communications earth & environment

Article

https://doi.org/10.1038/s43247-024-01424-5

Accounting for NO_x emissions from biomass burning and urbanization doubles existing inventories over South, Southeast and East Asia

Check for updates

Jian Liu [1,2], Jason Blake Cohen [2] ✉, Qin He [2], Pravash Tiwari [2] & Kai Qin [2]

Rapid urbanization and broad use of biomass burning have led to important changes in NO_x [sum of nitrogen dioxide and nitrous oxide] emissions across South, Southeast, and East Asia, frequently occurring on day-to-day time scales and over areas not identified by existing emissions databases. Here we compute NO_x emissions using remotely sensed NO_2 [nitrogen dioxide] and a model-free mass-conserving inverse method, resulting respectively in 61 kt d^{-1} and 40 kt d^{-1} from biomass burning in Northern and Southern Continental Southeast Asia, and 14.3 kt d^{-1} and 3.7 kt d^{-1} from urbanization in China and Eastern South Asia, a net increase more than double existing inventories. Three observationally based physical constraints consistent with theory are found which current chemical transport models cannot match: more NO_2 per unit of NO_x emissions, longer and more variable in-situ lifetime, and longer-range transport. This result provides quantitative support for mitigation efforts targeting specific events, processes, or geographies.

Nitrogen dioxide (NO_2) is a reactive, short-lived atmospheric trace gas which is formed both naturally from lightning and emitted by anthropogenic activities generating heat in the atmosphere such as fossil fuel use and biomass burning (BB)[1,2]. NO_2 reacts in the atmosphere with OH and clouds to form nitrate aerosol, an important fraction of $PM_{2.5,}$ as well as an important agent in coating and removing black carbon (BC) from the atmosphere[3]. NO_2 also has a very short lifetime in the atmosphere, rapidly exchanging with NO as defined by the pseudo steady-state assumption[4]. Therefore, in general works look at the total sum of NO and NO_2, hereafter referred to as NO_x. Both the ratio of NO_2 to NO as well as the total loading of NO_x, in combination with heat, ultraviolet radiation, and either carbon monoxide (CO) or volatile organic compounds (VOCs) leads to the production of ozone (O_3)[5]. Therefore, determining the emissions, and in-situ processes impacting both NO_2 and NO_x are essential for understanding their atmospheric distribution as well as their impact on multiple atmospheric, environmental, and climatological phenomena[6,7].

The column loading of NO_2 can be readily measured by analyzing backscattered blue sunlight in combination with differential optical absorption spectroscopy on the space-based downward looking platform OMI[8]. The patterns between these absorption spectra and atmospheric column loadings of NO_2 in-situ can be evaluated quantitatively within a degree of uncertainty[9]. Some recent works have investigated the relationships from OMI of NO_2 directly, with only a very small few venturing to other forms of nitrogen, including a study connecting column NO_2 and surface NO_x using big data, and a study connecting NO_2 and ammonia with a model[10,11].

There are many emission inventories currently used by the scientific community that provide emissions of NO_x and other co-emitted species such as CO and BC. The first way in which such data sets are derived is using a bottom-up approach, an approach that relies on aggregating economic, population, and other factors, together with emissions factors obtained from idealized cases or events, not the atmosphere in-situ, as is done by EDGAR[12,13]. Another approach is to scale bottom-up emissions datasets in tandem with remotely sensed measurements of land use change or fire radiative power at higher temporal and spatial frequency, and then couple these with emissions and mass factors based on small-scale studies which are built from the bottom-up, as is done by FINN[14]. These approaches rely upon a limited number of measurements from the laboratory or limited field experiment[15,16], strongly constraining their results based on a prior information. Therefore, these emissions datasets do not adapt well to new or

[1]School of Atmospheric Sciences, Sun Yat-Sen University, Zhuhai, China. [2]School of Environment and Spatial Informatics, China University of Mining and Technology, Xuzhou, China. ✉e-mail: jasonbc@alum.mit.edu; jasonbc@cumt.edu.cn

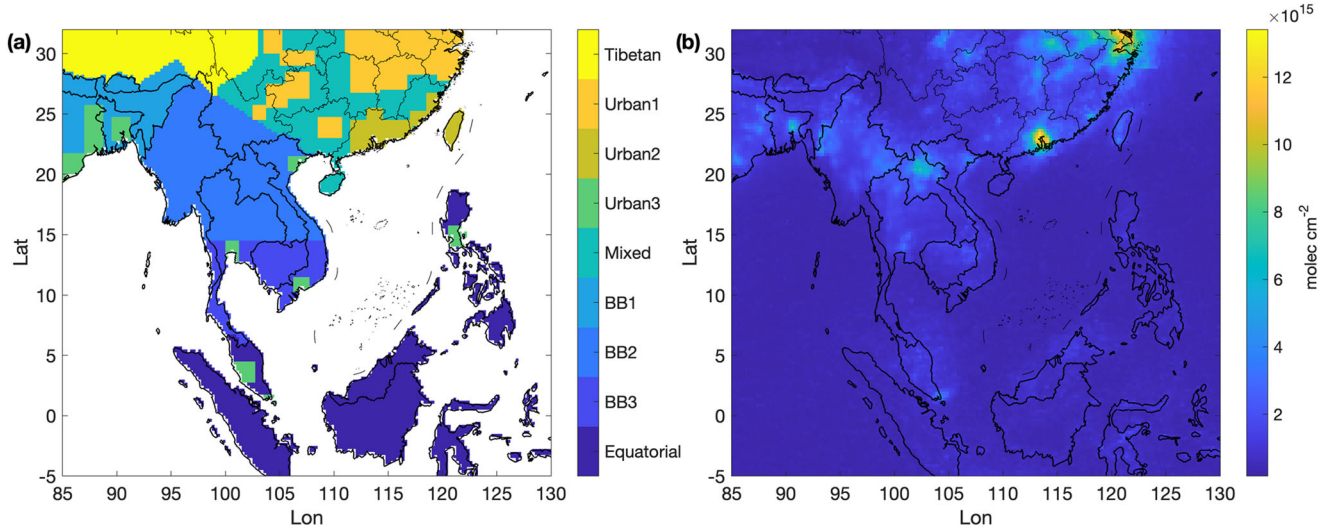

Fig. 1 | An overview of land-use categories and NO_2 observations. Spatial distribution of (**a**) different land-use categories and (**b**) climatological standard deviation of daily remotely sensed column NO_2 [molecules cm^{-2}].

rapidly changing sources, or when the underlying factors impacting emissions change[17,18]. Furthermore, rapid changes frequently impact the quality or availability of measurements themselves, including changes in albedo, land use, rain and clouds, and high loadings of aerosols from upwind fires, leading to hard-to-scale uncertainties[19,20].

Some works have scaled bottom-up emissions changes directly to column loadings from satellites, using a mixture from simple scaling to complex models, and data assimilation techniques, although in general these are not able to improve datasets in which the a priori is zero[21,22]. In these cases, the top-down results tend to be biased by the spatial and temporal assumptions underlying the a priori emissions and are strongly impacted by mis-characterization of the measurement uncertainty[23–25]. The applications of direct top-down estimation approaches are generally only applied to long-lived gasses (CH_4, CFCs, and N_2O), since their chemical decay is very slow compared with their transport processes, allowing for any observed perturbations that are larger than the measurement uncertainty to definitely be due to an emissions source, which then can be inverted based on an inversion of the meteorological fields[26]. There is only a limited set of papers addressing short-lived species, and always under idealized conditions including where there is a strong single point source surrounded by what is otherwise a source-free region[27,28] or using an underlying model to interactively approximate the chemical and transport properties of the short-lived species on average over a long period of time and then using these average conditions to make an average inversion over a single geographic region under temporal conditions which are climatologically similar[29].

Analyzing spatial and temporal extremes of remotely sensed measurements and attributing the variability in connection with a rigorous analysis of measurement and physical process error has also been used to demonstrate the spatial and temporal extent of missing emissions, albeit under limited conditions[30,31]. The frequency and spatial distribution of time-varying aerosol sources in Asia was identified using observations from MISR[32] and validated by independent measurements from CALIOP, OMI, and MOPITT[21,33]. A top-down Kalman Filter approach was utilized to estimate global-scale BC emissions[28,34]. Remotely sensed measurements of CO from MOPITT were used to constrain the temporal and spatial distribution of missing emissions sources, as well as a pseudo-magnitude[33]. An attempt at improving the representation of missing biomass burning and urban pollution sources at the daily scale was attempted using mathematical techniques of intermediate complexity using both empirical orthogonal functions and plume rise modeling, to determine the variability and vertical distribution of observed extremes of CO and NO_2 over a decade-long period[20,27]. There have also been attempts to constrain the chemical lifetime by making simple assumptions such as gaussian plumes or diffusion-based transport[35–37]. What none of these has been able to accomplish before is a direct calculation of emissions, free from a priori, under highly complex atmospheric, surface, and anthropogenic forcing conditions, with a focus on day-to-day extreme events.

This work utilizes a unique, fast, first-order approach using daily measurements of remotely sensed NO_2, winds, and mass-conserving estimates of in-situ chemical and physical processing to estimate the daily NO_x emissions of extreme events associated with both biomass burning and rapid changes in urbanization. This paper adds five aspects: first, the process quantifies the sources associated with extreme events on a day-to-day basis (with groups of extreme events ranging from 15 to 50 days); second, the approach uses measurements over heterogeneous geographical regions that include mountains, coasts, urban regions, forests, agricultural land, and various atmospheric phenomena including high cloud cover, high aerosol cover, large temperature gradients, extremes in surface UV radiation, a substantial contribution of both land and sea, and considerable vertical atmospheric mixing; third, the approach is able to capture sizable NO_x sources associated with both biomass burning and urbanization, ranging from hyper modern Singapore to least developed countries like Laos; fourth, the approach does not rely on any chemical or climate transport or climate models; and fifth, the approach is not limited to a single season or relatively consistent geographic region. Furthermore, by analyzing the first-order approximations of thermodynamical, chemical, and transport factors, this work presents a holistic and unequivocal approach to quantify both known and missing sources, as well as their uncertainty.

Results

Land surface characteristics over the domain from 5°S to 32°N and 85°E to 130°E, covering most of Continental Southeast Asia (CSA), large parts of Southern and Central China (SCC), Northeast India (NEI), and Bangladesh are given in (Fig. 1). This region is geographically heterogeneous with elevations ranging from over 8000 m to the surface, climate types ranging from third pole to tropical, and vegetation types ranging from tropical rainforest through savanna, as well as substantial ocean cover. Fig. 1 shows the climatological standard deviation of 2016 daily OMI NO_2 observations, which has been previously shown to indicate changes in source amount and variability[18,38]. The highest values are found in existing urban areas including conurbations (Shanghai to Nanjing, Hefei, and Hangzhou [YRD] and Hong Kong to Shenzhen and Guangzhou [PRD]), large urban basins of SCC, Dhaka, Bangkok, Hanoi, Ho Chi Minh City, and Singapore, as well as in rural regions in the mountains of NEI and CSA. This indicates there are

