NO_2 Vertical Column Densities (VCD) [10^{15} molecules cm^{-2}] for (a) WRF-Chem, (b) OMI recalculated retrievals and (c) the absolute difference (WRF-Chem minus OMI) between the two as well as (d) frequency distribution of both WRF-Chem and OMI over the whole domain and (e) the distribution of the absolute difference between the two per grid point including mean (dashed line), median (dotted line) and 90% confidence interval (dashed-dotted line).

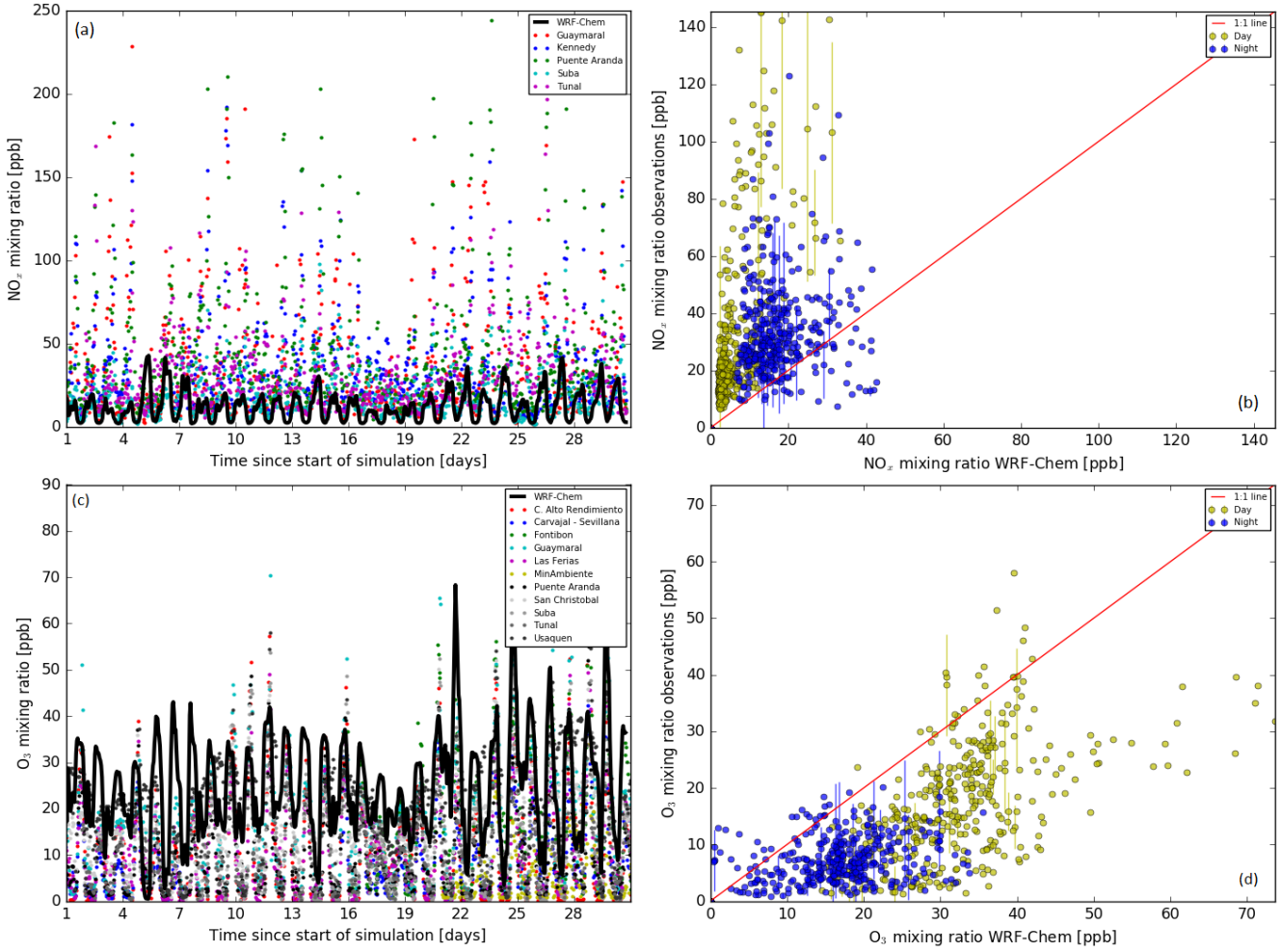

Figure 6. Temporal evolution of (a) NO_x and (c) O_3 mixing ratios [ppb] in Bogotá for WRF-Chem (black solid line) and all available observational stations (coloured points). Scatter plots of the WRF-Chem output compared with averaged (b) NO_x, (d) O_3 mixing ratios [ppb] from the stations are split up in day (yellow) and night (blue). The error bars indicate the standard deviation of the observational data from randomly sampled points (not all standard deviations are shown for visual purposes).

ing simulated NO_x, the average nocturnal mixing ratios are 20 ppb, with some day-to-day variation (standard deviation = 10 ppb), and a minimum of 2 ppb during daytime but with less day-to-day variation (Fig. 7a). The observations reach peak mixing ratios around 7:00 local time where vehicular emissions during rush hour are mixed in a shallow boundary layer increasing NO_x mixing ratios to 85 ppb on average. After rush hour mixing ratios decrease due to decreasing emissions, increasing boundary
300 layer height and decreasing NO_x lifetime. It is interesting to note that there does not seem to be a clear signal of evening rush hour in the NO_x measurements and simulation.

The averaged diurnal cycle of CO in WRF-Chem shows a similar pattern to that of NO_x (Fig. 7b). WRF-Chem shows daytime

mixing ratios of ∼150 ppb (well above rural background mixing ratios of 100 ppb) and ∼350 ppb during nighttime while the surface measurements show a significantly larger variation. Averaged surface measurements during rush hour exceed CO mixing ratios of 1500 ppb. Some measurement stations even report mixing ratios above 3000 ppb. Even though nighttime emissions are mixed over a smaller boundary layer we find that WRF-Chem underestimates surface mixing ratios of CO by a factor of 4 during rush-hour and by a factor of 2 for nighttime conditions. These ratios are similar to the NO_x ratios in Fig. 7a. Since CO has a relatively long lifetime compared to that of NO_x we argue that observed differences regarding simulated and observed CO mixing ratios reflect issues regarding the representativity of the WRF-Chem grid simulated pollutant levels, including the representation of emissions and online simulated meteorological conditions, relative to the footprint of the surface observations.

