First measurements of ambient $PM_{2.5}$ in Kinshasa, Democratic Republic of Congo and Brazzaville, Republic of Congo using field-calibrated low-cost sensors

Celeste McFarlane[1,2], Paulson Kasereka Isevulambire[3], Raymond Sinsi Lumbuenamo[4], Arnold Murphy Elouma Ndinga[5], Ranil Dhammapala[6], Xiaomeng Jin[7], V. Faye McNeill[2,8], Carl Malings[9,10], R Subramanian[9,11,12], Daniel M. Westervelt[1,13,*]

[1]Lamont-Doherty Earth Observatory of Columbia University, New York, USA
[2]Columbia University, Department of Chemical Engineering, New York, USA
[3]Ecole Régionale postuniversitaire d'Aménagement et de Gestion Intégrés des Forêts et Territoire tropicaux (ERAIFT) Kinshasa, Democratic Republic of Congo
[4]World Bank Group, Kinshasa, Democratic Republic of Congo
[5]Département de chimie, Université Marien Ngouabi, Brazzaville, Republic of Congo
[6]Washington State Department of Ecology, Washington, USA
[7]Department of Chemistry, University of California Berkeley, USA
[8]Columbia University, Department of Earth and Environmental Sciences, New York, USA
[9]OSU-EFLUVE - Observatoire Sciences de l'Univers-Enveloppes Fluides de la Ville à l'Exobiologie, Université Paris-Est-Créteil, France
[10]NASA Postdoctoral Program Fellow, Goddard Space Flight Center, Greenbelt, Maryland, USA
[11]Center for Atmospheric Particle Studies, Carnegie Mellon University, Pittsburgh, PA 15218, USA
[12]Kigali Collaborative Research Centre, Kigali, Rwanda
[13]NASA Goddard Institute for Space Studies, New York, USA

Abstract

Estimates of air pollution mortality in sub-Saharan Africa are limited by a lack of surface observations of fine particulate matter ($PM_{2.5}$). Despite being large metropolises, Kinshasa, Democratic Republic of the Congo (DRC), population 14.3 million, and Brazzaville, Republic of the Congo (ROC),

[*] *Corresponding author. Tel: 1-845-365-8194
Email address: danielmw@ldeo.columbia.edu

population 2.4 million, have no reference air pollution monitors at the time of writing. Recently, a few reference monitors have been deployed in other parts of sub-Saharan Africa, including Kampala, Uganda. A low-cost PurpleAir $PM_{2.5}$ monitor was collocated next to the Kampala US Embassy BAM-1020 (Met One Beta Attenuation Monitor) starting in August 2019. Raw PurpleAir data are strongly correlated with the BAM ($r^2 = 0.88$), but have a mean absolute error of approximately 14 μg m^{-3}. Two calibration models, multiple linear regression and a random forest approach, decrease mean absolute error (MAE) from 14.3 μg m^{-3} to 3.4 μg m^{-3} or less and improve the the r^2 from 0.88 to 0.96. Given the similarity in climate and emissions in Kampala, we apply the collocated field correction factors to four PurpleAir sensors in Kinshasa, DRC and one in neighboring Brazzaville, ROC deployed beginning April 2018. Annual average $PM_{2.5}$ for 2019 in Kinshasa is estimated at 43.5 μg m^{-3}, more than 4 times higher than WHO Interim Target 1 of 10 μg m^{-3}. Surface $PM_{2.5}$ and aerosol optical depth were each about 40% lower during the 2020 COVID19 lockdown period compared to the same time period in 2019, which cannot be explained by changes in meteorology or wildfire emissions alone. Our results highlight the need for clean air solutions implementation in the Congo.

Keywords: low-cost sensors; particulate matter; air quality; Africa

INTRODUCTION

Ambient air pollution is a major global public health crisis that causes an estimated 4.9 million premature deaths per year around the world. Air pollution is the fifth leading risk factor for all mortality (Health Effects Institute, 2019), and on average reduces life span by 20 months worldwide, rivaling the global impact of cigarette smoking. According to one air quality modeling study, air pollution may cause at least an estimated 780,000 premature deaths annually in Africa (Bauer et al., 2019), and a significant number of diseases (comorbidities) that are known to be worsened by chronic exposure to air pollution, like asthma, lung cancer, and chronic obstructive pulmonary disease (Burnett et al., 2018). However, sparse air pollution monitoring imparts high uncertainty to estimates of exposure and impact. Most of these ambient air pollution deaths come from $PM_{2.5}$, or particulate mass concentrations for particles with diameter less than 2.5 µm.