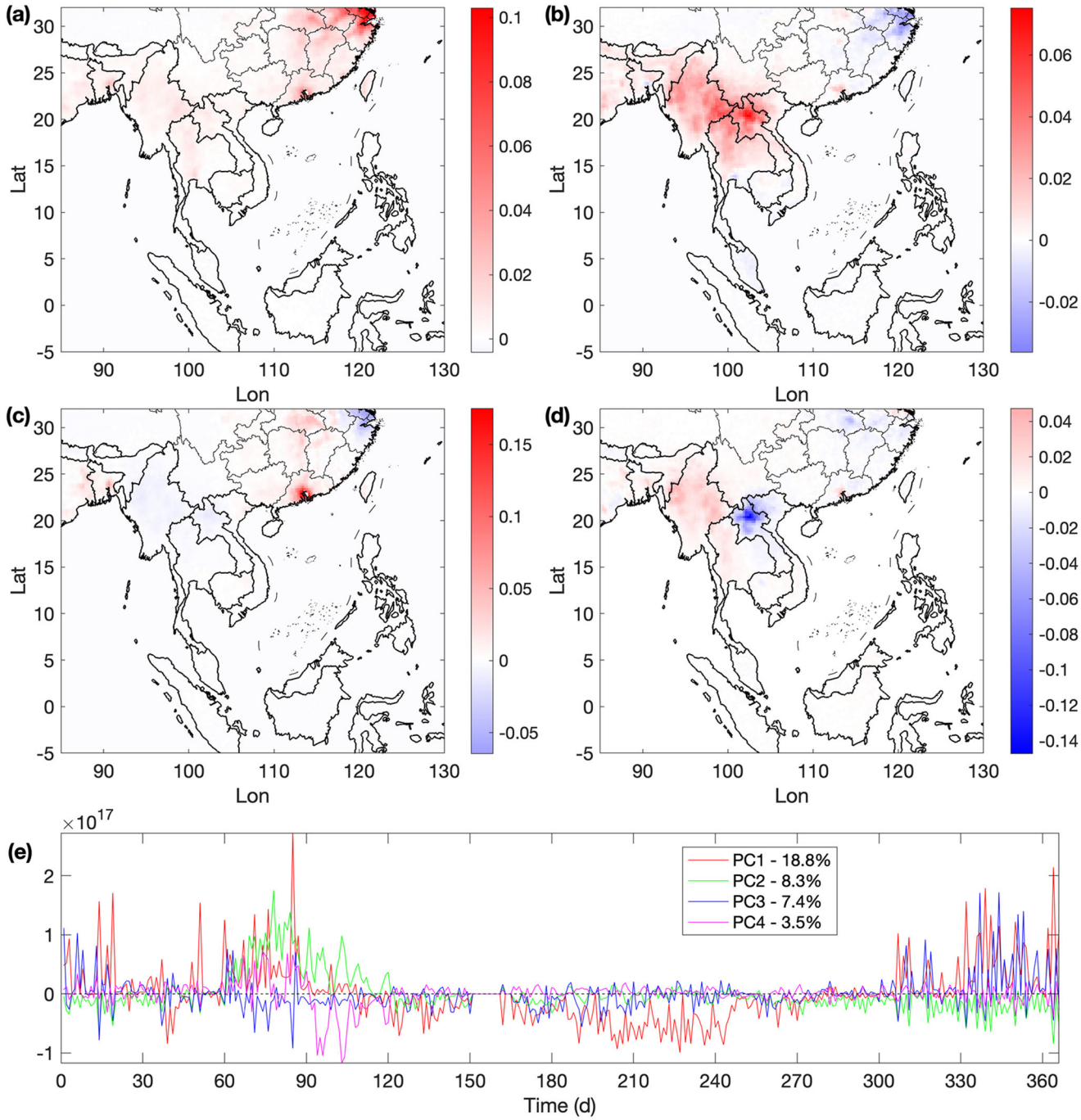

Fig. 2 | Spatial and temporal features extracted from NO_2 observations via variance maximization method. a–d Spatial distributions of the four main EOFs and (e) the corresponding time series of the PCs.

substantial changes in NO_2 occurring over short periods over areas which are usually polluted as well as over areas which are usually clean, indicating that both urbanization and biomass burning are sizable sources[32,39].

Spatial and temporal features of extreme conditions

A variance maximization technique has been applied to the daily NO_2 column measurements over the entire domain, which has derived orthogonal spatial and temporal patterns contributing to the maximum amount of variation in the field. Emissions and changes in chemistry, transport, and thermodynamics are strongly related to the change in variance, with the stronger signals being easier to decompose and analyze[31,32]. The four spatial [EOF] and temporal [PC] modes contributing the most variation of the observed daily NO_2 fields contribute 18.8%, 8.3%, 7.4%, and 3.5% of the total variability with the spatial patterns given by EOFs in Fig. 2 and the temporal patterns given by PCs in Fig. 2.

The first EOF demonstrates a signal in the urban areas in the PRD, YRD, and SCC, with additional moderate signals in Xiamen, Taizhong, Dhaka, Bangkok, and cities in Eastern India. Additionally, it demonstrates contribution from remote areas in Myanmar, Northern Thailand, and Northwestern Laos (Supplementary Fig. 1)[40]. The first PC is observed from days 45 to 90, and again from days 300 to 366, with a few individual peaks observed between day 1 and day 30.

The second EOF shows a signal across the biomass burning areas in CSA including Myanmar, Northern Laos, and Northern Thailand, and extending into the remote regions surrounding the PRD, Yunnan, NEI and Northern Bangladesh. There are smaller signals observed in the urban YRD,

Table 1 | NO_x Emission inventory source used in different kinds of regions

Area categories	Main emission source	Emission inventory	Covering regions
Urban1 area	Anthropogenic	EDGAR	Cities in Pearl River Delta and Fujian and Taiwan Province
Urban2 area	Anthropogenic	EDGAR	Cities located in Yangtze River Basin, including Shanghai, Wuhan, Chengdu and etc.
Urban3 area	Anthropogenic	EDGAR	Dhaka and urban areas in CSA
BB1 area	Biomass burning	FINN	BB areas in NEI
BB2 area	Biomass burning	FINN	BB areas in CSA
BB3 area	Biomass burning	FINN	BB areas in Central Thailand and Cambodia
Mixed area	Anthropogenic, biomass burning	EDGAR + FINN	The areas except BB and urban areas in China
Tibetan area	Biomass burning	FINN	Tibetan plateau and its surrounding mountain areas
Equatorial area	Biomass burning	FINN	Malaysia, Indonesia, and Philippines without main cities

parts of Southeastern China, the Malaysian Peninsula, Dhaka and Bangkok. The second PC is concentrated from days 60 to 120, and again from days 300 to 366.

The third and fourth EOF are weaker, although still have some regions with notable signal. The third EOF displays an urban signal in the PRD, YRD, Central China, NEI, and Bangladesh, while also showing a weak signal in the biomass burning areas in CSA. The third PC has sufficient magnitude, but only occurs over isolated peaks banded from days 1 to 90 and again from days 300 to 366. The fourth EOF outlines biomass burning areas west of 102°E, including Myanmar, Western Thailand, and some parts of NEI, as well as Laos. The fourth PC has only a few days' worth of peaks, all occurring from days 60 to 120.

The geographic and temporal features of the EOFs and PCs allow the map to be divided into urban and non-urban (including clean, biomass burning, and mixed) regions as demonstrated in Fig. 1 and Table 1, following variance maximization[27,32]. For non-urban areas, this work focuses on the mountainous tropical forests, agricultural lands, and densely populated non-urban areas found throughout CSA and NEI, regions with an annually occurring intense wet season and extended dry season during which time most fires will occur or expand[41,42].

The specific days and locations are chosen based on the variance maximization analysis of the underlying OMI NO_2 column measurements, so that they cover the spatial-temporal extent of the largest changes in emissions or in-situ processing of NO_x as demonstrated in Fig. 2.

The time series of the NO_2 weighted column loading over the portions of EOF1 and EOF3 corresponding to BB2 and the portions of EOF1 corresponding to Urban1 and Urban2 (see Fig. 1) are displayed with the respective spatial and temporal cutoffs in Fig. 3. The signal within central Myanmar, Northern Thailand and Northern Laos occurs a total of 30 days between days 60 and 110, while the signal throughout most of Myanmar and Laos occurs a total of 39 days from day 60 to day 120, indicating multiple stable phases of biomass burning occurring in different geographic areas at different times, with there being considerable overlap in Eastern Myanmar and Northern Laos, but not in other regions. This indicates clearly that there are distinct anthropogenic and natural driving forces behind the burning. Similarly, in the urban areas, the NO_2 weighted column loading over the PRD and SCC occur over 36 days at random from days 1 to 110 and from day 300 to day 360; over the PRD and parts of Fujian and Taiwan Island the signal occurs over 25 days at random from days 300 to 366 and another 11 days completely at random throughout the remainder of the year.

The observed weighted NO_2 in each BB and urban EOF/PC region at the selected times (in green) is larger than the weighted NO_2 over the entire region averaged over the year, with the weighted NO_2 enhanced by a factor of 8.7 and 9.5 over EOF1 BB2 and EOF3 BB2 respectively, and enhanced by a factor of 18.5 and 13.5 over EOF1 Urban1 and EOF1 Urban2 respectively. These results consistently and uniquely identify those regions on a day-to-day basis that are highly polluted.

Thermodynamic, chemical, and transport coefficients

Using the filtered spatial and temporal features, the best fit coefficients for thermodynamics of NO and NO_2 production (α_1), chemical loss of NO_x due to in-situ chemical reactions (α_2), and transport of NO_x (α_3) are computed following Eq. (5) (see Supplementary Table 1). α_1 is found to range from 1.12 to 17.1 [NO_x/NO_2], α_2 is found to range from 4.3 to 23.8 [hours], and α_3 is found to range from −19.2 to 11.9 [number of 0.25°x0.25° grids]. α_3 is more frequently negative (54.8% of the grids) consistent with the net flow from higher emissions grids into lower emissions grids, while the reverse is observed in 45.2% of grids, consistent with transport by pressure gradients also playing an essential role. This result is a step beyond current approaches of estimating emissions of NO_x, CH_4, and CO_2 from satellite[25,36,43], which limit themselves exclusively to consider flow only from high concentration to low concentration regions.

The ratio of NO_x/NO_2 in BB areas (1.18 to 17.1) is relatively larger than in urban areas (1.12 to 4.41) indicating relatively more production of NO in BB areas, consistent with the combustion temperature and efficiency in BB areas being lower than combustion found in urban areas (i.e., from power plants, factories, residential use, and transport). In particular, difference between the 25% and 75% value of α_1 over BB areas is broader than over urban areas, consistent with a more diverse set of fuel loadings, land-use types, temperatures and types of combustion (burning, flaming, smoldering, etc.), and energy efficiencies based on how biomass is consumed (fuel, heat, light, etc.), consistent with known thermodynamics associated with emission temperature[44]. While local atmospheric processing will also have an impact on this ratio, given that there is no substantial difference in the climatology between these regions, and the fact that both regions under high pollution conditions frequently have very low to nearly zero ozone concentrations, the net effect of the climate processing is not expected to be vastly different between these regions.

Chemical lifetime in general is slower over BB regions than urban regions. Most of the 25% percentile α_2 values over BB areas are more negative than over urban areas, consistent with the higher average OH levels and more rapid oxidization on average in urban areas[45]. This is further consistent with the fact that ultra-fast non-linear oxidation at the sub-kilometer level will not be represented at grid-scale[46] and in this work is assumed to be a function of the emissions following previous studies[47,48]. One special case is PC2-BB3-Pos, which has a chemical lifetime similar to urban areas, in part because its biomass burning emissions are adjacent to and interact with the urban chemistry associated with Bangkok, Ho Chi Minh City, and rapidly urbanizing regions around Phnom Penh and Tonle Sap (Supplementary Fig. 1)[40]. A second special case is found in Urban3, which has a wider range (particularly on the faster lifetime side) than Urban2 and Urban1, driven both by substantial local biomass burning sources within the urban airsheds and more intense UV radiation on average to drive photochemistry[40,49].