We find that for WRF-Chem most of the NO_x is present as NO_2 with NO mixing ratios being very close to 0 ppb (~~Fig. 7e~~not shown here). In contrast, the observations show that most of the NO_x is present as NO. For WRF-Chem we find a $[NO]/[NO_2]$ ratio of ∼0.32 (±0.13) during daytime and ∼0.07 (±0.04) during nighttime while for the surface measurements these ratios are ∼1.11 (±0.40) and ∼0.89 (±0.38) respectively. The observations that show a high $[NO]/[NO_2]$ ratio might be indicative of a location close to local sources, e.g. roads. The abundant fresh NO emissions at these locations quickly react with O_3 forming NO_2. The surplus NO, however, pushes the $[NO]/[NO_2]$ ratio up. Indeed, a simulated underestimation by WRF-Chem of 10 ppb NO during nighttime is consistent with a simulated overestimation of 10 ppb O_3 (Fig. 7~~d~~c). We also find that in WRF-Chem, the formation of O_3 immediately starts at 6:00 local time (sunrise) while for the observations we find the lowest mixing ratios at 7:00 local time due to the extra NO titration caused by rush hour. Nonetheless, it seems that chemical production and destruction rates of O_3, as well as other processes contributing to the overall magnitude and diurnal cycle in O_3, e.g., entrainment and deposition, are well captured by WRF-Chem considering the similar shape and amplitude of the diurnal cycle.

To test the hypothesis that the model-data mismatch over Bogatá is caused by a too coarse model resolution, or a misrepresentation of emissions, we conducted additional experiments with a Single Column chemistry-meteorological Model (SCM), also being previously applied for an analysis of observations of the plume of pollution downwind of the city of Manaus (Brasil) (Kuhn et al., 2010). The SCM simulates online, similar to WRF-Chem, atmospheric chemistry processes, including anthropogenic and natural emissions, gas-phase chemistry, wet and dry deposition and turbulent and convective tracer transport as a function of meteorological and hydrological drivers, surface cover, and land use properties (Ganzeveld et al., 2002b, 2008). For these urban area simulations with the SCM we have modified the surface cover properties by prescribing surface roughness at 1 meter, assuming a reduced vegetation fraction of 0.6, using a city area albedo of 0.18 and nudged the SCM meteorology with wind speed, moisture, and temperature profiles from the WRF-Chem simulation. The SCM is also nudged with long-lived tracers such as O_3, NO_x and CO above the boundary layer using WRF-Chem mixing ratios. Finally, we also used the same emissions, including diurnal cycle, as in the WRF-Chem simulation.

Using these settings in the SCM results in simulated January average diurnal cycles in NO_x ~~, NO~~ and CO, quite different from WRF-Chem but, in better agreement with the observations in terms of 30-day average diurnal cycles, maximum early morning peak and daytime minimum mixing ratios of ~~CO,~~ NO_x and ~~NO~~ CO (Fig. 7a-b). The skewed O_3 diurnal cycle (Fig. 7c) is also better reproduced compared to WRF-Chem although the overestimation of the maximum afternoon mixing ratios is

larger. ~~Interestingly, the SCMsimulation, showing especially a shallower nocturnal inversion layer compared to that simulated in WRF-Chem~~Figure 7d shows a comparison of the SCM, WRF-Chem simulated- as well as the ERA5 reanalysis boundary layer height for the grid point resembling the location of Bogotá. The SCM is showing a substantially deeper daytime maximum boundary layer with more day-to-day variation compared to WRF-Chem and ERA5 reanalysis data. The SCM also simulates a relatively fast afternoon transition to suppressed nocturnal mixing conditions reflected by a nocturnal inversion layer which agrees well with the ERA5 boundary layer height beging shallower than that simulated by WRF-Chem. Interestingly, the SCM simulation results in a better representation of the observed diurnal cycle of urban pollutant mixing ratios, especially regarding the observed early morning maximum CO and NO_x and minimum O_3 concentrations, without requiring the hypothesized enhancement in emissions. This stresses that, besides application of higher-resolution emission inventories and model experiments, the diurnal cycle in boundary layer dynamics (and advection) should be critically evaluated in models such as WRF-Chem which, however, would then also require urban boundary layer structure measurements.

5 Discussion

The integration of global emission inventories in a highly resolved coupled meteorology-air quality model (WRF-Chem), with roughly the same spatial scale, allowed us to assess the state of- and contribution by different sources to the air quality in Colombia and neighbouring countries, diagnosed with a focus on NO_x. We identified four major sources of NO_x in Colombia which were implemented in WRF-Chem partly through emission inventories (anthropogenic and biomass burning) and partly through emission models (soil NO and lightning). Using January emissions to infer a NO_x emission budget expressed per year we found that lightning NO_x emissions are the main source for the domain applied in this study, with 1258 Gg N yr^{-1}. These are followed by respectively anthropogenic (933 Gg N yr^{-1}), soil biogenic (187 Gg N yr^{-1}) and biomass burning (104 Gg N yr^{-1}) emissions. Figure 8 shows the averaged VCDs over the regions dominated by one of the four emissions classes (Fig. 4b). Figure 8a shows the yearly trends in OMI NO_2 VCDs. The domain averaged anthropogenic or lightning dominated regions seem to have relatively low interannual variability. The biogenic and biomass burning dominated regions show most interannual variability which also seem to correlate with El Niño years (https://origin.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ONI_v5.php, last access: October 2019), with the exception of 2015. Colombia is relatively warm and dry during El Niño years (Córdoba-Machado et al., 2015). Figure 8a indicates that biogenic- and biomass burning emissions might have increased during El Niño years reflected by higher January monthly mean VCDs above those regions. To further put the findings of the combined WRF-Chem and OMI VCDs for January 2014 in context, Figure 8b shows the seasonal variability in OMI NO_2 VCDs. We find that biogenic, biomass burning and anthropogenic emissions show a maximum at the end of the dry season (March). For biogenic and biomass burning this is most likely caused by increased emissions while for the domain dominted by anthropogenic emissions this is most likely caused by advection of NO_x, emitted by biogenic or biomass burning sources, located upwind. For lightning NO_2 VCDs we find a maximum in August/September. We find that this is caused by an increase in NO_2 VCDs in the south-eastern part of the domain (Amazon region), not shown here. The large standard deviation in the biomass burning NO_2 VCDs again indicates the large interannual variability. Based on this further analysis of the