There is currently no publicly available $PM_{2.5}$ monitoring in the city of Kinshasa, Democratic Republic of Congo or Brazzaville, Republic of Congo, either as regulatory monitors or citizen-deployed low-cost sensors. Therefore, it is currently difficult to know the level of exposure and the potential health impacts of air pollution in Kinshasa and Brazzaville, an alarming gap in a pair of capital megacities containing more than 16 million people and experiencing rapid population growth. To address the lack of data in both cities, we deployed and calibrated a small network of 5 low-cost air quality sensors starting in 2018.

Almost no prior published work has considered $PM_{2.5}$ or air quality in Kinshasa or Brazzaville to this point. Mbelambela et al. (2017) used a portable sensor to calculate personal $PM_{2.5}$, NO_2, and SO_2 exposure in a cohort of 517 subjects, mostly bus drivers. Data were only collected between April 20 and May 14 of 2015 and are more representative of personal exposure rather than ambient concentrations. The authors reported $PM_{2.5}$ exposure over the 3-week period ranging from 64 to 129 µg m^{-3}. The only other air quality study in the peer-reviewed literature in either city was a study of trace metals conducted in 1990 (Lobo et al., 1990), before the Democratic Republic of the Congo was known as such, and has limited relevance to present-day exposures. No studies on air quality in neighboring Brazzaville were found in the literature. Additionally, to our knowledge, there are no national ambient air quality standards in either country.

Low-cost sensors (LCS) have the potential to improve air quality data coverage throughout the world, especially in resource-limited areas (Amegah, 2018). For LCS to provide high quality data, understanding local conditions is vital. Calibration factors and sensor technical performance vary strongly with conditions such as temperature, relative humidity, particle size distribution, and particle loading (Hagan and Kroll, 2020; Levy Zamora et al., 2019; Tryner et al., 2020). Other factors such as sensor aging can also influence sensor performance and have been studied in both a laboratory and field setting (Malings et al., 2019; Tryner et al., 2020). Careful co-location, or side-by-side placement of LCS with reference monitors is an essential step to getting accurate data out of LCS. Several recent studies have performed either a field calibration or a laboratory calibration for a variety of different sensors in a variety of different environmental conditions, paving the way for additional studies in diverse environments (Hagan and Kroll, 2020; Holstius et al., 2014; Jayaratne et al., 2018; Jiao et al., 2016; Johnson et al., 2016; Kelly et al., 2017; Tryner et al., 2020, 2019). However, to our knowledge, published LCS co-location and performance evaluation studies within sub-Saharan Africa are limited (R Subramanian, 2020).

Here we present the first ever multi-year, field-calibrated ambient $PM_{2.5}$ dataset in Kinshasa and Brazzaville. We first build a multiple linear regression model based on a collocation of sensors and reference (Federal Equivalence Method, FEM) $PM_{2.5}$ monitors in Kampala, Uganda and develop a correction factor for low-cost sensors. We then apply the correction to our network in Kinshasa and Brazzaville. We analyze $PM_{2.5}$ on monthly, weekly, daily, and hourly timescales. We interpret the data in the context of changing meteorology and changing human activity coinciding with COVID19-related stay-at-home orders. Finally, we assess the air quality picture in Kinshasa as seen from satellite remote sensing.