The transport term α_3 is based on the atmospheric concentration, wind direction, and orientation between adjacent areas, with values close to zero

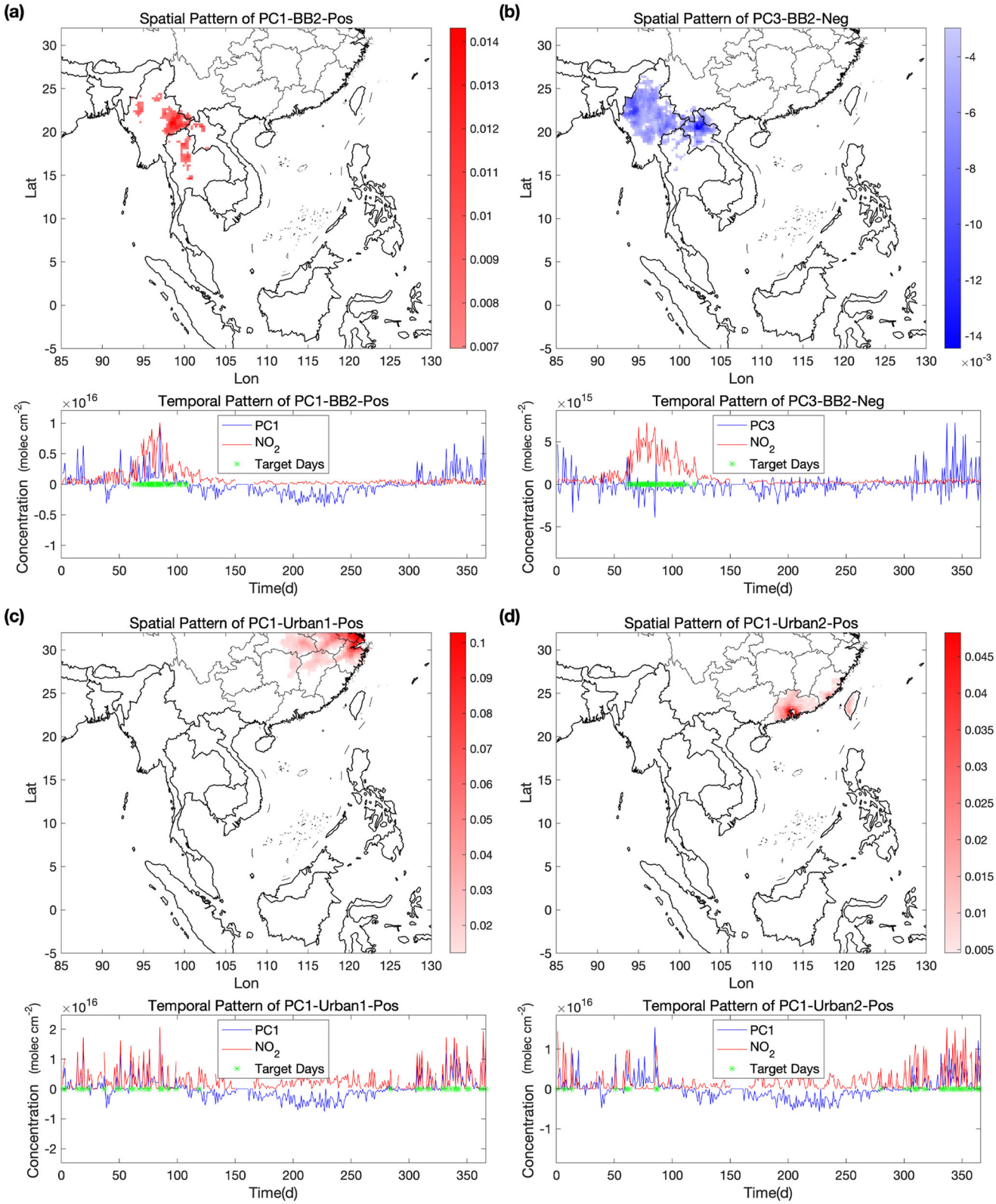

Fig. 3 | Spatial and temporal patterns of typical extreme. Different EOF distributions and PC time series aggregated over (**a**, **b**) BB2 areas, and (**c**, **d**) Urban1 and Urban2 areas.

being inconsequential. First, the transport distances from BB areas are smaller than from urban areas, consistent with both the faster chemical decay in BB regions and the lower surface roughness in urban areas. Second, the magnitude of the outward transport tends to be larger than inward transport, consistent with greater chemical loss occurring when polluted plumes are transported into more polluted regions. There are 4 special regions, three of which are always exporting regions (PC2-BB3-Pos always exports to surrounding urban areas, and PC2-Urban1-Neg and PC1-Urban3-Pos always export to surrounding suburban areas), and one of which is always an importing region (PC1-Urban2-Pos is surrounded by the

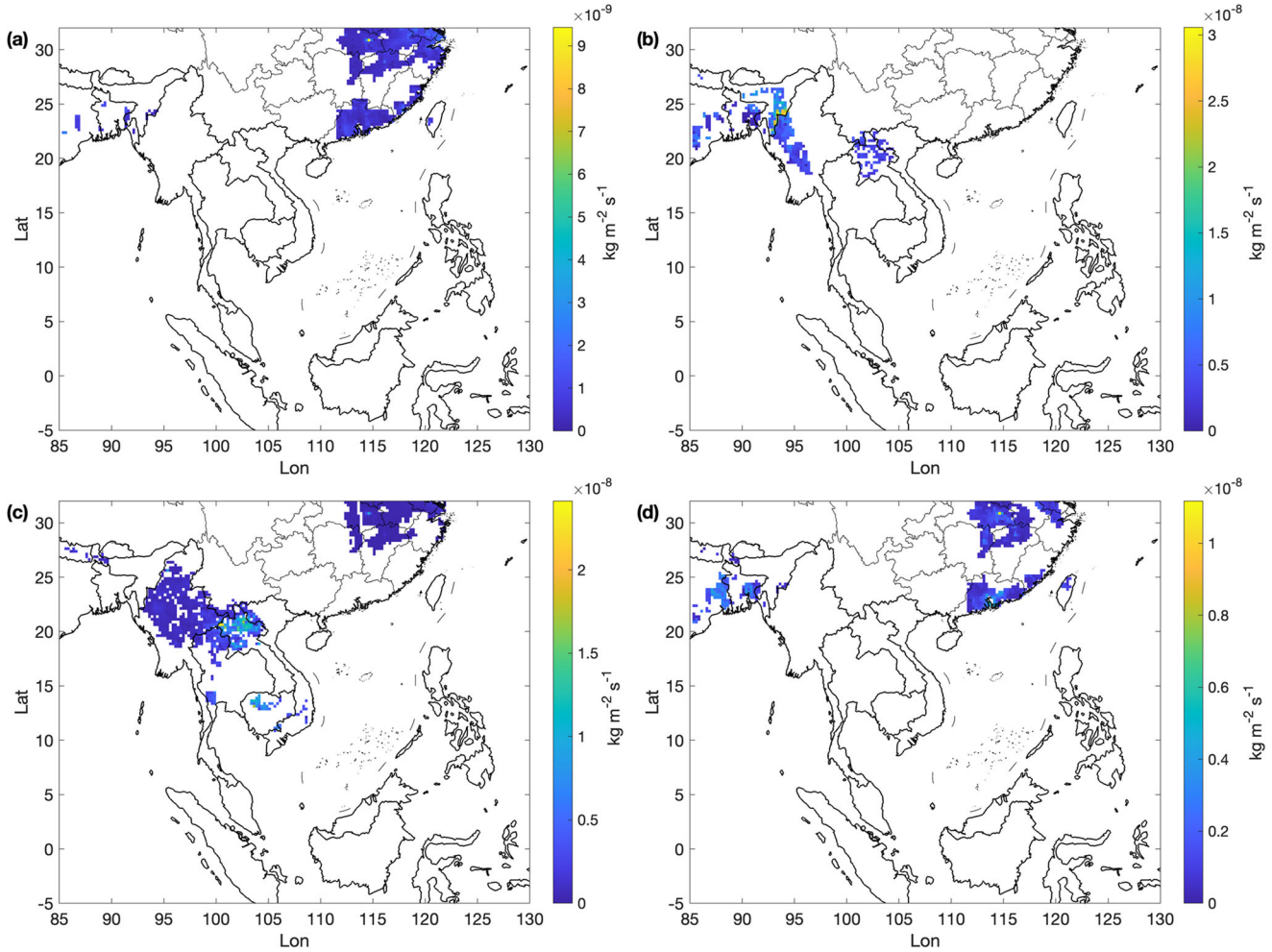

Fig. 4 | NO_x Emissions of four typical days. Daily NO_x Emissions computed on four distinctly characteristic days (**a**–**d**): Days 51, 72, 119, and 362.

even more intensely emitting urban and manufacturing centers of Guangzhou, Dongguan, Shenzhen, Xiamen, Xinbei and Taibei, as well as being downwind from biomass burning exported from NEI and the Northern CSA). The special case of BB3 is always negative as its emissions are advected away from the mountains on the West and the East into the central urban areas exclusively on days 1 to 90 with PC1-Urban3-Pos, which consequently is the same time as the high-speed west-to-east wind induced by the transition in the Asian Monsoon[50]. The special case of PC3-Urban3-Neg demonstrates a mostly positive environment which occurs infrequently from days 30 to 90 when the biomass burning emissions are high, but otherwise randomly throughout the year consistent with a blocking type of meteorological environment near the surface[39,41]. Consequentially, some of the regions with the greatest amount of import and export may been dependent on meteorological and climate factors driven far from the source regions[51].

NO_x emissions

The daily emissions computed on days 51 and 362 have large urban sources of NO_x emissions (Fig. 4). It is clearly demonstrated that the geospatial distribution of emissions captures the largest urban areas including a cluster in SCC and YRD, a second cluster in the PRD, Fujian and Taiwan Island, and a third cluster around NEI and Bangladesh. By comparison, the daily emissions computed on days 72 and 119 represent high emissions from BB regions (Fig. 4). The geospatial distribution of emissions clustered in Northern CSA, Southern CSA, and NEI. These results demonstrate both large sources of NO_x emissions from Myanmar, Laos and NEI as well as urban regions in the outer parts of the YRD and SCC, although the a priori inventories are very low or non-existent on these days. The ability to identify and quantify emissions occurring at either the wrong time and/or the wrong place, or sources which are heavily impacted by changes in climate or other dynamical effects, is essential for attribution work, since areas with zero emissions cannot be scaled or otherwise easily accounted for[48,52]. This finding is clearly demonstrated in the case of the Hong Kong Special Administrative Region's inverted NO_x emissions, which matches the well-known west to east high to low emissions gradient (Supplementary Fig. 2), as well as general magnitude (15.3 ± 3.5 kt yr^{-1} over a total of 48 days (325 ± 75 t d^{-1}), as compared to the Hong Kong Government's value of 89.7 kt yr^{-1} (249 t d^{-1})), even though major emissions data sets (including EDGAR and MEIC) do not have emissions data for Hong Kong in 2016.

The annual mean and normalized standard deviation of the extreme event emissions calculated in this work and the merged EDGAR and FINN emissions are given in Fig. 5. The net a priori emissions show a wider geospatial coverage due to the 22.3% of the domain identified herein that does not have an OMI column NO_2 signal which changes enough in at least one of either space or time to be identified using this approach. Over the regions identified, the a priori emissions on average are a factor of 7.5 times lower than the mean, with considerable variability occurring in different regions. In the rapidly developing urban areas of Dhaka, Bangkok, Ho Chi Minh, Manila, Kuala Lumpur, and Hanoi, the emissions are 11 times higher than the a priori, consistent with rapid increases in energy use and economic development and biomass burning in upwind areas. The mean emissions in the PRD is a factor of 4.2 higher than the a priori, consistent with the rapid growth in most of the region outside of the core areas of Guangzhou, Foshan, Dongguan, Hong Kong, and Shenzhen, as observed in both the

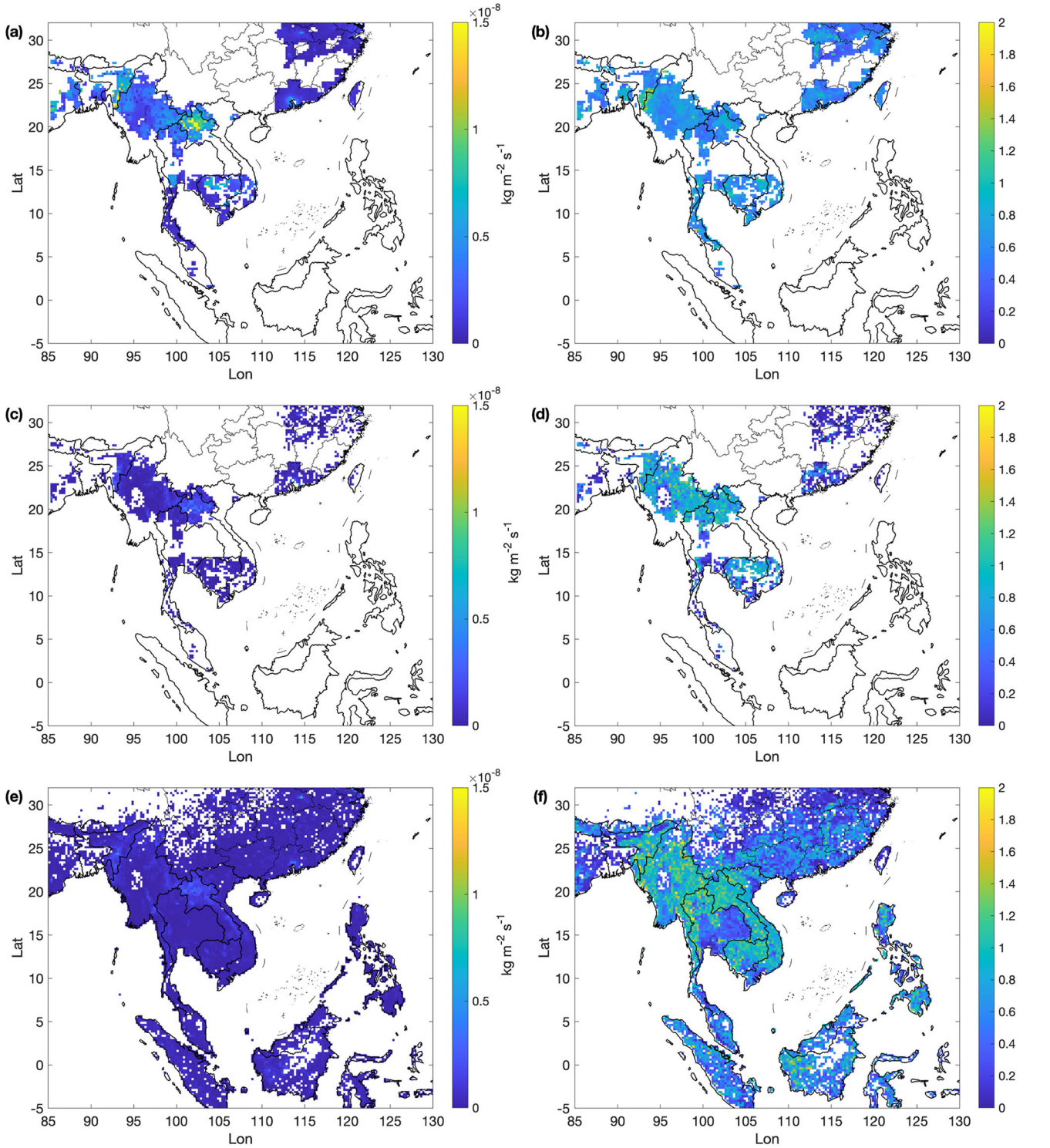

Fig. 5 | Statistics of daily emissions over the EOF regions defined during different specific days. a, **b** are days defined by the PCs (extreme conditions), (**c**, **d**) are days defined by FINN (biomass burning), and (**e**, **f**) are days given by EDGAR (entire year). **a**, **c**, **e** represent the mean [kg m^{-2} s^{-1}] and **b**, **d**, **f** the ratio of standard deviation to mean respectively of NO_x emissions.