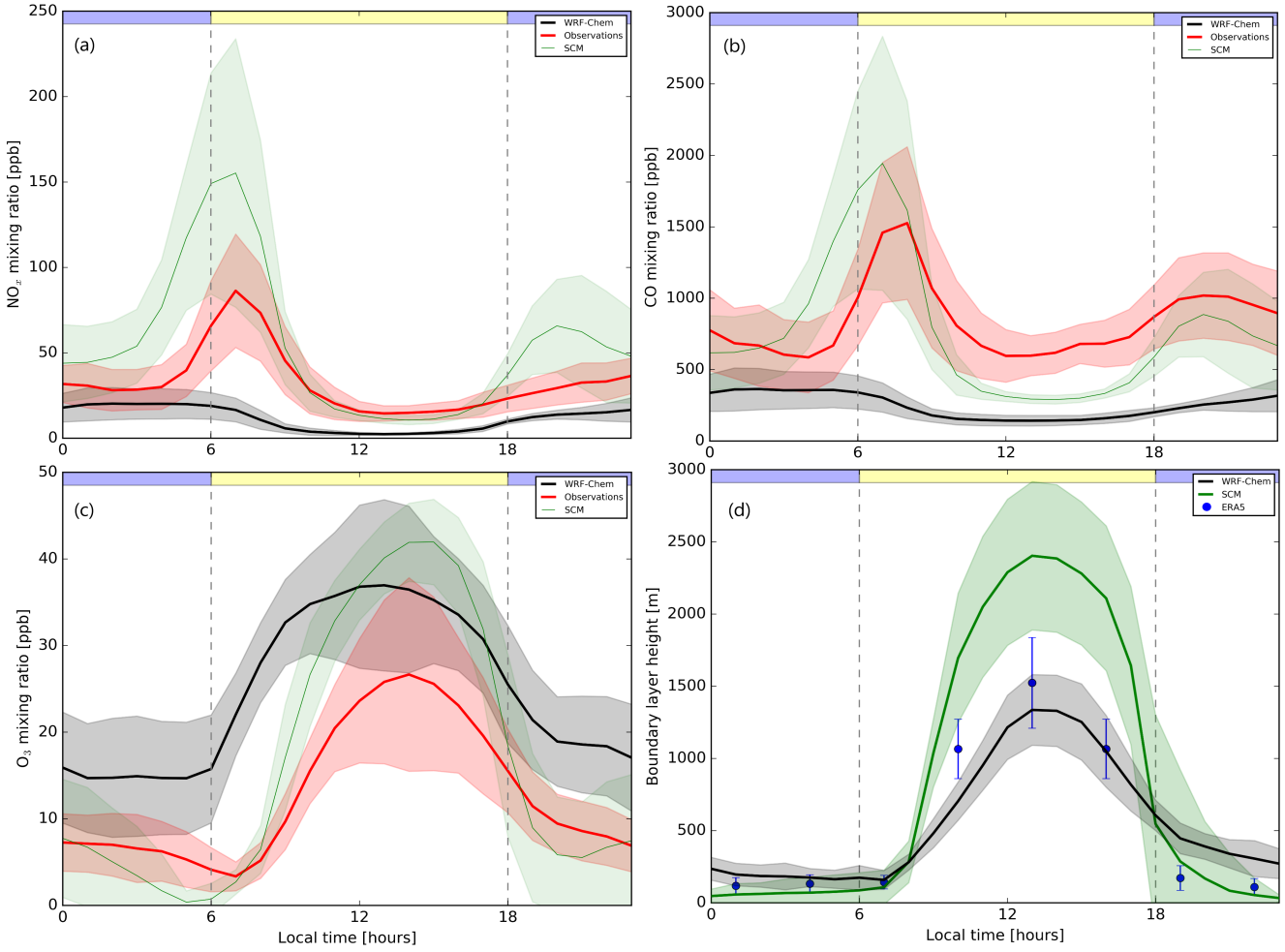

Figure 7. Averaged diurnal cycle of (a) NO_x, (b) CO, (c) ~~NO and (d)~~ O_3 mixing ratios [ppb] and (d) Boundary layer height [m] in Bogotá for WRF-Chem (black solid line), averaged observational data (red solid line)~~and~~ , SCM (green solid line) and ERA5 reanalysis data (blue dots). The black, red and green shadings and blue error bars indicate the 30-day standard deviation of WRF-Chem, observations~~and~~ , SCM and ERA5 respectively. The vertical lines, blue (night) and yellow (day) shading indicate daytime and nighttime.

long-term trends in OMI NO_2 VCDs, we argue that the 2014 simulation and remote sensing data analysis is a reasonably good approximation of the baseline state of air quality in Colombia, at least regarding NO_x. However, we have to take into account the interannual and seasonal variability in NO_x emissions in interpreting the OMI data and WRF-Chem results.