METHODS

Description of PurpleAir low-cost sensors

We deploy 5 PurpleAir PA-II-SD low-cost $PM_{2.5}$ devices (www.purpleair.com) in the Kinshasa-Brazzaville area and one PurpleAir device in Kampala for calibration purposes. PurpleAir uses dual Plantower PMS 5003 optical sensors to estimate $PM_{2.5}$ mass concentrations and a Bosch BME280 sensor to estimate temperature, relative humidity, and pressure. Data are transmitted via wireless connectivity in real-time and recorded to an on-board 16GB microSD card. All PurpleAir data is available online on their website. The devices measure sensor readings in six size bins ranging from 300 nm to 10 µm at approximately a 60 second interval. A proprietary algorithm converts raw sensor measurements to PM_1, $PM_{2.5}$, and PM_{10} mass using assumptions about particle shape and density. We use the "CF=ATM" data field as provided by PurpleAir, which is actually the higher "CF=1" data field on account of mislabeled columns by PurpleAir. The CF=1 data has not been transformed nonlinearly and is a better input into regression models. PurpleAir $PM_{2.5}$ is known to strongly correlate ($r > 0.9$) with reference grade monitors but are subject to biases at high relative humidity in particular (Jayaratne et al., 2018; Magi et al., 2020; Malings et al., 2019; Tryner et al., 2020). Sensitivity to relative humidity has also been identified in other low-cost air pollution monitoring devices (Di Antonio et al., 2018; Hagan and Kroll, 2020; He et al., 2020; Jayaratne et al., 2018; Kelleher et al., 2018). PA-II sensors cost approximately $250 USD per unit which are about 100 times cheaper than reference $PM_{2.5}$ monitors, making them attractive for multi-sensor networks. Their use, however, requires careful field calibration in order to achieve high quality data. Malings et al. (2019) developed a multiple linear regression-based calibration method which will be utilized in the following sections.

Sampling locations and periods

Kampala

A PurpleAir was collocated at the US Embassy in Kampala, located at Plot 1577 Ggaba Road, about 3 kilometers from the city center, 0.301268 N latitude and 32.591711 E longitude. Sampling began in September 2019 and continues through the present. Also located at the US Embassy is a MetOne Beta Attenuation Monitor (BAM) 1020, which provides a calibration point for the PurpleAir. Data collection of $PM_{2.5}$ with the BAM-1020 has been ongoing at the US Embassy since January 2017 (https://www.airnow.gov/international/us-embassies-and-consulates/, last accessed 26 August 2020). We use the 2019 September through February overlapping dataset with the BAM-1020 and PurpleAir for our calibration (see the Field Calibration Section of Methods).

Kinshasa and Brazzaville

Four PurpleAir sensors were deployed throughout Kinshasa with a fifth in Brazzaville. Starting in March 2018 and continuing through present, a device was located at the US Embassy in Kinshasa at 310 Avenue des Aviateurs, latitude 4.3002 S and longitude 15.3138 E. In November of 2019, sensors were added at L'université pédagogique nationale (UPN) located on Route de Matadi, latitude 4.4039 S and longitude 15.2572 E, and L'Ecole Régionale Postuniversitaire d'Aménagement et de Gestion Intégrés des Forêts et Territoires tropicaux (ERAIFT), latitude 4.4103 S and 15.3065 E. A fourth sensor was located in a residential area in the Kintambo area of Kinshasa at the Belle Vue Villas (abbrieviated CBV), latitude 4.3278 S and longitude 15.2722 E, starting in November 2019. Finally, a fifth sensor was located across the Congo River in Brazzaville, Republic of Congo, at the US Embassy in Brazzaville at latitude 4.2751 S and longitude 15.2561 E, starting in February 2020. A map of all sites is shown in Fig. 1.

Field calibration

Multiple Linear Regression (MLR)

We use both a multiple linear regression and a random forest approach for bias-correcting the PurpleAir data towards the reference monitor in Kampala. The collocation of the PurpleAir and the FEM monitor took place between September 2019 and March 2020. A randomly selected 75% of data during that time period is used to build the multiple linear regression model, and the remaining 25% is used for validation. This model was developed using the base R statistics package. The multiple linear regression approach follows a similar methodology as in Malings et al. (2019) in which daily-averaged raw PurpleAir $PM_{2.5}$, relative humidity, and temperature are used as explanatory variables to predict corrected $PM_{2.5}$ concentration:

$$PM_{2.5} = \beta_0 + \beta_1 \times purpleairPM_{2.5} + \beta_2 \times T(°C) + \beta_3 \times RH(\%) \quad (1)$$

The multiple linear regression model was evaluated based on coefficients of determination (r^2) and Mean Absolute Error (MAE) defined as:

$$MAE = \frac{\sum_{i=1}^{i=N} \left| PM_i^{ref} - PM_i^{LCS} \right|}{N} \quad (2)$$

where PM refers to $PM_{2.5}$ (in µg m^{-3}), "ref" is the BAM-1020 reference monitor, "LCS" refers to the low-cost sensor data, N is total number of observations, and i is the timeseries variable. Since the MLR and RF methods result in similar correlation and mean bias improvements when the same averaging time periods are used and the MLR approach is more transparent (see Results and Discussion), we present further calibrated $PM_{2.5}$ results based on MLR. Results obtained using the RF approach are not substantially different so they are presented in the Supplementary Information. The results of the MLR model correction are shown results subsection "Kampala co-location analysis".