amount and geospatial distribution of factories and transport sources as well as a broadening of the residential density (both a magnitude and an extension of geographic extent of emissions). This finding is consistent with the ongoing changes to rapidly increase the development of the PRD into an integrated economic region inclusive of Hong Kong and Macau[53]. The mean emissions in the biomass burning areas of Myanmar, Northern Laos, and Northern Thailand is higher than the a priori emission by a factor of 2.5, consistent with increased forest fires. Inland urbanizing regions in China, including Wuhan, Changsha and Nanchang, have mean emissions 2.1 times higher than the a priori, consistent with the overall push by China to increase economic expansion and development in central regions[54]. However, the most highly developed coastal cities of the region, Shanghai, Suzhou, Hangzhou, Fuzhou, Xiamen, Taibei, and Singapore, have the smallest increase in mean emissions from the a priori, being a relatively smaller factor

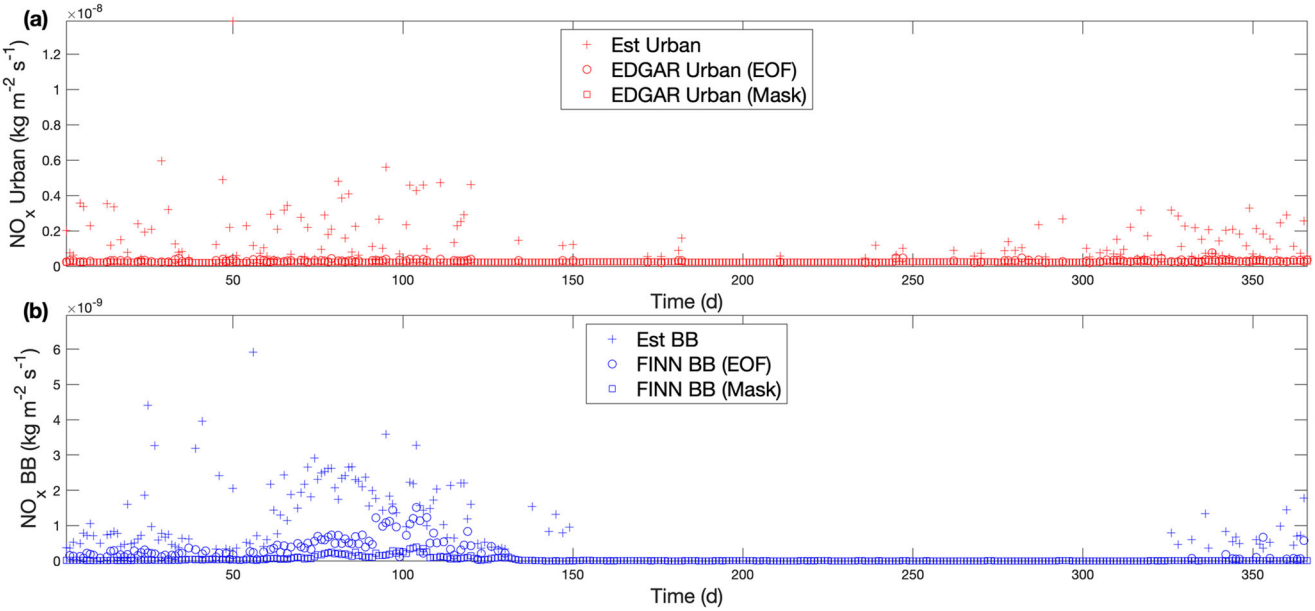

Fig. 6 | Time series of estimated and emission inventory-based NO_x emissions. **a** Est Urban, EDGAR Urban (EOF), and EDGAR Urban (Mask) denote the time series of daily mean estimated NO_x emissions over urban regions, EDGAR emissions over urban EOF regions and EDGAR emissions over urban mask regions, (**b**) Est BB, FINN BB (EOF) and FINN BB (Mask) denote time series of daily mean estimated NO_x emissions over BB regions, FINN emissions over BB EOF regions and FINN emissions over BB mask regions.

of 1.8 higher, consistent with higher quality and less rapid growth, as well as better a priori knowledge[55]. The normalized standard deviation of emissions in Northern Laos, most of Myanmar, Northern Thailand, NEI, and Cambodia is a factor of 0.87 lower than the normalized standard deviation of the a priori inventories (Fig. 5), consistent with the fact that there are both 11 days of emissions covering 5.7% of the domain which this approach has identified but are missed by FINN.

Over the entirety of the domain, the day-to-day statistics of NO_x emissions are given in Fig. 6. Due to low variability, the overall EDGAR a priori NO_x emission over the EOF regions has an average value of 3.2×10^{-10} kg m^{-2} s^{-1}, compared with the average value over the EDGAR domain of 2.4×10^{-10} kg m^{-2} s^{-1}. The FINN emission over the EOF regions has an average of 1.4×10^{-10} kg m^{-2} s^{-1}, with a high value from day 1 to day 130 and a low value otherwise, compared with the average value over the FINN domain of 3.3×10^{-11} kg m^{-2} s^{-1}. This work's NO_x emissions exhibit a larger mean and expanded variability, with [mean, daily variation, and uncertainty] respectively of [1.8×10^{-9}, 1.6×10^{-9}, 5.1×10^{-10}] kg m^{-2} s^{-1} and [1.4×10^{-9}, 1.1×10^{-9}, 5.4×10^{-10}] kg m^{-2} s^{-1} in urban and biomass burning areas. This work's emissions are particularly high in urban areas from days 1 to 35, days 50 to 125, and days 280 to 366 with the [mean, daily variation, and uncertainty] respectively [1.9×10^{-9}, 1.7×10^{-9}, 5.0×10^{-10}] kg m^{-2} s^{-1}, corresponding to when there is decreased UV radiation and slower chemical loss, as well as increased coal combustion for heating and end-of-year factory production. The emissions are particularly low from days 125 to 280 with [mean, daily variation, and uncertainty] respectively [9.3×10^{-10}, 3.4×10^{-10}, 5.3×10^{-10}] kg m^{-2} s^{-1}, corresponding to more UV radiation and faster atmospheric chemistry, coupled with less power demand required for heating and more available hydropower during the rainy season.

Over BB regions, the differences in emissions are even larger between the high and the low periods, with combustion being the primary difference. The [mean, daily variation, and uncertainty] of NO_x emissions over BB areas from days 60 to 150 is [1.9×10^{-9}, 6.9×10^{-10}, 7.5×10^{-10}] kg m^{-2} s^{-1} while the respective values are [1.0×10^{-9}, 1.2×10^{-9}, 3.6×10^{-10}] kg m^{-2} s^{-1} from days 1 to 60 and days 325 to 366. It is noted that the average NO_x emissions are both very high and relatively consistent during the biomass burning period from days 60 to 150, with low day-to-day variability and relatively low error. During the other biomass burning periods, the average is less high, the day-to-day variability is higher, and the error is quite low, indicating that day-to-day variability is driving the biomass burning emissions from days 1 to 60 and from days 325 to 366. This clearly demonstrates that there are both continuous and variable phases of the burning occurring in these regions, consistent with some amount of anthropogenic forcing involved.

It is noted that the uncertainty is only larger than the day-to-day variability in urban areas during very low emissions days in urban areas. The uncertainty is smaller than the day-to-day variability in all other cases (urban areas with medium and high emissions, and all biomass burning areas). The uncertainty is also smaller than the mean value under all conditions, indicating that the results are nearly always statistically relevant.

The computed day-to-day emissions total mean±error over all of the EOF regions and times are found to be 61.0 ± 32.6 kt d^{-1} from biomass burning in Northern CSA (44.3 kt d^{-1} more than FINN), 4.0 ± 2.3 kt d^{-1} from biomass burning in Southern CSA (3.2 kt d^{-1} more than FINN), 14.3 ± 6.1 kt d^{-1} from urbanization in China (5.0 kt d^{-1} more than EDGAR), and 5.1 ± 3.2 kt d^{-1} from urbanization in NEI and Bangladesh (3.7 kt d^{-1} more than EDGAR). The net NO_x emissions is 88.2 kt d^{-1}, compared with the 29.4 kt d^{-1} sum of FINN and EDGAR over the same region.

There is also a quantified connection between the spatial-temporal distribution of NO_x emissions and the land-use type (Fig. 7). Cropland areas have emissions that are lower (65.6% of emissions under 1.0×10^{-9}, 15.2% of emissions between 1.0×10^{-9} and 2.0×10^{-9}, and 19.2% between 2.0×10^{-9} and 5×10^{-9}), consistent with biomass burning being carefully controlled to clear new land for future agriculture and to clean out organic rubbish. Savanna regions have emissions that are higher than cropland emissions (39.7% of emissions under 1.0×10^{-9}, 33.1% of emissions between 1.0×10^{-9} and 2.0×10^{-9}, and 24.4% between 2.0×10^{-9} and 5×10^{-9}), consistent with these drier regions being more prone to accidental fires and spread, as well as some amount of peat burning. Broadleaf regions have the highest and most variable emissions of all land-use types (27.3% of emissions under 1.0×10^{-9}, 31.5% of emissions between 1.0×10^{-9} and 2.0×10^{-9}, and 37.8% between 2.0×10^{-9} and 5×10^{-9}), consistent with very hot fires when large trees burn, spread due to upwind/downwind slope effects, burning in remote areas being hard to control, and active control when clearing land for

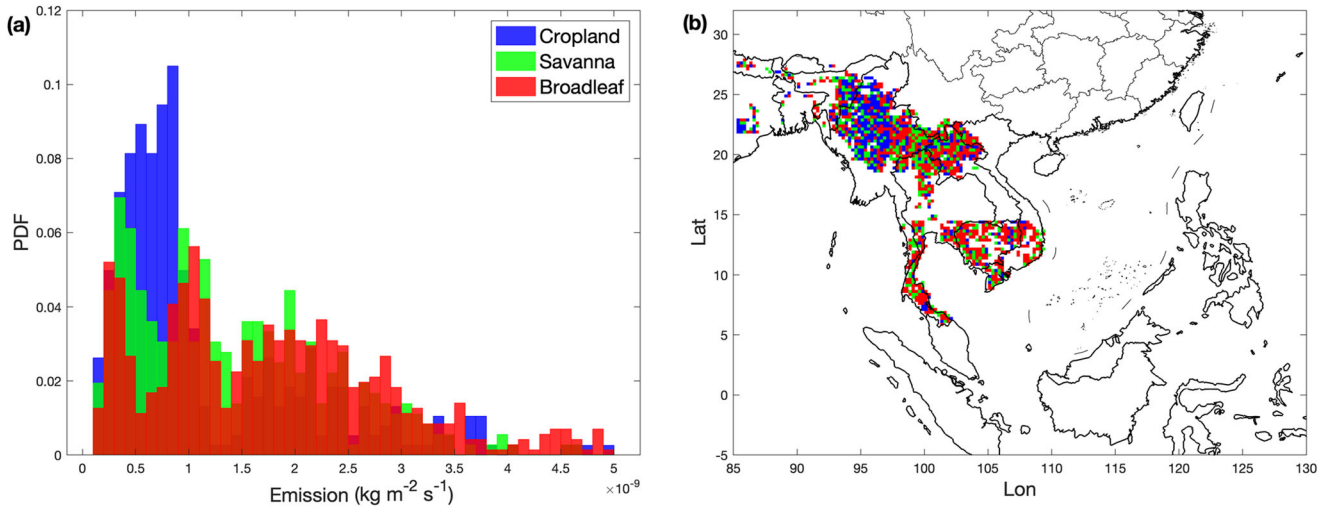

Fig. 7 | Relations between land cover types and spatial distribution of NO_x emission. a Probability density function of annual mean daily emission over the (**b**) geolocations of three different aggregated land cover types observed within the net BB areas.

new agriculture at the edges of existing land, or combustion of underbrush to support commercial trees such as wood, rubber, or palm, etc. In addition, there are further differences in similar land-use types as a function of policy, with differences in emissions observed within the same land use type on opposite sides of national boarders in: NEI and Bangladesh, NEI and Myanmar, the triangular region between Myanmar, Laos and Thailand, Thailand and Cambodia, and Cambodia and Vietnam.