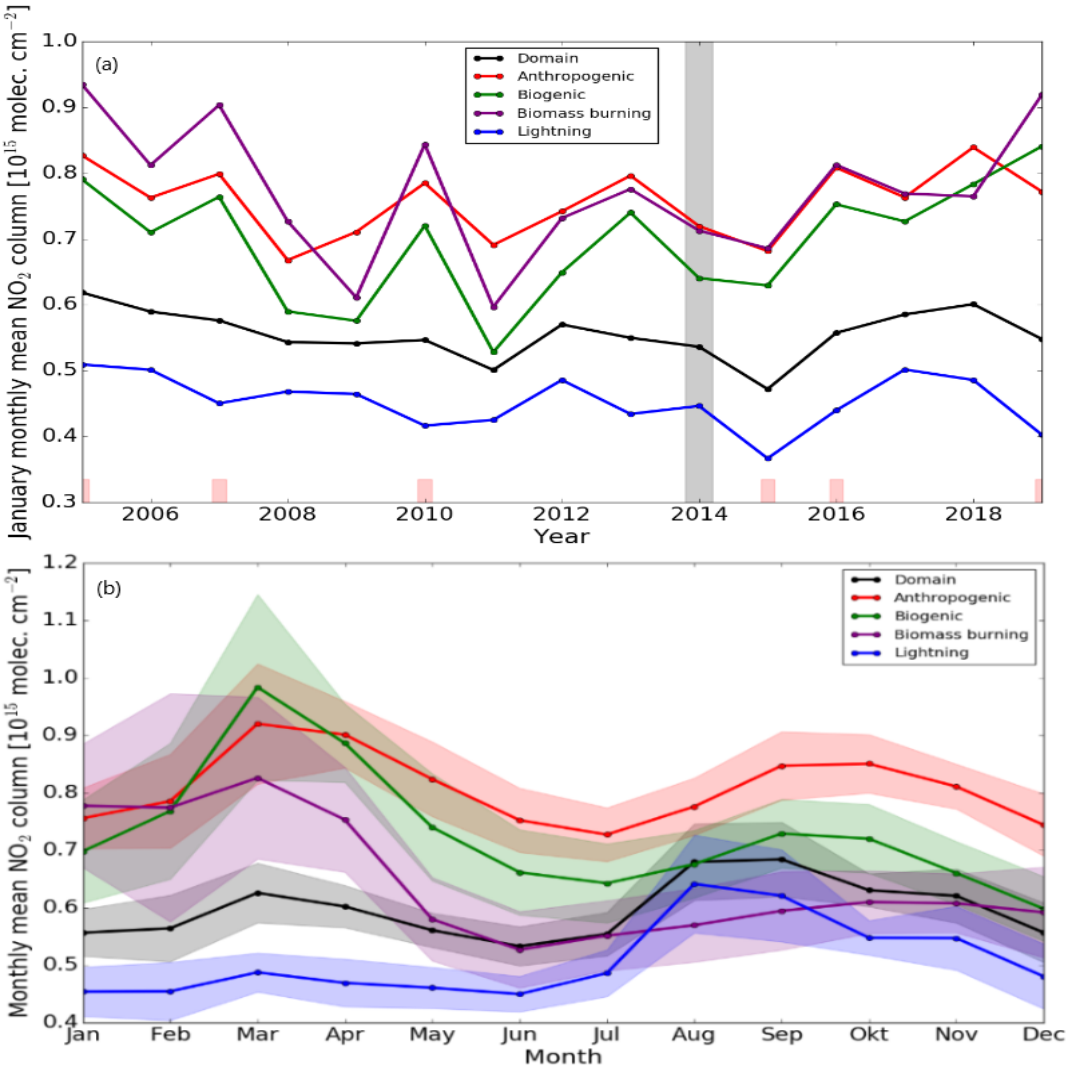

Figure 8. (a) January monthly averaged NO_2 vertical column densities [10^{15} molecules cm^{-2}] retrieved from OMI for 2005-2019 for the whole domain (black), regions with dominating anthropogenic (red), biogenic (green), biomass burning (yellow) and lightning (blue) emissions. The grey vertical bar highlights the WRF-Chem simulated year 2014. The red bars indicate El Niño years (2005, 2007, 2010, 2015, 2016, 2019). (b) Monthly averaged OMI NO_2 vertical column densities [10^{15} molecules cm^{-2}] for 2005-2019. The shadings indicate +/- 1 standard deviation.

The top-down validation approach, using satellite retrievals, is a valuable tool to evaluate air quality in remote regions (Bailey

et al., 2006; Webley et al., 2012) with a missing network of air quality monitoring in both urban and rural sites. The daily global coverage and retrievals of NO_2 by OMI (Levelt et al., 2006) were used to assess the quality of all emission inventories over the whole domain. However, 63% of the data is lost for this specific model setup mainly due to the continuous presence of clouds. Thus, longer simulation times have to be considered in the tropics compared to mid-latitudes. The vertical distribution of NO_x within a modeling environment is key to identify discrepancies for a top-down validation approach using satellite retrievals. It has to be recognized that the satellite sensitivity is reduced towards the surface (Boersma et al., 2016), inducing enhanced differences between observed and modeled profiles. However, this can be overcome by replacing a priori TM5 profiles with those from the applied model (Boersma et al., 2016).

In contrast to the bottom-up validation approach, where WRF-Chem showed a significant underestimation of NO_x compared to the in situ measurements, we found that WRF-Chem does not systematically underestimate urban VCDs. This suggests that the problem is indeed bound to representativeness of WRF-Chem with respect to sub-grid scale emissions and other processes and not so much to the magnitude of anthropogenic emissions. The underestimation by WRF-Chem in the Orinoco region, where biogenic and biomass burning emissions make up a great part of the emission budget, indicate an underestimation of biomass burning emissions. Biogenic emissions are expected to show a more homogeneous distribution over a larger area with less pronounced peak emissions.. Therefore, they are also not expected to explain VCDs over $2{\cdot}10^{15}$ molecules cm^{-2} we found in OMI retrievals. This connects to the findings of Grajales and Baquero-Bernal (2014) who concluded that high VCDs in this region are most likely related to biomass burning, which is apparently underestimated by the emission inventory we applied in this study. Soil NO emissions might be underestimated due to the missing anthropogenic term (fertilizer and manure application) (Visser et al., 2019). However, enhanced soil NO emissions due to the use of fertilizer is estimated to only contribute $\sim$1.3% to the total soil NO flux in the global chemistry-climate model EMAC ($\sim 2.8°$) for this domain (Ganzeveld et al., 2010). Furthermore, the simulated soil biogenic NO flux (187 Gg N yr^{-1}) from MEGAN is in the range between the total soil NO flux (230 Gg N yr^{-1}) and the NO flux at the top of the canopy (105 Gg N yr^{-1}) estimated by EMAC.

We find a large area with overestimations of modeled VCDs in the region dominated by lightning NO_x emissions. These findings are in contrast with Grajales and Baquero-Bernal (2014) who found in their study with the GEOS-Chem modeling system that in remote regions without biomass burning there is an underestimation of modeled VCDs. Our study indicates that lightning NO_x emissions are a major source of NO_x which might explain the discrepancy in the study by Grajales and Baquero-Bernal (2014) in which this source was not considered. Also, the use of WRF-Chem, having a spatial resolution approximately the same size as the OMI observations, can be advantageous over coarser models such as GEOS-Chem used by Grajales and Baquero-Bernal (2014). Further attention is required not only regarding the lightning NO_x parametrization scheme, but also the model representation of convection and clouds, in follow-up studies on atmospheric NO_x over Colombia, or other regions where lightning is a dominant source of NO_x. This study does not aim to provide comprehensive estimates of any of the emission sources using OMI data. Rather, we show the potential use of satellite data in a region with a limited air quality monitoring network in determining the regional scale air quality and NO_x source regions. The use of cloud covered OMI observations to get a more comprehensive estimate of lightning NO_x emissions (Beirle et al., 2010; Pickering et al., 2016) would make a very interesting follow up study.

The air-quality monitoring network in Colombia is limited to four major cities. This implies that the validation is limited to urban areas where anthropogenic emissions are the dominant source of pollution. A comparison with in situ data showed that WRF-Chem systematically underestimates urban surface mixing ratios of NO_x and CO. The surface observations showed a clear signal of morning rush-hour emissions with average observed NO_x mixing ratios up to 90 ppb and single observations not rarely exceeding 150 ppb. Similar to González et al. (2018), who focused on O_3 dynamics in Manizales (medium sized Andean city), we find an overestimation of O_3 by WRF-Chem both during nighttime and daytime. For Manizales, NO_x measurements were not available (González et al., 2018) and were proposed to explain most of the inferred discrepancies between the observed and simulated O_3 mixing ratios. In this study we found that the underestimation of NO by $\sim$10 ppb translates to an overestimation of $\sim$10 ppb O_3. Even though O_3 production and destruction seems to be well captured by WRF-Chem, local emission inventories, including a more detailed spatial resolution around cities, can provide the extra detail needed for sub-grid scale analysis of the interactions between local-scale emissions, chemistry, mixing and resulting pollutant concentrations (González et al., 2018). However, as shown in Sect. 4.3, a nested domain with local, high-resolution emission inventories might not be the main solution to properly simulate urban pollutant concentrations. EDGAR emissions as included in WRF-Chem but then applied in the SCM simulations resulted in averaged diurnal cycles of O_3, CO and NO_x that agreed reasonably well with the observed diurnal cycles. The main difference being that the SCM simulations were especially showing differences regarding the nocturnal inversion compared to WRF-Chem.