Satellite data

We use remote sensing data from the NASA Aqua and Terra satellites, which contain the Moderate Resolution Imaging Spectroradiometer (MODIS). MODIS views the Earth surface over the Congo region once per day with a daily overpass time around 12:30pm. We specifically use the Aerosol Optical Depth (AOD) at 550 nm wavelength using the Multi-Angle Implementation of Atmospheric Correction (MAIAC) Level 2 gridded data over land surfaces at 1 km pixel resolution (short name MCD19A2) (Lyapustin et al., 2018). We convert the pixel data in the H19V9 tile to latitude and longitude points using the MODLAND Tile Calculator (https://landweb.modaps.eosdis.nasa.gov/cgi-bin/developer/tilemap.cgi, last accessed 25 Aug 2020). Cloudy pixel data are removed as part of this data product, and only "best quality" flagged data are used in our analysis. Data are available on the NASA Earth Data repository (https://earthdata.nasa.gov/, last accessed 25 Aug 2020). We use daily MAIAC data from January 2018 through July 2020 in our analysis.

Meteorological data

We use the NOAA National Centers for Environmental Information land-based station data for meteorology. In Kinshasa and Brazzaville, the most complete temperature records are at the N'Djili International Airport (Kinshasa) and Maya-Maya International Airport (Brazzaville). Temperature, relative humidity, and wind speed, is mostly a complete record at these locations. Greater than 58% of precipitation data between September 2017 and August 2020 are missing at each of these stations, rendering any precipitation analysis impossible. We therefore use temperature, relative humidity, and wind speed data from N'Djili Airport to analyze the impact of weather on $PM_{2.5}$ observations.

RESULTS AND DISCUSSION

Kampala co-location analysis

In Table 1 we present the regression parameters for the Kampala co-location. Correlation between raw PurpleAir data and Embassy BAM-1020 is high (r^2 = 0.88). The overall mean absolute error (MAE) in the raw data compared to the Embassy BAM is 14.8 μg m^{-3} which is partially driven by a large overprediction during times of very poor air quality (values greater than 100 μg m^{-3}). Values less than 100 μg m^{-3} show less bias compared to the BAM-1020. Using the MLR model to correct the raw $PM_{2.5}$ data towards FEM standard using the BAM-1020 results in a reduction of MAE to 3.4 μg m^{-3}, quantified using only the remaining 25% of the collocated data reserved for testing and validation. The corrected PurpleAir $PM_{2.5}$ data (turquoise line in Fig. 2) shows a substantial improvement in comparison against BAM $PM_{2.5}$. Correlation is also improved by the MLR model (r^2 = 0.96). Though hourly-averaged $PM_{2.5}$ data from the PurpleAir and the BAM-1020 are noisy and therefore not as well suited for MLR, we also build an MLR model for the hourly-averaged $PM_{2.5}$ data, which we only apply to any analysis that uses hourly mean data instead of daily mean data. Table 2 also contains the hourly-based MLR statistics.

We apply the correction factors developed from the long term MLR collocation in Kampala to each of the 5 PurpleAir $PM_{2.5}$ monitors deployed in Kinshasa and Brazzaville. Kampala is just under 2000 km away from Kinshasa; however, this is the closest reference $PM_{2.5}$ monitor to Kinshasa. The climates of Kinshasa and Kampala are quite similar as both lie within the center of the tropics. Both cities have two wet seasons that peak around October-November and March-April. Unlike Kampala, Kinshasa has a true dry season in which there is no or very little precipitation. Kampala's "drier" season coincides with Kinshasa's (June-August) dry season. Annual mean temperature and relative humidity are 30.4 ºC and 80% in Kinshasa and 27.8 ºC and 75% in Kampala. Differences in emissions characteristics in the two cities are also a possible influence on applicability of our calibration. Thus, we also analyzed emissions data from the Diffuse and Inefficient Combustion Emissions Inventory in Africa (DICE-Africa), comparing source profiles for SO_2, BC, and OC emissions in a 25 km by 25 km grid boundary in each city. Results (Fig. S1) indicate that the source profiles have many similarities in the two cities, in particular household fuel usage which comprises about 50-75% of the total emissions of SO_2, black carbon (BC), and organic carbon (OC) in each city. Though the total emissions are slightly higher in Kinshasa compared to Kampala, the similar source mix further justifies applicability of sensor calibrations between the two cities. Qualitative information about particle size in the two cities can be obtained through satellite retrievals of the Angstrom exponent, which we present in Fig. S2. The 2019 annual mean Angstrom exponent as retrieved by MODIS Terra Deep Blue retrieval algorithm is about 1.5 in Kampala and 1.6 in Kinshasa, indicating qualitative similarity in particle size distributions in each city. Though the geographic distance between Kinshasa and Kampala is a limitation of the work, given the similarities between the two environments, estimates of emissions sources, qualitative similarity in particle size, and the paucity of reference monitoring and resources for such monitoring in sub-Saharan Africa, we consider our sub-continental calibration to be sufficient for further analysis.