These results can be applied to finer resolution satellite observations, such as TROPOMI NO_2 observations and to longer time series from merged OMI and GOME NO_2 observations, to obtain improved emissions in terms of spatial and temporal coverage. The method can be enhanced to account for new advances with respect to uncertainties of NO_2 retrievals, chemical methods, dynamics, and transport. The method has limitations based on satellite retrieval issues due to high cloud cover and the limited coverage of polar orbiting satellites. Furthermore, there are additional uncertainties associated with changes in the land surface due to burning and urbanization, changes in co-emitted BC aerosols leading to additional changes in the retrieved NO_2 signal, uncertainties associated with the vertical profile of the emissions, uncertainties associated with hourly and other high frequency extreme events, and other extreme events which are not sensitive enough to be picked up. More access to ground measurements, the next-generation geostationary satellites and improved knowledge and/or observations of the changes in the climate itself, could be used to address many of these issues. Finally, it is hoped that with more cooperation across the emissions community, that the improved emissions estimations in this would lead to a next generation of a priori emission databases, which in turn would iteratively assist the next generation of results using this approach.

Sensitivity of NO_x emissions

Two sensitivity runs are performed to examine how the uncertainty in OMI observations may impact the final emissions results. The first run scales all OMI measurements by −40% [herein called Base-40%] while the second run scales all OMI measurements by +40% [herein called Base+40%] compared to the default run [herein called Base]. The reason for this selection is that ±40% is considered the largest and smallest possible uncertainty in the retrieved OMI NO_2 column value found in the literature[56–63]. Both of these values consider issues including additional absorption by BC above what the a priori model driving the air mass factor computes[64–66], differences in the plume rise heights from the underlying models (which tend to be less over the regions studied here, although may be more over other regions of the world)[27,67,68], and upward looking observations from MAX-DOAS[58]. The point of these runs is not to comprehensively diagnose the uncertainties, but instead to put a maximum and minimum value bound on them, allowing a minimum and maximum value of the emissions to be quantified, and on grids and during days when this is larger than the signal, the user can then choose to include or discard them.

For each case, a new set of emissions is computed following the same procedures and using the same reanalysis and a priori emissions data. As demonstrated in Fig. 8, the differences on a grid-by-grid and day-by-day basis in urban areas are unbiased within the top 85% of data, while the differences on both bases in biomass burning areas are unbiased within the top 67% of data. On a grid-by-grid basis, the difference between Base-40% and Base is always less negative than -1.49×10^{-9} kg m^{-2} s^{-1}, while the difference between Base+40% and Base is always smaller than 0.33×10^{-9} kg m^{-2} s^{-1}. On a day-by-day basis the differences in urban areas are generally larger than the differences in biomass burning areas, where specifically the difference in urban areas between Base-40% and Base is always less negative than -2.31×10^{-9} kg m^{-2} s^{-1} and the difference between Base+40% and Base is always smaller than 2.44×10^{-9} kg m^{-2} s^{-1}, and the difference in biomass burning areas between Base-40% and Base is always less negative than -1.17×10^{-9} kg m^{-2} s^{-1} and the difference between Base+40% and Base is always smaller than 0.13×10^{-9} kg m^{-2} s^{-1}. It is important to note that the total number of grids in the three cases Base +40%, Base, and Base-40% are not the same (Supplementary Tables 1–3), and therefore all comparisons are being made over the regions of respective overlap.

The daily and yearly emission and their differences over different regions and total areas are further investigated as shown in Table 2. The ratio of daily and yearly differences between Base-40% and Base to the Base case over the total area are found respectively to be −22% and −23.6%, so are 9% and 10.9% for the differences between Base+40% and Base to the Base case. In all cases the differences are smaller than the 40% changes imposed on the OMI NO_2 column loadings, confirming that the approach is robust. A point of interest is that in general the difference between Base+40% and Base yields a much smaller magnitude in emissions than the difference between Base-40% and Base, especially in biomass burning areas. In all cases the low mean and variability in the a priori emissions data, particularly so in biomass burning areas, coupled with constraints placed on the ranges of α_1 and α_2, lead to a net larger constraint on the computed emissions. It is clear from this sensitivity study that the emissions computed in this work may be too low, although they are already higher than the a priori, especially in biomass burning regions.

The a priori emissions mean plus variability is generally so low that it does not offer enough data to constrain α_1 and α_2 in a physically reasonable way, in particular over biomass burning regions. This result is clearly demonstrated in terms of the decreased number of fit points per

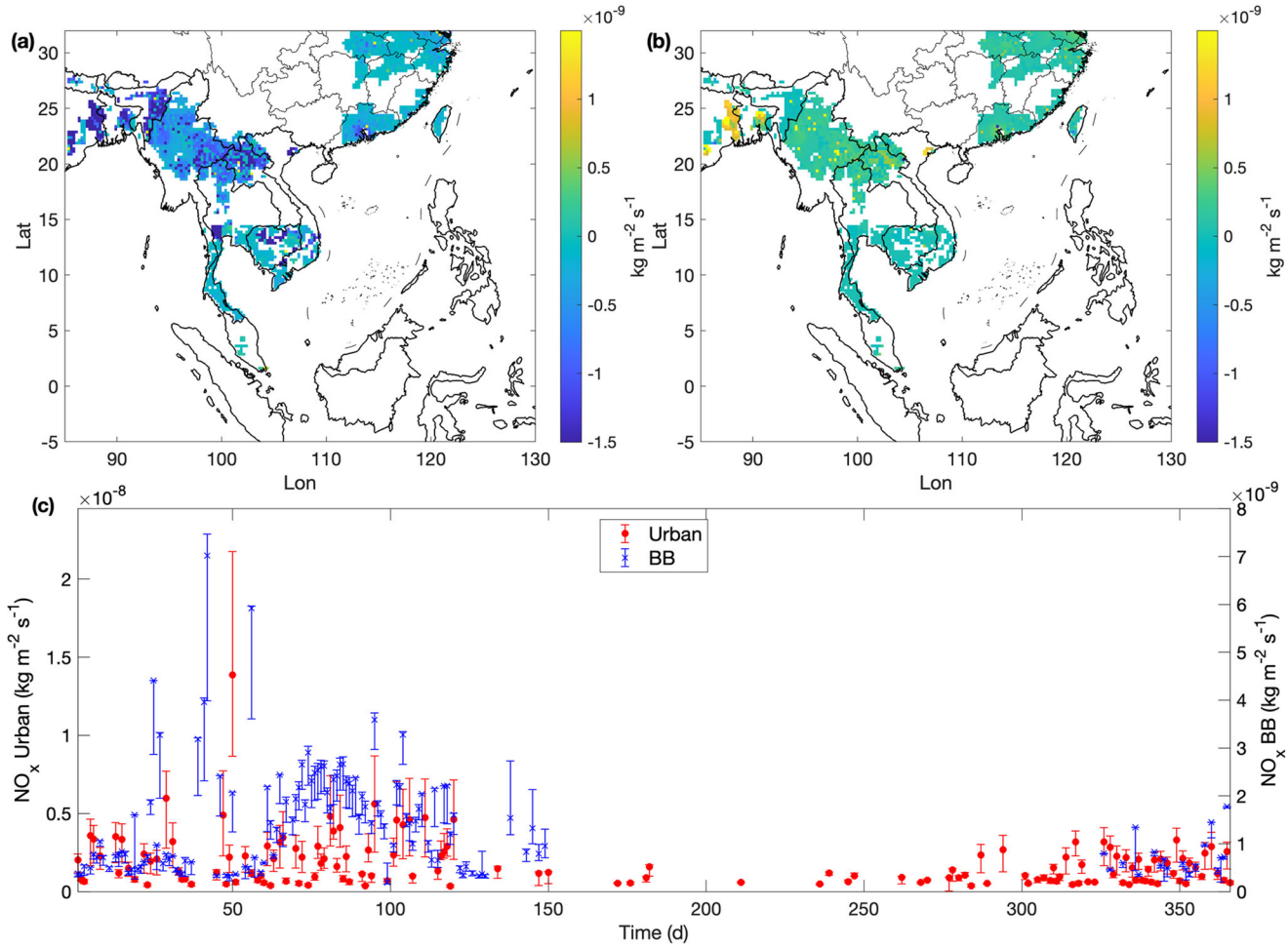

Fig. 8 | Spatial and temporal comparisons within the sensitivity analysis. Grid-by-grid and day-by-day differences of emissions sensitivity analysis: (**a**) Base-40% – Base, (**b**) Base+40% – Base, and (**c**) time series of day-by-day emission and the respective maximum and minimum values corresponding to the cases Base+40% and Base-40% over urban areas (red solid circle) and BB areas (blue asterisk), respectively.

EOF/PCA due to the physical constraints in terms of pairing in the Base +40% case as compared to standard emissions, as well as the slight increase in the number of fit points in the Base-40% case (Supplementary Tables 1–3). As can be observed, the values which are observed in the Base +40% case for α_1 are shifted very low while the values of α_2 are shifted very high, particularly in the biomass case, which are both consistent with colder air, as is expected from a larger fraction of the emissions being rapidly emitted at height. In the urban regions, the majority of the difference in case Base+40% is observed in the increase of α_3, which is more consistent with missing suburban sources or small industries located away from city centers, which is consistent with the fact that new sources are rapidly changing due to rapid economic growth and expansion, especially into regions which may not be as well regulated. The physical constraints supplied by the solution space of the three driving coefficients needing to be realistic applies additional constraints on the calculated emissions on a grid-by-grid and day-by-day basis. In fact, there is a number of grids for which OMI NO_2 data is available in each of these cases which are ultimately not considered in the emissions calculation, with the number of grids being discarded slightly different under each case (Base+40%, Base, and Base-40%), with the specific numbers and resulting products described in Supplementary Tables 1–3.

As observed, the emissions a priori seem very low compared to the inverted emissions, possibly requiring an iterative process between the top-down results given here and bottom-up processes done by others. Perhaps in a step-by-step manner, top-down and bottom-up communities can be used to improve each other. This is especially true since there are underlying non-linear physical issues, including but not limited to: rapid vertical lofting of emitted NO_x from both biomass burning and upslope/downslope effects, vast differences in scattered and absorbed UV radiation, and retrieval issues related to the introduction of new sources identified in regions previously assumed to have a value of zero.

Methods

Selected spatial and temporal domain of study

El Niño, a natural anomaly effecting the climate throughout East Asia, South Asia, and CSA, was particularly active from January to November 2016, resulting in major changes in equatorial water temperatures, air temperatures, and precipitation throughout most of the region studied in this work[50,54]. Furthermore, in 2016 the global distribution of the Indian Ocean Dipole and the North Atlantic Oscillation were also observed in CSA, with a similar set of effects. These net effects resulted in increased land-surface dryness, decreased atmospheric wet removal, stronger winds, and lower cloudiness, all leading to more drought and likely higher than normal emissions[50]. In fact, the period in January and February 2016 was the second hottest on record during the modern era, with the present year 2024 being the hottest[69]. For these reasons, it is expected that extremes of NO_x emissions should be relatively higher in 2016, presenting an excellent test of the methods in this paper, as well as providing more insight into how emissions will behave under the effects of global climate change over the next few decades. Therefore, the total year of 2016 is selected as the research period.