One of the regions that is currently undergoing major land-use changes is the Orinoco. Its traditional agriculture and extensive grazing shift rapidly towards a more intensified production of food, biofuels and rubber (Lavelle et al., 2014). Especially oil palm, which is one of the world's most rapidly expanding crops (Fitzherbert et al., 2008), is becoming more and more dominant in the Orinoco region (Vargas et al., 2015). Also, urbanization in Colombia is continuously increasing (Samad et al., 2012). Ongoing and anticipated future transformation of both rural and urban areas, in combination with expected increases in temperature and changes in the hydrological cycle, imply changes in emission budgets affecting air quality in the future. Further consistent coupling of land-use classes with emission representations may provide valuable information of future predicted air quality in Colombia. This includes anthropogenic-, biomass burning-, biogenic-, and lightning emissions apparently all having a generally dominant role in atmospheric NO_x cycling in different regions of Colombia.

6 Conclusions

This study presented an analysis of the baseline state of air quality in Colombia, focusing on NO_x as main metric. Using a highly resolved coupled meteorology-air quality model (WRF-Chem), with roughly the same scale as both global emission inventories as well as satellite retrievals (OMI), allowed us to identify sources of pollution and the baseline state of air quality in Colombia. The main findings illustrate that, within the modeling domain, lightning (1258 Gg N yr^{-1}), anthropogenic (933 Gg N yr^{-1}), soil biogenic (187 Gg N yr^{-1}) and biomass burning emissions (104 Gg N yr^{-1}) all contribute to the total nitrogen emission budget. Especially the spatial distribution, clearly identifying regions with different dominating NO_x sources, shows the importance of providing good estimates of each individual NO_x source.

The top-down validation approach, using OMI retrievals, indicated a mean bias of NO_2 Vertical Column Densities (VCDs) of $0.02 \cdot 10^{15}$ molecules cm^{-2}, which is <5% of the mean column, with a 90% confidence interval of $(-0.43, 0.70) \cdot 10^{15}$ molecules cm^{-2}. The VCDs in the Amazon region are overestimated in WRF-Chem, even after an already strongly reduced production efficiency, with respect to the low cloud free VCDs in OMI which is operating near the detection limit. This is a region where lightning NO_x emissions are the only significant source of NO_2. Additionally, the comparison indicates that GFED biomass burning emissions are potentially underestimated for January 2014 since OMI showed some strong enhancements in NO_2 not being reproduced by WRF-Chem. The biomass burning emission inventory shows some presence of wildfires in that region but the model only produces estimates of VCDs of $\sim 1 \cdot 10^{15}$ molecules cm^{-2}, compared to OMI VCDs up to $2 \cdot 10^{15}$ molecules cm^{-2}, in regions where it is known to have significant biomass burning sources. Air Mass Factors (AMFs) were recalculated based on the vertical distribution of NO_2 within WRF-Chem with respect to the coarse (1°x1°) a priori profiles for a more consistent model satellite comparison. An analysis of the past one and a half decade of OMI NO_2 VCD data showed that the selected simulation period is representative for the baseline state of air quality in Colombia but that interannual and seasonal variability has to be taken into account in interpreting the OMI data and WRF-Chem simulations. The interannual variability in NO_2 columns over the different source regions can be attributed to specific events such as ENSO whereas the seasonal variability shows a strong enhancement of NO_2 VCDs above biogenic and biomass burning regions at the end of the dry season.

The bottom-up validation approach using air quality monitoring stations in urban areas showed that WRF-Chem, at the relative coarse resolution, does not reproduce these observations given the role of large heterogeneity in the emissions and other processes determining pollution levels. Application of the anthropogenic EDGAR emission inventory (0.1°x0.1° resolution) in WRF-Chem resulted in a simulated underestimation of NO_x and CO mixing ratios with respect to the local urban surface measurements. However, WRF-Chem was able to simulate the diurnal amplitude in O_3 reasonably well for all urban locations. It seems that the underestimation of $\sim$10 ppb O_3 both during day- and nighttime can be attributed to the underestimation of NO by $\sim$10 ppb. Additional sensitivity simulations were performed with a Single Column Model (SCM) showed especially a shallower nocturnal inversion layer compared to that simulated in WRF-Chem. This resulted in a better representation of the observed diurnal cycle of urban pollutant mixing ratio without the hypothesized enhancement in emissions. This indicated that besides the use of local emissions inventories in highly resolved modeling systems, it is also essential to carefully assess the role of boundary layer dynamics, in particular the representation of nocturnal mixing conditions, to evaluate simulations of pollutant concentrations.

In this study we presented a concise method, integrating both in situ and remote sensing observations with a mesoscale modeling system, to arrive at a quantification of air quality in regions with a limited measurement network to cover the large spatial heterogeneity in air pollution source distribution. Results obtained in this study provide insight in the baseline state of air quality in Colombia and which is essential to apply the presented combined modeling and measurement approach also to assess how air quality will further change due to future industrialization and land use changes.

Code and data availability. OMI data, in situ data and WRF-Chem output are available upon request as well as scripts to recalculate the tropospheric AMF.

Author contributions. JGMB and LNG designed the experiment. JGMB performed the WRF-Chem simulations. LNG performed the single column model simulations. JGMB performed the analysis and wrote the manuscript, with contributions from all co-authors.

Competing interests. The authors declare no competing interests.

Acknowledgements. ~~The~~ We greatly appreciate the two anonymous reviewers for their critical and constructive comments. Their effort has contributed to a major improvements of this manuscript. Furthermore, the authors acknowledge Oscar Julian Guerrero Molina from the Instituto de Hidrología, Meteorología y Estudios Ambientales (IDEAM) for providing us with air quality monitoring data. We acknowledge WRF-Chem developers and emission inventory (EDGAR, MEGAN, GFED) developpers. We also acknowledge the QA4ECV consortium for making OMI NO_2 data publicly available. We thank Folkert Boersma for his input on the OMI analysis.