$PM_{2.5}$ time series at each sampling location

Figure 3 shows both the corrected (turquoise) and raw (purple) monthly averaged PurpleAir $PM_{2.5}$ data collected at the US Embassy in Kinshasa (see Methods Section and Fig. 1) from March of 2018 through July 2020. Often the calibrated data and raw data are within 10 µg m^{-3} of each other, except when the $PM_{2.5}$ is above 100 µg m^{-3} or below 25 µg m^{-3}. The 95% confidence interval for daily averaged raw data ranges between ±0.1 to ±0.5 μg m^{-3}, depending on the exact month in the data. $PM_{2.5}$ data are nearly always above WHO air quality guidelines (10 μg m^{-3} on an annual mean basis, 25 μg m^{-3} on a 24-hour mean basis). The 2018 average $PM_{2.5}$ was 54.4 μg m^{-3}, which was followed by a decrease in 2019 (43.5 μg m^{-3}). Through July 2020 (time of last data collection), the average is 35.7 μg m^{-3}, though $PM_{2.5}$ is typically high in August and later in the year. The 2018 average also does not include the months of January and February, which are typically lower in $PM_{2.5}$. In addition to this, the decrease in $PM_{2.5}$ between 2018 and 2019 is not expected to be related to emissions reductions. Emissions trends in recent years for DRC are not well known. However, to our knowledge there has been no emissions control mitigation or measures in the last 3 years. Meteorological data are also limited in Kinshasa and Brazzaville. In Fig. 3 we also plot the Kinshasa station available meteorological variables from January of 2018 to August 2020. To first order, several previous studies have reported that on a large-scale, $PM_{2.5}$ can be positively correlated with temperature, negatively correlated with wind speed, and positively correlated with relative humidity (Fiore et al., 2015; Jacob and Winner, 2009; Westervelt et al., 2016). As seen in Fig. 3, dry season July 2019 temperatures are approximately the same between 2018, 2019, and 2020, suggesting that the temperature effects on $PM_{2.5}$ cannot explain the observed decrease in $PM_{2.5}$ in 2020. Relative humidity in 2020 is slightly higher or roughly the same compared to either 2019 or 2018, inconsistent with the strong 2020 $PM_{2.5}$ decreases. Wind speed is slightly lower in 2020 compared to 2018 or 2019, which would also have the effect of likely increasing $PM_{2.5}$, and therefore cannot explain the observed decreases in $PM_{2.5}$. We conclude that the decrease in $PM_{2.5}$ between 2018, 2019, and 2020 cannot be directly attributed to at least these three meteorological factors. Another plausible explanation is a potential decrease in wildfire activity in the Congolese rainforest. Burned area was shown to decline by ~1.3% yr^{-1} between 2003 and 2017 in the Central African Republic and South Sudan (Jiang et al., 2020), thousands of kilometers from Kinshasa. Closer to Kinshasa, burned area was found to largely remain unchanged over the same time period. Given these contrasting results, it is not likely that fire activity is the main driver of observed $PM_{2.5}$ decreases either, although it may contribute partially. We will later attribute the 2020 decrease at least partially to decreases in activity associated with COVID19 lockdown.