Table 2 | Summary of emissions over five aggregated regions on a day-to-day basis (top half) and a PCA integrated annual basis (bottom half)

Daily area	NO_x (Base)	NO_x (EI)	NO_x (EI - Base)	NO_x (Base-40%)	NO_x (Base-40% - Base)	NO_x (Base+40%)	NO_x (Base+40% - Base)
BB_NEI	3.66	1.18	−2.48	2.83	−0.83	3.68	0.02
BB_ICP	61.0	16.7	−44.3	49.2	−11.8	65.0	4.03
BB_Cam	4.05	0.82	−3.23	2.8	−1.25	4.08	0.03
UrbanChina	14.3	9.33	−4.98	11.0	−3.29	16.6	2.31
Urban3	5.13	1.42	−3.71	2.94	−2.19	6.74	1.61
Total Area	88.1	29.5	−58.7	68.8	−19.4	96.1	8.00
Unit for daily statistics above is kt d^{-1}.							
Annual area	**NO_x (Base)**	**NO_x (EI)**	**NO_x (EI - Base)**	**NO_x (Base-40%)**	**NO_x (Base-40% - Base)**	**NO_x (Base+40%)**	**NO_x (Base+40% - Base)**
BB_NEI	311.	80.6	−230.	240.	−70.4	313.	2.23
BB_ICP	2620	717.	−1910	2120	−508.	2.80×10^3	173.
BB_Cam	320.	62.3	−257.	221.	−98.3	322.	2.52
UrbanChina	1320	858.	−458.	1010	−302.	1530	213.
Urban3	513.	142.	−371.	294.	−219.	674.	161.
Total Area	5080	1860	−3230	3890	-1.20×10^3	5640	552.
Unit for annual statistics above is kt yr^{-1}.							

The left column is the Base emissions, the second column is the a priori emissions (EI), the third column is the difference between EI and Base, the fourth column is the Base-40% case emissions, the fifth column is the difference between Base-40% and Base, the sixth column is the Base+40% case, and the right column is the difference between Base+40% and Base.

OMI NO_2 dataset

NASA launched Aura on July 15, 2004, in a sun-synchronous, near polar (98.2-degree inclination) orbit 705 km above the Earth. OMI is a key instrument onboard Aura, capable of observing solar backscattered radiation utilizing hyperspectral imaging from the visible and ultraviolet parts of the spectrum[8], allowing detection of O_3, NO_2, SO_2, and aerosols. OMI provides a high resolution (13×24 km^2 at nadir) imaging capability, capable of tracking column loadings of NO_2 pollution at the scale of urban centers and large biomass burning sources. This work uses Level-3 daily global gridded NO_2 product with a $0.25 \times 0.25°$ resolution, including both cloud clearing and validation (Cloud Fraction<30%). This product has a known reliability of approximately 1.0×10^{15} molecules cm^{-2} [70]. All retrieved values which do not pass the quality assurance, are too low to be reliable, or are otherwise missing are not considered further in this work. While this may lead to an underestimation in the total coverage, interpolation has also been shown to present its own errors and additional uncertainties[11].

Supplementary Fig. 3 shows the daily measured climatological average of all observations of NO_2 in 2016. A higher value appears over known urban and biomass burning regions, including Southern China, NEI and parts of Southeast Asia such as Thailand, Laos and Vietnam. Large urban areas which have a higher level of economic development, mature industries, and very large populations are particularly pronounced, including the YRD and the PRD. Other rapidly developing large urban areas such as Bangkok, Dhaka, Wuhan, Chongqing, Chengdu, Xiamen, Changsha, Nanchang, and Hanoi, and highly developed but smaller urban areas such as Singapore and Taibei, are also clearly observed.

There are two important aspects of OMI NO_2 column retrieval bias[56,57]. The major issues discussed include co-absorption of the NO_2 by BC (which is a function of the SSA of the BC and the height of the NO_2 and BC) and the general underestimation of total columns as retrieved from satellite with respect to direct sun retrievals made from surface platforms. Given that the SSA in the regions of interest in this paper are observed to be very low (from 0.84 to 0.91 as measured by AERONET[71,72]) and the vertical height tends to be relatively high (from 1000 m to 3000 m)[73].

Land cover dataset

The MODIS land cover product includes annually computed land cover classifications using Terra and Aqua MODIS data in connection with a decision-tree classification method. Three different land cover classification schemes are applied to derive the Leaf Area Index (LAI) at a $0.05 \times 0.05°$ spatial resolution. This specific work uses the Version 6 land surface type data product specifically from 2016 based on the measurements of LAI, as displayed in Supplementary Fig. 4. Over the area in this study, the land use types of largest area are water (61%), savannah (16%), evergreen broadleaf forest (11%), and grasses/cereal (7.9%).

To analyze land cover in connection with emissions, the individual categories of land cover use are obtained over the same grid points of the net EOF BB regions. Similar land cover classes are grouped into three larger categories: Croplands corresponding to drier regions which are naturally irrigated or do not require irrigation and intensively used for agriculture including grass, cereal, and shrubs; Savanna which are drier and require intensive artificial irrigation or other intervention to be used for intensive agriculture; and Broadleaf which corresponds to the major forest lands, in particular found in hilly or inland jungle conditions.

Atmospheric reanalysis data

ERA-5 is an atmospheric reanalysis product by ECMWF, providing global-scale atmospheric wind speed and other physical and dynamical products from 1950 to the present, with a horizontal resolution of 0.25°[62]. The product is based on geophysical model physics including: long-wave radiation, a simplified linearized parametrization scheme for surface processes, ozone, improved land component and ocean waves, all within an all-sky data assimilation framework. The product is widely used in atmospheric models and remote sensing applications, including the AMF calculations underlying the NO_2 retrievals used in this work[63]. This work specifically uses horizontal wind near the surface (975hPa) and at high-altitude (800hPa), to drive the model for urban and biomass burning areas[20,33].

A priori emission inventories

FINN is a fire emissions product based on the 1-km level-2 active fire product derived from MODIS TERRA and AURA infrared measurements made in bands: B21(3.929-3.989 μm), B22(3.940-4.001 μm), and B31 (10.780-11.280 μm)[74], with the emissions from each NIR/IR plume based on the fuel availability, plume size, and intensity[14]. Due to coverage gaps between the adjacent orbits on a day-to-day basis, cloudiness, and optically thick smoke with an AOD larger than 2.0, there are many missed observations particularly so in the equatorial region between 30°S and 30°N latitudes[75]. These results have been used in a bottom-up manner with laboratory constrained emission factors and estimated environmental fuel loadings at a scale of 1×1 km^2 to estimate the emissions of selected trace

gasses and aerosols from biomass burning. This product has been used in many common atmospheric chemical transport models and is considered a standard emissions source[76]. It has been demonstrated that the product faces difficulty in cases of diagnosing small fires or cloud covered fires, which have caused models to suffer difficulty reconciling the actual environment with this dataset[77]. For this combination of reasons, this work utilizes FINNv1.5 from 2016 to retrieve the NO_x emission for biomass burning areas in Southeast Asia.

The Emission Database for Global Atmospheric Research (EDGAR) is an anthropogenic emission inventory product, computed using a bottom-up technology-based emission factor approach to calculate the emission for countries over the world and consists of greenhouse gases and air pollutants, such as carbon dioxide, methane, CO, NO_x, VOC, ammonia, etc[13]. This work uses the version 2.2 of EDGAR-HTAP, with a global gridded ($0.1° \times 0.1°$) anthropogenic emissions product, at monthly temporal resolution, for the years 2008 and 2010 (there is no current product for 2016). This is also a standard dataset used by many studies in the modeling and impacts communities[48,78].

Both emission inventories are interpolated to the corresponding OMI NO_2 grids ($0.25° \times 0.25°$) to be used in the following estimation processes shown in Supplementary Fig. 5.

Variance maximization

EOF is a mathematical method that decomposes a dataset in terms of orthogonal basis functions into the factors which contribute the most to the variability of the underlying dataset[32]. We use this method to extract the spatial and temporal features of the extremes of the remotely sensed NO_2 fields. This method has been used in the past for monthly-average climatological AOD, weekly-average climatological CO, and weekly-average climatological NO_2[27,32].

When performing an EOF analysis, the first step is to quantify the most relevant EOFs and PCs as a function of the magnitude of the eigenvalue. The first step to separating urban regions is to identify those regions which both (a) do have known large urban use land types, and (b) contain a large population, or industrial or residential economic usage. To this extent, this section follows the results as generated in Table 1. The cutoffs of the spatial and temporal modes are quantitatively determined employing a recursive cutoff based on the mean and one standard deviation, following the approach in Lin[27]. First, the weighted average NO_2 over the domain ($EOF(i,j)*NO_2(i,j)$) for all values i,j inside the domain is computed using the nine equally spaced percentiles from 10% through 90% of the distribution of EOF(i,j) over the domain. Second, the temporal correlation is calculated between the weighted NO_2 and the PC. The combination of lowest possible cutoff value with highest possible absolute value of the R statistic (as good a fit in time with the peaks in the PC) and the lowest possible increase in additional RMSE error for each datapoint added (as good a fit in magnitude with the peaks in the PC) yields the appropriate domain in space. Third, in terms of temporal fit, the appropriate subset of times is obtained from all data in which the PCs fall outside of the mean plus/minus one standard deviation. When the same geographic area is identified in multiple EOFs the time series of these respective modes are different and can be aggregated based on the magnitude of the EOF, so as to use the orthogonality to do a complete reconstruction.

In all cases, this approach guarantees that our spatial-temporal domain contains the most substantial signals of the remotely sensed NO_2 fields, which in turn contribute to the maximum changes. These changes are in turn those most responsible for changes in emissions of NO_x. These sources which are changing the most are those which are most likely to be mis-diagnosed by current emissions inventories for two major reasons. First, current emissions inventories do not consider day-to-day and other high-frequency variations on a grid-by-grid basis, which this approach makes explicit. Second, some fraction of emissions from biomass burning and regions undergoing rapid urbanization are completely missing in existing emissions inventories[20,32], which this work also captures, since it is analyzing the tropospheric column, not merely the surface.

Selection of urban and non-urban areas

Due to the considerably economic and political diversity in CSA, there are many people who live in rapidly changing communities in countries such as Vietnam, Laos, Cambodia, Thailand, Myanmar, Indonesia and Malaysia. Within sizable portions of each of these countries, there still is a large amount of biomass used for cooking and clearing of farmland, with most occurring during the local dry season from mid-February through mid-April. The conditions are climatologically similar, although with a very different economic and political profile in NEI and Bangladesh. Due to the higher level of economic development and more strict government policies, there are fewer people in southern China cooking with biomass. On top of this, there is a rapidly increasing number of automobiles, wide-spread city construction, and even small but frequent wildfires that are observed in the local dry season. Overall, throughout the regions of interest there are sizable sources from both biomass and urban sources, which contribute to time-varying intense air pollution events across many broad air pollutants.

Due to their orthogonal and specific nature, this work utilizes as an a priori the respective NO_x emission from FINN for biomass burning locations and EDGAR for urban regions. The total emissions have been distributed into 9 separate groups, as shown in Table 1. There are 3 specific BB areas: BB1, BB2, BB3, three specific urban areas: Urban1, Urban2, Urban3 and three other areas: a Tibetan area, a Mixed area and an Equatorial area. The Mixed Area is the only one which uses data from both FINN and EDGAR, when and where overlap actually occurs.

The region defined as Urban1 area mainly covers the urban regions over provinces along the Yangtze River, such as Sichuan, Chongqing, Guizhou, Yunnan, Guangxi, Hunan, Hubei, Anhui, Jiangxi, Zhejiang, Shanghai and Jiangsu, which are mainly impacted by the subtropical mid-latitude climate. In these areas, cities in the YRD are more developed and have a larger population density than the other cities. The region defined as Urban2 covers the same-latitude urban regions of Guangdong, Fujian, Taiwan Island, and Hong Kong, which are climatologically impacted by sea, mountains, and the Asian Monsoon, and have a high population density and extensive industry. The region defined as Urban3 includes the major cities of CSA including Bangkok, Kuala Lumpur, Singapore, Hanoi, Ho Chi Minh City, Manila, as well as Bangladesh's Dhaka, and some cities in NEI, like Calcutta. Cities here are impacted by tropical climate and generally have both a large population and high population density, and in general are less developed economically than the former two areas (with the exception of Singapore). The region defined as the Mixed area includes the remaining areas of China not previously mentioned, and not contained within BB2.