Appendix A: ~~Long term analysis of OMI VCD~~

~~(a) January monthly averaged NO_2 vertical column densities 10^{15} molecules cm^{-2}retrieved from OMI for 2005-2019 for the whole domain (black), regions with dominating anthropogenic (red), biogenic (green), biomass burning (yellow) and lightning (blue) emissions. The grey vertical bar highlights the WRF-Chem simulated year 2014. The red bars indicate El Niño years (2005, 2007, 2010, 2015, 2016, 2019). (b) Monthly averaged OMI NO_2 vertical column densities 10^{15} molecules $cm^{-2}$for 2005-2019. The shadings indicate +/- 1 standard deviation.~~

Appendix A: Complete overview in situ data

Appendix B: ~~O_3 and NO_x mixing ratios in Bogotá for January 2014~~

~~Temporal evolution of (a) NO_x and (c) O_3 mixing ratios ppbin Bogotá for WRF-Chem (black solid line) and all available observational stations (coloured points). Scatter plots of the WRF-Chem output compared with averaged (b) NO_x, (d) O_3 mixing ratios ppbfrom the stations are split up in day (yellow) and night (blue). The error bars indicate the standard deviation of the observational data from randomly sampled points (not all standard deviations are shown for visual purposes).~~

Table A1. Available air quality monitoring stations including city, location and measured compounds.

Station name	City	Latitude	Longitude	CO	NO	NO_2	O_3
Pance	Cali	3.305	-76.533				✓
Universidad del Valle	Cali	3.378	-76.534			✓	✓
Compartir	Cali	3.428	-76.467				✓
C. Alto Rendimiento	Bogotá	4.658	-74.084	✓			✓
Carvajal - Sevillana	Bogotá	4.596	-74.149	✓			✓
Fontibon	Bogotá	4.670	-74.142	✓			✓
Kennedy	Bogotá	4.625	-74.161	✓	✓	✓	
Las Ferias	Bogotá	4.691	-74.083	✓			✓
MinAmbiente	Bogotá	4.626	-74.067				✓
Puente Aranda	Bogotá	4.632	-74.118	✓	✓	✓	✓
San Christobal	Bogotá	4.573	-74.084				✓
Tunal	Bogotá	4.576	-74.131	✓	✓	✓	✓
Guaymaral	Bogotá	4.784	-74.044		✓	✓	✓
Suba	Bogotá	4.761	-74.094		✓	✓	✓
Usaquen	Bogotá	4.710	-74.030				✓
CAL-Corp. Lasallista	Medellín	6.102	-75.642				✓
ITA-Casa Justicia	Medellín	6.188	-75.601		✓	✓	
ITA-Col. Concejo	Medellín	6.171	-75.648				✓
MED-Politecnico JIC	Medellín	6.212	-75.581		✓	✓	
MED-Politecnico JIC (S)	Medellín	6.212	-75.581		✓		
BEL-U.S. Buenaventura	Medellín	6.331	-75.569		✓	✓	✓
MED-Museo Antioquia	Medellín	6.253	-75.570	✓			
MED-UN Fac. Minas	Medellín	6.274	-75.593		✓	✓	
MED-UN Nucleo Volador	Medellín	6.266	-75.580		✓	✓	✓
MED-Univ. Medellín	Medellín	6.256	-75.559				✓
MED-Villahermosa	Medellín	6.256	-75.559				✓
BAR-Parque Las Aguas	Medellín	6.409	-75.417				✓
Cabecera	Bucaramanga	7.113	-73.111	✓			
Centro	Bucaramanga	7.119	-73.127	✓	✓	✓	
Ciudadela	Bucaramanga	7.106	-73.124	✓			

References

Archer-Nicholls, S., Lowe, D., Utembe, S., Allan, J., Zaveri, R. A., Fast, J. D., Hodnebrog, Ø., Denier van der Gon, H., and McFiggans, G.: Gaseous chemistry and aerosol mechanism developments for version 3.5. 1 of the online regional model, WRF-Chem, Geoscientific Model Development, 7, 2557–2579, 2014.

Ashmore, M. R. and Marshall, F.: Ozone impacts on agriculture: an issue of global concern, in: Advances in Botanical Research, vol. 29, pp. 31–52, Elsevier, 1998.

Bailey, J. E., Dean, K. G., Dehn, J., Webley, P. W., Power, J., Coombs, M., and Freymueller, J.: Integrated satellite observations of the 2006 eruption of Augustine Volcano, The, pp. 481–506, 2006.

Baklanov, A., Schlünzen, K., Suppan, P., Baldasano, J., Brunner, D., Aksoyoglu, S., Carmichael, G., Douros, J., Flemming, J., Forkel, R., et al.: Online coupled regional meteorology chemistry models in Europe: current status and prospects, Atmospheric Chemistry and Physics, 14, 317–398, 2014.

Bauer, P., Thorpe, A., and Brunet, G.: The quiet revolution of numerical weather prediction, Nature, 525, 47, 2015.

Beirle, S., Huntrieser, H., and Wagner, T.: Direct satellite observation of lightning-produced NOx, Atmospheric Chemistry and Physics, pp. 10 965–10 986, 2010.

Boersma, K., Eskes, H., Veefkind, J., Brinksma, E., Van Der A, R., Sneep, M., Van Den Oord, G., Levelt, P., Stammes, P., Gleason, J., et al.: Near-real time retrieval of tropospheric NO 2 from OMI, Atmospheric Chemistry and Physics, 7, 2103–2118, 2007.

Boersma, K., Eskes, H., Dirksen, R., Veefkind, J., Stammes, P., Huijnen, V., Kleipool, Q., Sneep, M., Claas, J., Leitão, J., et al.: An improved tropospheric NO2 column retrieval algorithm for the Ozone Monitoring Instrument, Atmospheric Measurement Techniques, 4, 1905, 2011.

Boersma, K., Vinken, G., and Eskes, H.: Representativeness errors in comparing chemistry transport and chemistry climate models with satellite UV-Vis tropospheric column retrievals, Geoscientific Model Development, 9, 875, 2016.

Boersma, K. F., Eskes, H. J., Richter, A., Smedt, I. D., Lorente, A., Beirle, S., Van Geffen, J. H., Zara, M., Peters, E., Roozendael, M. V., et al.: Improving algorithms and uncertainty estimates for satellite NO 2 retrievals: results from the quality assurance for the essential climate variables (QA4ECV) project, Atmospheric Measurement Techniques, 11, 6651–6678, 2018.

Bond, D. W., Steiger, S., Zhang, R., Tie, X., and Orville, R. E.: The importance of NOx production by lightning in the tropics, Atmospheric Environment, 36, 1509–1519, 2002.

Bradshaw, J., Davis, D., Grodzinsky, G., Smyth, S., Newell, R., Sandholm, S., and Liu, S.: Observed distributions of nitrogen oxides in the remote free troposphere from the NASA global tropospheric experiment programs, Reviews of Geophysics, 38, 61–116, 2000.