In Fig. 4 we plot the monthly averaged corrected $PM_{2.5}$ for each of the four other sites (see Fig. 1). Data collection began at 3 of these 4 sites in November of 2019 (UPN, ERAIFT, and Cite Belle Vue), and in February of 2020 at the Brazzaville US Embassy, continuing through September of 2020. Each of the sites show a similar seasonality, with lower concentrations in November 2019 through April 2020, coinciding with the rainy season. $PM_{2.5}$ concentrations in June and July are higher at 60-70 µg m^{-3} from 40 µg m^{-3} in the rainy season. This is qualitatively consistent with the typical seasonal behavior at the longer-term US Embassy Kinshasa location, though the dry season rebound in $PM_{2.5}$ in 2020 at the Kinshasa embassy is smaller than at the other 4 sites. As with the Kinshasa embassy site, corrected $PM_{2.5}$ levels are nearly always above WHO guidelines, even in the rainy season.

Diurnal and weekly profiles at each sampling site

We present in Fig. 5 the $PM_{2.5}$ over the diurnal cycle in each of the Kinshasa-Brazzaville locations using the entire datasets available for each site. Each site tends to have similar diurnal variability, though the magnitude of the $PM_{2.5}$ concentrations is different at each location. There is typically a very early morning minimum in the diurnal trend at all sites, around 5:00 West Africa Time (WAT), followed by a morning peak at around 8:00. Local traffic emissions and household cooking are expected to be a major contributor to these peaks. Except at UPN-DRC, there is a steady increase throughout the course of the day, as local activity such as cooking and preparing of food, open burning, and other sources begin to contribute to $PM_{2.5}$. The highest $PM_{2.5}$ values of the day then peak at about 18:00 - 20:00, coinciding with evening vehicle traffic and cooking. After 20:00, $PM_{2.5}$ values start to drop quickly overnight. The highest $PM_{2.5}$ values occur at Brazzaville Embassy and UPN-DRC, peaking at around 60 μg m^{-3} in the evening. The lowest $PM_{2.5}$ values occur at Cité Belle Vue, which is a higher-end residential area where some foreign diplomats tend to live. These areas are less likely to use high emitting cooking and open burning practices, potentially explaining the lower $PM_{2.5}$. However, the $PM_{2.5}$ data at Cité Belle Vue is also only available for the mostly wet season, which can also potentially explain the lower $PM_{2.5}$ values. The 5-site average is shown in black in Fig. 5 and ranges from about 40 μg m^{-3} at minimum to above 50 μg m^{-3} at the peak.

Average $PM_{2.5}$ concentration by day of the week is plotted for each site and the area average in Fig. 6. Generally, there are substantial differences in the weekly $PM_{2.5}$ variation between each location. As in the diurnal variation (Fig. 5), Brazzaville and UPN-DRC have the highest $PM_{2.5}$ of all the sites, around 50-55 μg m^{-3}. CBV on average is about 10 μg m^{-3} lower, though this is likely due to a limited data range that mainly includes the wet season. Sunday is generally one of the lowest days for $PM_{2.5}$ in all sites (except CBV), coinciding with a low point in economic activity during the week. Each site has typically two days of the week in which $PM_{2.5}$ peaks, though which day varies by site. At UPN-DRC, there is an early week peak Wednesday and then a lower Saturday peak, with a minimum on Thursday. Conversely, Kinshasa Embassy and Brazzaville Embassy are highest on Thursday, and ERAIFT-DRC is highest on Friday. This variation among weekday $PM_{2.5}$ is likely explained by variation in type of location. Residential, educational, and diplomatic locations are represented among the 5 sites, each of which have different source signatures, emissions patterns, and micrometeorology.

Figure 7 summarizes all daily mean data at all 5 sites and the 5-site average in a violin plot. Distributions of $PM_{2.5}$ are mostly unimodal and right-skewed, with the majority of data points located around 50 µg m^{-3}. As previously discussed, UPN-DRC and Brazzaville Embassy have the highest median and quartile ranges for $PM_{2.5}$. The highest daily mean extreme value (nearly 180 µg m^{-3}) is observed at the Kinshasa Embassy. However, Kinshasa embassy has the most extreme values with several daily means greater than 100 µg m^{-3}, though it also represents the longest data record length. The $PM_{2.5}$ distributions at every site except for Cité Belle Vue each have "long tails" with nonzero density as concentrations approach 100 µg m^{-3}, indicating a higher frequency of poor air quality at these locations compared to other sites. The 5-site average median daily $PM_{2.5}$ value is 42.1 µg m^{-3}, a factor of 4 higher than the WHO annual mean guideline of 10 µg m^{-3}.