The region defined as BB1 covers the BB areas in NEI, Nepal, Bangladesh and Bhutan. The region defined as BB2 covers Southwestern Yunnan Province, a small part of NEI and most of Northern CSA, including Myanmar, Northern Thailand, Laos and Northern Vietnam. The region defined as BB3 covers the other areas of CSA except for Peninsular Malaysia and Singapore. These BB areas are mainly impacted by a Tropical monsoon climate, including both the Indian Monsoon and the Asian Monsoon.

The region defined as the Tibetan Area mainly includes the Tibet and the mountains of Western Sichuan, as well as some nearby high-altitudes regions. Overall, this area has a very low population density, a low level of atmospheric pollution, and a sizable fraction of their emissions coming from wild fires. The Mixed Area has a combination of small to medium cities, mountains and forests, and therefore has characteristics of the anthropogenic and biomass burning types. The Equatorial Area includes the islands around the equator in the Maritime Continent except for Singapore (Indonesia, Malaysia and Brunei) as well as the Philippines. This area has a high occurrence of wild fires each year in August through October, and a large amount of rapidly developing and highly dense urban areas. However, this region is also heavily impacted by rain, and therefore does not have many days of measurements available.

Mass conserving equation

The basic mass-conserving equation for a tracer in the atmosphere (in this case NO_x) is shown in Eq. (1). The terms include the time rate of change between the previous day's and current day's loading of NO_x, $\frac{\partial(V_{NO_x})}{\partial t}$, the emissions of NO_x, E_{NO_x}, the loss of NO_x due to chemical decay, $Loss_{NO_x}$, and the two transport terms of NO_x corresponding to advective transport and pressure-transport, $\nabla\left(\bar{\mathbf{u}} \cdot V_{NO_x}\right)$, where $\bar{u}$ denotes the horizontal wind vector (in terms of both u and v) and V denotes the tropospheric column as the previous studies[5]. Three simplifications are then made to allow the solution to be readily solved within the context of the available measurements from OMI. First, there is a linearization between NO_x and NO_2, α_1, whereby the ratio of NO_x/NO_2 in Eq. (2), is based on the thermodynamics of the NO_x formation in the flame[1]. Second, the loss term is linearized in Eq. (3), where α_2 is the rate of reaction times the concentration of OH, C_{OH}, responsible for the major conversion of atmospheric NO_x into nitric acid[79]. The third simplification as given in Eq. (4), α_3 denotes the weighted distance of the horizontal grid over which the transport of NO_x occurs. Overall, these terms are merged together in Eq. (5). In this equation, the grids missing OMI NO_2 data will be discarded, only the grids with OMI NO_2 observations will be used. The a priori emission inventories will also be used after interpolation to the same resolution (0.25°x0.25°) based on the OMI NO_2 product.

$$\frac{\partial(V_{NO_x})}{\partial t} = E_{NO_x} - \mathrm{Loss}_{NO_x} + \nabla\left(\bar{\mathbf{u}} \cdot V_{NO_x}\right) \tag{1}$$

$$\frac{\partial(V_{NO_x})}{\partial t} = \frac{d(\alpha_1 V_{NO_2})}{dt} = \alpha_1 \frac{dV_{NO_2}}{dt} \tag{2}$$

$$\mathrm{Loss}_{NO_x} \approx \alpha_2{}' * C_{OH} * V_{NO_x} = \alpha_2{}' * C_{OH} * \alpha_1 * V_{NO_2} = \alpha_2 \alpha_1 V_{NO_2} \tag{3}$$

$$\nabla\left(\bar{\mathbf{u}} \cdot V_{NO_x}\right) = \alpha_3 \nabla\left(\bar{\mathbf{u}} \cdot \alpha_1 V_{NO_2}\right) = \alpha_3 \alpha_1 \nabla\left(\bar{\mathbf{u}} \cdot V_{NO_2}\right) \tag{4}$$

$$\alpha_1 \frac{dV_{NO_2}}{dt} = E_{NO_x} - \alpha_2 \alpha_1 V_{NO_2} + \alpha_3 \alpha_1 \nabla\left(\bar{\mathbf{u}} \cdot V_{NO_2}\right) \tag{5}$$

To find the best fit values for α_1, α_2, and α_3, a least square method is used. This is a statistical procedure to find the best fit by minimizing the sum of the squares of the residuals[80], as given by Eq. (6), where $f(x_i)$ is based on Eq. (5), and y_i is given based on the measured a priori emission value. In this work, a prior is given by the sum of the NO_x emissions from FINN and EDGAR.

$$\min\left\{ \mathrm{f}(x) = \sum_{i=1}^{m} \left[y_i - f(x_i) \right]^2 \right\} \tag{6}$$

The fits of α_1, α_2, and α_3, and values of computed weighted NO_x over the EOF fields in comparison to the PCs are further explained using correlation, the rmse, and simple probability density functions (PDF) and analysis.

Bootstrapping

Bootstrapping is a statistical technique which utilizes random sampling with repeatability of replacement, specifically capable of handling all types of PDFs. This method is capable of being computed on a laptop or desktop using commonly available software, in this work MATLAB was used. This method can be widely applied to estimate the variation of statistics (bias, variance, confidence intervals, etc.)[81]. In this work, this method has been used to sample the values of α_1, α_2, and α_3, which are then in turn used to compute different permutations of emissions from the already constrained PDFs of the coefficient values and their associated uncertainty ranges. The uncertainty ranges are computed specifically by computing half of the difference between the 25th-percentile and 75th-percentile results of the bootstrap distribution on a pixel-by-pixel and day-by-day basis. These numbers are then combined in space and time via the root-mean-square techniques to form a region-by-region uncertainty on a day-by-day and year-by-year basis. Finally, the mean and standard deviation of these derived individually computed emissions are finalized over each specific sub-domain, each month, and then finally merged to cover the entire domains of interest, in a bottom-up manner.

Emissions calculation

Based on the workflow chart of NO_x estimation presented in Supplementary Fig. 5, daily emissions of NO_x are calculated throughout the entire year using Eq. (5) once the values of α_1, α_2, and α_3 as derived and bootstrapping has been undertaken. Spatially averaged emissions are calculated over the extracted respective urban and EOF BB regions at the corresponding time periods. Masked NO_x emissions are calculated over the default urban and BB areas shown in Fig. 1, without time constraints.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The satellite NO_2 datasets used in this study are available at https://doi.org/10.5067/Aura/OMI/DATA3007, the ERA-5 reanalysis product is available at https://doi.org/10.24381/cds.bd0915c6. FINN and EDGAR are retrieved from https://www.acom.ucar.edu/Data/fire and https://edgar.jrc.ec.europa.eu/dataset_htap_v2 respectively. The MODIS land cover product is distributed at https://doi.org/10.5067/MODIS/MCD12C1.006. All of the data and underlying figures are available for download at https://doi.org/10.6084/m9.figshare.19334180.

Code availability

MATLAB code used to calculate the emissions is available at https://doi.org/10.6084/m9.figshare.19334180.

Received: 7 November 2023; Accepted: 25 April 2024;
Published online: 13 May 2024

References

1. Andreae, M. O. & Merlet, P. Emission of trace gases and aerosols from biomass burning. *Glob. Biogeochem. Cycles* **15**, 955–966 (2001).
2. Van Der, A. R. J. et al. Trends, seasonal variability and dominant NO_x source derived from a ten year record of NO_2 measured from space. *J. Geophys. Res.* **113**, D04302 (2008).
3. Tabor, K., Gutzwiller, L. & Rossi, M. J. The heterogeneous interaction of NO_2 with amorphous carbon. *Geophys. Res. Lett.* **20**, 1431–1434 (1993).
4. Singh, H. B. Reactive nitrogen in the troposphere. *Environ. Sci. Technol.* **21**, 320–327 (1987).
5. Seinfeld, J. H. & Pandis, S. N. *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* (John Wiley & Sons, New York, 1998).
6. Clapp, L. Analysis of the relationship between ambient levels of O_3, NO_2 and NO as a function of NO_x in the UK. *Atmos. Environ.* **35**, 6391–6405 (2001).
7. Cyrys, J. et al. Variation of NO_2 and NO_x concentrations between and within 36 European study areas: results from the ESCAPE study. *Atmos. Environ.* **62**, 374–390 (2012).
8. Levelt, P. F. et al. The ozone monitoring instrument. *IEEE Trans. Geosci. Remote Sens.* **44**, 1093–1101 (2006).
9. Chen, C. H. & Peter Ho, P.-G. Statistical pattern recognition in remote sensing. *Pattern Recognit.* **41**, 2731–2741 (2008).
10. Gu, B. et al. Abating ammonia is more cost-effective than nitrogen oxides for mitigating $PM_{2.5}$ air pollution. *Science* **374**, 758–762 (2021).
11. He, Q. et al. Spatially and temporally coherent reconstruction of tropospheric NO_2 over China combining OMI and GOME-2B measurements. *Environ. Res. Lett.* **15**, 125011 (2020).