Carslaw, D. C.: Evidence of an increasing NO2/NOx emissions ratio from road traffic emissions, Atmospheric Environment, 39, 4793–4802, 2005.

Chen, F. and Dudhia, J.: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity, Monthly Weather Review, 129, 569–585, 2001.

Córdoba-Machado, S., Palomino-Lemus, R., Gámiz-Fortis, S. R., Castro-Díez, Y., and Esteban-Parra, M. J.: Influence of tropical Pacific SST on seasonal precipitation in Colombia: prediction using El Niño and El Niño Modoki, Climate Dynamics, 44, 1293–1310, 2015.

Dodman, D.: Blaming cities for climate change? An analysis of urban greenhouse gas emissions inventories, Environment and urbanization, 21, 185–201, 2009.

Dudhia, J.: Numerical study of convection observed during the winter monsoon experiment using a mesoscale two-dimensional model, Journal of the Atmospheric Sciences, 46, 3077–3107, 1989.

Fitzherbert, E. B., Struebig, M. J., Morel, A., Danielsen, F., Brühl, C. A., Donald, P. F., and Phalan, B.: How will oil palm expansion affect biodiversity?, Trends in ecology & evolution, 23, 538–545, 2008.

Ganzeveld, L., Lelieveld, J., Dentener, F., Krol, M., Bouwman, A., and Roelofs, G.-J.: Global soil-biogenic NOx emissions and the role of canopy processes, Journal of Geophysical Research: Atmospheres, 107, 2002a.

Ganzeveld, L., Lelieveld, J., Dentener, F., Krol, M., and Roelofs, G.-J.: Atmosphere-biosphere trace gas exchanges simulated with a single-column model, Journal of Geophysical Research: Atmospheres, 107, ACH–8, 2002b.

Ganzeveld, L., Eerdekens, G., Feig, G., Fischer, H., Harder, H., Königstedt, R., Kubistin, D., Martinez, M., Meixner, F., Scheeren, H., et al.: Surface and boundary layer exchanges of volatile organic compounds, nitrogen oxides and ozone during the GABRIEL campaign, Atmospheric Chemistry and Physics, 8, 6223–6243, 2008.

Ganzeveld, L., Bouwman, L., Stehfest, E., van Vuuren, D. P., Eickhout, B., and Lelieveld, J.: Impact of future land use and land cover changes on atmospheric chemistry-climate interactions, Journal of Geophysical Research: Atmospheres, 115, 2010.

Gery, M. W., Whitten, G. Z., Killus, J. P., and Dodge, M. C.: A photochemical kinetics mechanism for urban and regional scale computer modeling, Journal of Geophysical Research: Atmospheres, 94, 12 925–12 956, 1989.

Ghude, S. D., Pfister, G. G., Jena, C., van der A, R., Emmons, L. K., and Kumar, R.: Satellite constraints of nitrogen oxide (NOx) emissions from India based on OMI observations and WRF-Chem simulations, Geophysical Research Letters, 40, 423–428, 2013.

González, C., Ynoue, R., Vara-Vela, A., Rojas, N., and Aristizábal, B.: High-resolution air quality modeling in a medium-sized city in the tropical Andes: Assessment of local and global emissions in understanding ozone and PM 10 dynamics, Atmospheric Pollution Research, 2018.

Grajales, J. F. and Baquero-Bernal, A.: Inference of surface concentrations of nitrogen dioxide (NO 2) in Colombia from tropospheric columns of the ozone measurement instrument (OMI), Atmósfera, 27, 193–214, 2014.

Grell, G. A. and Freitas, S. R.: A scale and aerosol aware stochastic convective parameterization for weather and air quality modeling., Atmospheric Chemistry & Physics Discussions, 13, 2013.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmospheric Environment, 39, 6957–6975, 2005.

Guenther, A., Jiang, X., Heald, C., Sakulyanontvittaya, T., Duhl, T., Emmons, L., and Wang, X.: The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2. 1): an extended and updated framework for modeling biogenic emissions, 2012.

Gupta, M. and Mohan, M.: Validation of WRF/Chem model and sensitivity of chemical mechanisms to ozone simulation over megacity Delhi, Atmospheric Environment, 122, 220–229, 2015.

Hong, S.-Y., Noh, Y., and Dudhia, J.: A new vertical diffusion package with an explicit treatment of entrainment processes, Monthly weather review, 134, 2318–2341, 2006.

Huang, G., Brook, R., Crippa, M., Janssens-Maenhout, G., Schieberle, C., Dore, C., Guizzardi, D., Muntean, M., Schaaf, E., and Friedrich, R.: Speciation of anthropogenic emissions of non-methane volatile organic compounds: a global gridded data set for 1970-2012, Atmospheric Chemistry and Physics, 17, 7683–7683, 2017.

Jaeglé, L., Steinberger, L., Martin, R. V., and Chance, K.: Global partitioning of NO x sources using satellite observations: Relative roles of fossil fuel combustion, biomass burning and soil emissions, Faraday discussions, 130, 407–423, 2005.

Janić, Z. I.: Nonsingular implementation of the Mellor-Yamada level 2.5 scheme in the NCEP Meso model, US Department of Commerce, National Oceanic and Atmospheric Administration, National Weather Service, National Centers for Environmental Prediction, 2001.

Janssens-Maenhout, G., Crippa, M., Guizzardi, D., Muntean, M., Schaaf, E., Dentener, F., Bergamaschi, P., Pagliari, V., Olivier, J., Peters, J., et al.: EDGAR v4.3.2 Global Atlas of the three major Greenhouse Gas Emissions for the period 1970–2012, Earth Syst. Sci. Data Discuss, 2017.

Kim, S.-W., Heckel, A., Frost, G., Richter, A., Gleason, J., Burrows, J., McKeen, S., Hsie, E.-Y., Granier, C., and Trainer, M.: NO2 columns in the western United States observed from space and simulated by a regional chemistry model and their implications for NOx emissions, Journal of Geophysical Research: Atmospheres, 114, 2009.

Krotkov, N. A., Lamsal, L. N., Celarier, E. A., Swartz, W. H., Marchenko, S. V., Bucsela, E. J., Chan, K. L., Wenig, M., and Zara, M.: The version 3 OMI NO 2 standard product, Atmospheric Measurement Techniques, 10, 3133–3149, 2017.

Kumar, A., Jiménez, R., Belalcázar, L. C., and Rojas, N. Y.: Application of WRF-Chem Model to Simulate PM10 Concentration over Bogota, Aerosol and Air Quality Research, 16, 1206–1221, 2016.

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., et al.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmospheric Chemistry and Physics, 10, 7017–7039, 2010.