Impact of COVID-19 lockdown on $PM_{2.5}$ in Kinshasa

We use our 29-month dataset (beginning March 2018) at the Kinshasa Embassy to assess the potential changes in $PM_{2.5}$ concentrations attributable to COVID-19 related stay-at-home orders in Kinshasa. The Governor of Kinshasa announced a total lockdown for the city starting on March 28, which was later postponed. Starting April 6, Gombe, the administrative and commercial center of Kinshasa where the US Embassy is located, was closed for two weeks. We plot and compare the April for 2019 and 2020 in Fig. 8. On average, April 2020 daily mean $PM_{2.5}$ is about 14.7 µg m^{-3} lower than the same time period in 2019, about a 40% decrease, larger than the amount of $PM_{2.5}$ reduction in China during COVID-19 quarantine orders (Shi and Brasseur, 2020). In particular, the evening peak around 20:00 seen clearly in 2019 is mostly absent in 2020. Other than a flattening of the evening peak, the 2020 trend follows 2019 very closely, though offset by at least 10 µg m^{-3}. Natural variability including fire activity may play a role in these decreases, though our preliminary meteorological analysis (see Fig. 3 and associated discussion) suggests that meteorology cannot explain the observed $PM_{2.5}$ decreases. Daily mean $PM_{2.5}$ in April 2018 was 3.8 μg m^{-3} lower compared to April 2019, indicating that there was not an April $PM_{2.5}$ decreasing trend prior to the COVID-19 pandemic.

We also analyze satellite observations of AOD at 550 nm wavelength using the MAIAC Level 2 gridded data over land surfaces at 1 km pixel resolution. Figure 9 shows MAIAC AOD for January-June 2018, 2019, and 2020, and a 2020 versus 2019 difference over the Congo region. AOD levels are elevated over the larger region and especially over Kinshasa and Brazzaville (located around 15.3 E latitude and 4.3 S latitude). Contrary to the surface level $PM_{2.5}$ data, AOD was higher in January through July 2019 than 2018. Columnar AOD data is a very different measure than surface level $PM_{2.5}$ and do not necessarily vary together (Li et al., 2015; Van Donkelaar et al., 2016). Compared to 2019, AOD in 2020 is about 0.1-0.2 lower than 2019, but only 0.05 lower than 2018. These decreases are qualitatively consistent with our hypothesis of COVID19 lockdown impacts on air pollution in the Congo region, though could also be consistent with changes in fire activity. Malings, Westervelt, et al. (2020) further explored connections between MAIAC AOD and surface $PM_{2.5}$ in both Africa and North America, including applications for converting column AOD to surface $PM_{2.5}$.

CONCLUSIONS

Air pollution causes millions of premature deaths and a host of illnesses in the world's cities each year. This is especially true in sub-Saharan Africa, where sparse air pollution monitoring imparts high uncertainty to estimates of exposure and impact. Together, the Kinshasa-Brazzaville megalopolis is home to about 14 million inhabitants, and yet long-term, publicly available, systematic air quality data did not exist prior to this work. We developed a correction factor built from a MLR model developed in Kampala, a city of similar environmental conditions, and deployed 5 low-cost sensor in the Kinshasa-Brazzaville area, starting in March 2018. Study-average network-wide $PM_{2.5}$ is 46.1 μg m^{-3}, indicating severely impaired air quality in the two cities. $PM_{2.5}$ is highest in the dry season of June, July, and August, and lower in the remaining months where wet scavenging plays a dominant role. Wet season $PM_{2.5}$ is about 5-10 μg m^{-3} lower on average than dry season, yet still four times higher than the WHO guideline. Decreases in $PM_{2.5}$ between 2018 and 2020 cannot be explained by changing meteorology; however, a paucity of available data limits our analysis. $PM_{2.5}$ varies by site to some extent between our 5 sites, though average values are within 10 μg m^{-3} of each other, suggesting a coherent set of sampling locations indicative of average city-wide conditions. $PM_{2.5}$ is generally highest around 20:00 WAT, corresponding to an early evening activity peak. Likewise, $PM_{2.5}$ is lowest on Sundays, when activity is limited. During the 2020 COVID19 lockdown, $PM_{2.5}$ decreased by 14.7 μg m^{-3} compared to an identical time period in 2019, which cannot alone be explained by changes in meteorology. Satellite observations of aerosol optical depth qualitatively confirm our surface findings.