12. Bond, T. C. et al. A technology-based global inventory of black and organic carbon emissions from combustion. *J. Geophys. Res.: Atmos*. **109**, D14203 (2004).
13. Janssens-Maenhout, G. et al. HTAP_v2.2: a mosaic of regional and global emission grid maps for 2008 and 2010 to study hemispheric transport of air pollution. *Atmos. Chem. Phys.* **15**, 11411–11432 (2015).
14. Wiedinmyer, C. et al. The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning. *Geosci. Model Dev.* **4**, 625–641 (2011).
15. Roden, C. A. et al. Laboratory and field investigations of particulate and carbon monoxide emissions from traditional and improved cookstoves. *Atmos. Environ.* **43**, 1170–1181 (2009).
16. Yan, X., Ohara, T. & Akimoto, H. Bottom-up estimate of biomass burning in mainland China. *Atmos. Environ.* **40**, 5262–5273 (2006).
17. Paton-Walsh, C., Smith, T. E. L., Young, E. L., Griffith, D. W. T. & Guérette, É.-A. New emission factors for Australian vegetation fires measured using open-path Fourier transform infrared spectroscopy – Part 1: Methods and Australian temperate forest fires. *Atmos. Chem. Phys.* **14**, 11313–11333 (2014).
18. van der Werf, G. R. et al. Global fire emissions and the contribution of deforestation, savanna, forest, agricultural, and peat fires (1997–2009). *Atmos. Chem. Phys.* **10**, 11707–11735 (2010).
19. Andreae, M. O. Emission of trace gases and aerosols from biomass burning – an updated assessment. *Atmos. Chem. Phys.* **19**, 8523–8546 (2019).
20. Wang, S., Cohen, J. B., Lin, C. & Deng, W. Constraining the relationships between aerosol height, aerosol optical depth and total column trace gas measurements using remote sensing and models. *Atmos. Chem. Phys.* **20**, 15401–15426 (2020).
21. Meirink, J. F., Bergamaschi, P. & Krol, M. C. Four-dimensional variational data assimilation for inverse modelling of atmospheric methane emissions: method and comparison with synthesis inversion. *Atmos. Chem. Phys.* **8**, 6341–6353 (2008).
22. Miyazaki, K., Eskes, H. J. & Sudo, K. Global NO_x emission estimates derived from an assimilation of OMI tropospheric NO_2 columns. *Atmos. Chem. Phys.* **12**, 2263–2288 (2012).
23. Sayer, A. M., Hsu, N. C., Bettenhausen, C. & Jeong, M.-J. Validation and uncertainty estimates for MODIS collection 6 "Deep Blue" aerosol data. *J. Geophys. Res.: Atmos.* **118**, 7864–7872 (2013).
24. Levy, R. C. et al. Global evaluation of the Collection 5 MODIS dark-target aerosol products over land. *Atmos. Chem. Phys.* **10**, 10399–10420 (2010).
25. Zheng, B. et al. Record-high CO_2 emissions from boreal fires in 2021. *Science* **379**, 912–917 (2023).
26. Chen, Y.-H. & Prinn, R. G. Estimation of atmospheric methane emissions between 1996 and 2001 using a three-dimensional global chemical transport model. *J. Geophys. Res.: Atmos*. **111**, D10307 (2006).
27. Lin, C., Cohen, J. B., Wang, S. & Lan, R. Application of a combined standard deviation and mean based approach to MOPITT CO column data, and resulting improved representation of biomass burning and urban air pollution sources. *Remote Sens. Environ.* **241**, 111720 (2020).
28. Cohen, J. B. & Wang, C. Estimating global black carbon emissions using a top-down Kalman Filter approach. *J. Geophys. Res.: Atmos.* **119**, 307–323 (2014).
29. Kong, H. et al. Considerable unaccounted local sources of NO_x emissions in China revealed from satellite. *Environ. Sci. Technol.* **56**, 7131–7142 (2022).
30. McLinden, C. A. et al. Space-based detection of missing sulfur dioxide sources of global air pollution. *Nat. Geosci.* **9**, 496–500 (2016).
31. Russell, A. R., Valin, L. C., Bucsela, E. J., Wenig, M. O. & Cohen, R. C. Space-based constraints on spatial and temporal patterns of NO_x emissions in California, 2005−2008. *Environ. Sci. Technol.* **44**, 3608–3615 (2010).
32. Cohen, J. B. Quantifying the occurrence and magnitude of the Southeast Asian fire climatology. *Environ. Res. Lett.* **9**, 114018 (2014).
33. Lin, C., Cohen, J. B., Wang, S., Lan, R. & Deng, W. A new perspective on the spatial, temporal, and vertical distribution of biomass burning: quantifying a significant increase in CO emissions. *Environ. Res. Lett.* **15**, 104091 (2020).
34. Weiss, R. F. & Prinn, R. G. Quantifying greenhouse-gas emissions from atmospheric measurements: a critical reality check for climate legislation. *Phil. Trans. R. Soc. A* **369**, 1925–1942 (2011).
35. Li, C. & Cohen, R. C. Space-borne estimation of volcanic sulfate aerosol lifetime. *J. Geophys. Res. Atmos.* **126**, e2020JD033883 (2021).
36. Beirle, S. et al. Pinpointing nitrogen oxide emissions from space. *Sci. Adv.* **5**, eaax9800 (2019).
37. Fibiger, D. L. et al. Wintertime overnight NO_x removal in a Southeastern United States coal-fired power plant plume: a model for understanding winter NO_x processing and its implications. *J. Geophys. Res. Atmos.* **123**, 1412–1425 (2018).
38. Vadrevu, K. P., Giglio, L. & Justice, C. Satellite based analysis of fire–carbon monoxide relationships from forest and agricultural residue burning (2003–2011). *Atmos. Environ.* **64**, 179–191 (2013).
39. Aouizerats, B., Van Der Werf, G. R., Balasubramanian, R. & Betha, R. Importance of transboundary transport of biomass burning emissions to regional air quality in Southeast Asia during a high fire event. *Atmos. Chem. Phys.* **15**, 363–373 (2015).
40. Yadav, I. C. et al. Biomass burning in Indo-China Peninsula and its impacts on regional air quality and global climate change-a review. *Environ. Pollut.* **227**, 414–427 (2017).
41. Lee, H.-H., Bar-Or, R. Z. & Wang, C. Biomass burning aerosols and the low-visibility events in Southeast Asia. *Atmos. Chem. Phys.* **17**, 965–980 (2017).
42. Yin, S. et al. Influence of biomass burning on local air pollution in mainland Southeast Asia from 2001 to 2016. *Environ. Pollut.* **254**, 112949 (2019).
43. Lu, X. et al. Observation-derived 2010–2019 trends in methane emissions and intensities from US oil and gas fields tied to activity metrics. *Proc. Natl. Acad. Sci. USA* **120**, e2217900120 (2023).
44. Cox, L. *Nitrogen Oxides (NO_x), Why and How They Are Controlled*. http://www.epa.gov/ttn/catc (1999).
45. Laughner, J. L. & Cohen, R. C. Direct observation of changing NO_x lifetime in North American cities. *Science* **366**, 723–727 (2019).
46. Ahern, A. T. et al. Production of secondary organic aerosol during aging of biomass burning smoke from fresh fuels and its relationship to VOC precursors. *J. Geophys. Res. Atmos.* **124**, 3583–3606 (2019).
47. Cohen, J. B., Prinn, R. G. & Wang, C. The impact of detailed urban-scale processing on the composition, distribution, and radiative forcing of anthropogenic aerosols. *Geophys. Res. Lett*. **38**, L10808 (2011).
48. Cohen, J. B. & Prinn, R. G. Development of a fast, urban chemistry metamodel for inclusion in global models. *Atmos. Chem. Phys.* **11**, 7629–7656 (2011).
49. Zhang, B.-N. & Kim Oanh, N. T. Photochemical smog pollution in the Bangkok Metropolitan Region of Thailand in relation to O_3 precursor concentrations and meteorological conditions. *Atmos. Environ.* **36**, 4211–4222 (2002).
50. Wang, S., Cohen, J. B., Deng, W., Qin, K. & Guo, J. Using a new top-down constrained emissions inventory to attribute the previously unknown source of extreme aerosol loadings observed annually in the monsoon asia free troposphere. *Earth's Future* **9**, 2021EF002167 (2021).
51. Randel, W. J. et al. Asian monsoon transport of pollution to the stratosphere. *Science* **328**, 611–613 (2010).

52. Marlier, M. E. et al. El Niño and health risks from landscape fire emissions in Southeast Asia. *Nat. Clim. Change* **3**, 131–136 (2013).
53. Wu, M., Wu, J. & Zang, C. A comprehensive evaluation of the eco-carrying capacity and green economy in the Guangdong-Hong Kong-Macao Greater Bay Area. *China. J. Cleaner Prod.* **281**, 124945 (2021).
54. Guan, X., Wei, H., Lu, S., Dai, Q. & Su, H. Assessment on the urbanization strategy in China: achievements, challenges and reflections. *Habitat Int*. **71**, 97–109 (2018).
55. Su, C.-W., Song, Y. & Umar, M. Financial aspects of marine economic growth: from the perspective of coastal provinces and regions in China. *Ocean Coastal Manage*. **204**, 105550 (2021).
56. Chimot, J., Veefkind, J. P., de Haan, J. F., Stammes, P. & Levelt, P. F. Minimizing aerosol effects on the OMI tropospheric NO_2 retrieval – An improved use of the 477nm $O_2 - O_2$ band and an estimation of the aerosol correction uncertainty. *Atmos. Meas. Tech.* **12**, 491–516 (2019).
57. Compernolle, S. et al. Validation of Aura-OMI QA4ECV NO_2 climate data records with ground-based DOAS networks: the role of measurement and comparison uncertainties. *Atmos. Chem. Phys.* **20**, 8017–8045 (2020).
58. Verhoelst, T. et al. Ground-based validation of the copernicus Sentinel-5P TROPOMI NO_2 measurements with the NDACC ZSL-DOAS, MAX-DOAS and Pandonia global networks. *Atmos. Meas. Tech.* **14**, 481–510 (2021).
59. Pinardi, G. et al. Validation of tropospheric NO_2 column measurements of GOME-2A and OMI using MAX-DOAS and direct sun network observations. *Atmos. Meas. Tech.* **13**, 6141–6174 (2020).
60. Irie, H. et al. Quantitative bias estimates for tropospheric NO_2 columns retrieved from SCIAMACHY, OMI, and GOME-2 using a common standard for East Asia. *Atmos. Meas. Tech.* **5**, 2403–2411 (2012).
61. Boersma, K. F., Eskes, H. J. & Brinksma, E. J. Error analysis for tropospheric NO_2 retrieval from space. *J. Geophys. Res.: Atmos*. **109**, D04311 (2004).
62. Wang, Y. et al. Validation of OMI, GOME-2A and GOME-2B tropospheric NO_2, SO_2 and HCHO products using MAX-DOAS observations from 2011 to 2014 in Wuxi, China: investigation of the effects of priori profiles and aerosols on the satellite products. *Atmos. Chem. Phys.* **17**, 5007–5033 (2017).
63. Wang, C., Wang, T., Wang, P. & Rakitin, V. Comparison and validation of TROPOMI and OMI NO_2 observations over China. *Atmosphere* **11**, 636 (2020).
64. Tiwari, P., Cohen, J. B., Wang, X., Wang, S. & Qin, K. Radiative forcing bias calculation based on COSMO (Core-Shell Mie model Optimization) and AERONET data. *npj Clim. Atmos. Sci.* **6**, 1–14 (2023).
65. Giles, D. M. et al. An analysis of AERONET aerosol absorption properties and classifications representative of aerosol source regions. *J. Geophys. Res.: Atmos*. **117**, D17203 (2012).
66. Che, H. et al. Aerosol optical properties and direct radiative forcing based on measurements from the china aerosol remote sensing network (CARSNET) in eastern China. *Atmos. Chem. Phys.* **18**, 405–425 (2018).
67. Reid, J. S. et al. A review of biomass burning emissions part III: Intensive optical properties of biomass burning particles. *Atmos. Chem. Phys.* **5**, 827–849 (2005).
68. Lin, N.-H. et al. An overview of regional experiments on biomass burning aerosols and related pollutants in Southeast Asia: From BASE-ASIA and the Dongsha Experiment to 7-SEAS. *Atmos. Environ.* **78**, 1–19 (2013).
69. Copernicus. February 2024 was globally the warmest on record – Global Sea Surface Temperatures at record high | https://climate.copernicus.eu/copernicus-february-2024-was-globally-warmest-record-global-sea-surface-temperatures-record-high (2024).
70. Krotkov, N. A. et al. The version 3 OMI NO_2 standard product. *Atmos. Meas. Tech.* **10**, 3133–3149 (2017).
71. Wang, S., Wang, X., Cohen, J. B. & Qin, K. Inferring polluted asian absorbing aerosol properties using decadal scale AERONET measurements and a MIE model. *Geophys. Res. Lett*. **48** (2021).
72. Lin, N.-H. et al. Interactions between biomass-burning aerosols and clouds over Southeast Asia: current status, challenges, and perspectives. *Environ. Pollut.* **195**, 292–307 (2014).
73. Cohen, J. B., Ng, D. H. L., Lim, A. W. L. & Chua, X. R. Vertical distribution of aerosols over the maritime continent during El Niño. *Atmos. Chem. Phys.* **18**, 7095–7108 (2018).
74. Lorente, A. et al. Structural uncertainty in air mass factor calculation for NO_2 and HCHO satellite retrievals. *Atmos. Meas. Tech.* **10**, 759–782 (2017).
75. Levelt, P. F. et al. The ozone monitoring instrument: overview of 14 years in space. *Atmos. Chem. Phys.* **18**, 5699–5745 (2018).
76. Li, F. et al. Historical (1700–2012) global multi-model estimates of the fire emissions from the Fire Modeling Intercomparison Project (FireMIP). *Atmos. Chem. Phys.* **19**, 12545–12567 (2019).
77. Kahn, R. A global perspective on wildfires. *Eos* **101**, https://doi.org/10.1029/2020EO138260 (2020).
78. Galmarini, S. et al. Technical note: coordination and harmonization of the multi-scale, multi-model activities HTAP2, AQMEII3, and MICS-Asia3: simulations, emission inventories, boundary conditions, and model output formats. *Atmos. Chem. Phys.* **17**, 1543–1555 (2017).
79. Kelly, T. J., Stedman, D. H., Ritter, J. A. & Harvey, R. B. Measurements of oxides of nitrogen and nitric acid in clean air. *J. Geophys. Res.: Oceans* **85**, 7417–7425 (1980).
80. Cantrell, C. A. Technical note: review of methods for linear least-squares fitting of data and application to atmospheric chemistry problems. *Atmos. Chem. Phys.* **8**, 5477–5487 (2008).
81. IPCC. *2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories* (IPCC, Switzerland, 2019).

Acknowledgements

This work is supported by the National Natural Science Foundation of China (42075147) and the Chinese National Young Thousand Talents Program (41180002). This work thanks the PIs of the OMI/Aura, ERA-5, ACOM-FINN, EDGAR, and MODIS/Landcover products for making their data available.

Author contributions

The first draft of the review was prepared by Jian Liu: Writing - original draft, Data curation, Writing - review and editing, investigation, formal analysis, and visualization. Jason Blake Cohen: conceptualization, data curation, writing - original draft, Writing - review and editing, investigation, validation, and supervision. Qin He: writing - review and editing. Pravash Tiwari: writing - review and editing. Kai Qin: writing - review and editing and project administration.

Competing interests

The authors declare no competing interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-024-01424-5.

Correspondence and requests for materials should be addressed to Jason Blake Cohen.

Peer review information *Communications Earth and Environment* thanks the anonymous reviewers for their contribution to the peer review of this work. Primary Handling Editors: Clare Davis. A peer review file is available.