Lavelle, P., Rodríguez, N., Arguello, O., Bernal, J., Botero, C., Chaparro, P., Gómez, Y., Gutiérrez, A., del Pilar Hurtado, M., Loaiza, S., et al.: Soil ecosystem services and land use in the rapidly changing Orinoco River Basin of Colombia, Agriculture, ecosystems & environment, 185, 106–117, 2014.

Levelt, P. F., van den Oord, G. H., Dobber, M. R., Malkki, A., Visser, H., de Vries, J., Stammes, P., Lundell, J. O., and Saari, H.: The ozone monitoring instrument, IEEE Transactions on geoscience and remote sensing, 44, 1093–1101, 2006.

Lorente, A., Boersma, K. F., Yu, H., Dörner, S., Hilboll, A., Richter, A., Liu, M., Lamsal, L. N., Barkley, M., De Smedt, I., et al.: Structural uncertainty in air mass factor calculation for NO 2 and HCHO satellite retrievals, Atmospheric Measurement Techniques, 10, 759–782, 2017.

Miyazaki, K., Eskes, H., Sudo, K., and Zhang, C.: Global lightning NOx production estimated by an assimilation of multiple satellite data sets, Atmospheric Chemistry and Physics, 14, 3277–3305, 2014.

Mlawer, E. J., Taubman, S. J., Brown, P. D., Iacono, M. J., and Clough, S. A.: Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave, Journal of Geophysical Research: Atmospheres, 102, 16 663–16 682, 1997.

Morrison, H., Thompson, G., and Tatarskii, V.: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one-and two-moment schemes, Monthly Weather Review, 137, 991–1007, 2009.

Murray, L. T.: Lightning NOx and impacts on air quality, Current Pollution Reports, 2, 115–133, 2016.

Ott, L. E., Pickering, K. E., Stenchikov, G. L., Allen, D. J., DeCaria, A. J., Ridley, B., Lin, R.-F., Lang, S., and Tao, W.-K.: Production of lightning NOx and its vertical distribution calculated from three-dimensional cloud-scale chemical transport model simulations, Journal of Geophysical Research: Atmospheres, 115, 2010.

Panella, M., Tommasini, V., Binotti, M., Palin, L., and Bona, G.: Monitoring nitrogen dioxide and its effects on asthmatic patients: Two different strategies compared, Environmental monitoring and assessment, 63, 447–458, 2000.

Pickering, K. E., Bucsela, E., Allen, D., Ring, A., Holzworth, R., and Krotkov, N.: Estimates of lightning NOx production based on OMI NO2 observations over the Gulf of Mexico, Journal of Geophysical Research: Atmospheres, 121, 8668–8691, 2016.

Price, C. and Rind, D.: A simple lightning parameterization for calculating global lightning distributions, Journal of Geophysical Research: Atmospheres, 97, 9919–9933, 1992.

Randerson, J., van der Werf, G., Giglio, L., Collatz, G., and Kasibhatla, P.: Global Fire Emissions Database, Version 4,(GFEDv4), ORNL DAAC, Oak Ridge, Tennessee, USA, 2015.

Saide, P., Spak, S., Carmichael, G., Mena-Carrasco, M., Yang, Q., Howell, S., Leon, D., Snider, J. R., Bandy, A. R., Collett, J. L., et al.: Evaluating WRF-Chem aerosol indirect effects in Southeast Pacific marine stratocumulus during VOCALS-REx, Atmospheric Chemistry and Physics, 12, 3045, 2012.

Samad, T., Lozano-Gracia, N., and Panman, A.: Colombia urbanization review: amplifying the gains from the urban transition, World Bank Publications, 2012.

Silvern, R., Jacob, D., Travis, K., Sherwen, T., Evans, M., Cohen, R., Laughner, J., Hall, S., Ullmann, K., Crounse, J., et al.: Observed NO/NO2 Ratios in the Upper Troposphere Imply Errors in NO-NO2-O3 Cycling Kinetics or an Unaccounted NOx Reservoir, Geophysical Research Letters, 45, 4466–4474, 2018.

Streets, D. G., Bond, T., Carmichael, G., Fernandes, S., Fu, Q., He, D., Klimont, Z., Nelson, S., Tsai, N., Wang, M. Q., et al.: An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, Journal of Geophysical Research: Atmospheres, 108, 2003.

Tie, X., Madronich, S., Walters, S., Zhang, R., Rasch, P., and Collins, W.: Effect of clouds on photolysis and oxidants in the troposphere, Journal of Geophysical Research: Atmospheres, 108, 2003.

Vargas, L. E. P., Laurance, W. F., Clements, G. R., and Edwards, W.: The impacts of oil palm agriculture on Colombia's biodiversity: what we know and still need to know, Tropical Conservation Science, 8, 828–845, 2015.

Vinken, G., Boersma, K., Maasakkers, J., Adon, M., and Martin, R.: Worldwide biogenic soil NO x emissions inferred from OMI NO 2 observations, Atmospheric Chemistry and Physics, 14, 10 363–10 381, 2014.

Visser, A. J., Boersma, K. F., Ganzeveld, L. N., and Krol, M. C.: European NO x emissions in WRF-Chem derived from OMI: impacts on summertime surface ozone, Atmospheric Chemistry and Physics, 19, 11 821–11 841, 2019.

Wang, Q. and Li, R.: Sino-Venezuelan oil-for-loan deal–the Chinese strategic gamble?, Renewable and Sustainable Energy Reviews, 64, 817–822, 2016.

Webley, P., Steensen, T., Stuefer, M., Grell, G., Freitas, S., and Pavolonis, M.: Analyzing the Eyjafjallajökull 2010 eruption using satellite remote sensing, lidar and WRF-Chem dispersion and tracking model, Journal of Geophysical Research: Atmospheres, 117, 2012.

WHO: Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide: report on a WHO working group, Bonn, Germany, 2003.

Williams, J. E., Boersma, K. F., Sager, P. L., and Verstraeten, W. W.: The high-resolution version of TM5-MP for optimized satellite retrievals: description and validation, Geoscientific Model Development, 10, 721–750, 2017.

Wolfe, A. H. and Patz, J. A.: Reactive nitrogen and human health: acute and long-term implications, Ambio: A journal of the human environment, 31, 120–125, 2002.

Zárate, E., Belalcazar, L. C., Clappier, A., Manzi, V., and Van den Bergh, H.: Air quality modelling over Bogota, Colombia: Combined techniques to estimate and evaluate emission inventories, Atmospheric Environment, 41, 6302–6318, 2007.

Zaveri, R. A. and Peters, L. K.: A new lumped structure photochemical mechanism for large-scale applications, Journal of Geophysical Research: Atmospheres, 104, 30 387–30 415, 1999.