This work represents a first step at understanding air quality in a fast-growing megacity in sub-Saharan Africa. Key limitations of the work include the lack of a reference grade monitor locally in Kinshasa, lack of robust weather data, and limited time periods of $PM_{2.5}$ data. Future work should address these issues, and also explore the potential for air quality models and satellite observations to be better adapted to use over Kinshasa and the Congo.

ACKNOWLEDGEMENTS

CMM acknowledges a SEAS summer internship and an Earth Institute Student Research Fellowship. DMW acknowledges a Columbia University Climate and Life Fellowship and a Columbia University Climate Center grant. RS acknowledges State assistance managed by the National Research Agency under the "Programme d'Investissements d'Avenir" under the reference "ANR-18-MPGA-0011" ("Make our planet great again" initiative), as well as equipment support from PurpleAir, LLC. VFM acknowledges the Columbia University Earth Institute for support. CM acknowledges support by an appointment to the NASA Postdoctoral Program at the Goddard Space Flight Center, administered by USRA through a contract with NASA. We acknowledge the following US State Department diplomats: Jeremy Chen, Dovas Saulys, Lindy Nester, and Thomas Hilleary in Kinshasa and David Kelm in Brazzaville. We acknowledge Paul Green, Engineer Bainomugisha, and the AirQo Kampala project for assistance in retrieving data. The findings of this paper do not represent the views of the United States Department of State and are solely the views of the authors. CMM and DMW created figures, led the data analysis and wrote the manuscript. DMW conceived the project and provided equipment. RS provided equipment. PK and RSL provided sensor support. RD provided data quality control and assurance guidance. VFM provided equipment and other guidance. All authors contributed to editing the paper. All low-cost data are publicly available on the PurpleAir website. (https://www.purpleair.com/sensorlist). Regulatory instrument data from the US Embassies are available through the US AirNow Department of State data website (https://www.airnow.gov/international/us-embassies-and-consulates/). Satellite data are available from NASA's EarthData website (https://earthdata.nasa.gov/).

REFERENCES

Amegah, A.K. (2018). Proliferation of low-cost sensors. What prospects for air pollution epidemiologic research in Sub-Saharan Africa? Environmental Pollution. https://doi.org/10.1016/j.envpol.2018.06.044

Bauer, S.E., Im, U., Mezuman, K., Gao, C.Y. (2019). Desert Dust, Industrialization, and

Agricultural Fires: Health Impacts of Outdoor Air Pollution in Africa. J. Geophys. Res. Atmos. 124, 4104–4120. https://doi.org/10.1029/2018JD029336

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C.A., Apte, J.S., Brauer, M., Cohen, A., Weichenthal, S., Coggins, J., Di, Q., Brunekreef, B., Frostad, J., Lim, S.S., Kan, H., Walker, K.D., Thurston, G.D., Hayes, R.B., Lim, C.C., et al. (2018). Global estimates of mortality associated with longterm exposure to outdoor fine particulate matter. Proc. Natl. Acad. Sci. U. S. A. 115, 9592–9597. https://doi.org/10.1073/pnas.1803222115

Di Antonio, A., Popoola, O., Ouyang, B., Saffell, J., Jones, R. (2018). Developing a Relative Humidity Correction for Low-Cost Sensors Measuring Ambient Particulate Matter. Sensors 18, 2790. https://doi.org/10.3390/s18092790

Fiore, A.M., Naik, V., Leibensperger, E.M. (2015). Air Quality and Climate Connections. J. Air Waste Manage. Assoc. 65, 645–685. https://doi.org/10.1080/10962247.2015.1040526

Hagan, D.H., Kroll, J.H. (2020). Assessing the accuracy of low-cost optical particle sensors using a physics-based approach. Atmos. Meas. Tech. 13, 6343–6355. https://doi.org/10.5194/amt-13-6343-2020

He, M., Kuerbanjiang, N., Dhaniyala, S. (2020). Performance characteristics of the low-cost Plantower PMS optical sensor. Aerosol Sci. Technol. 54, 232–241. https://doi.org/10.1080/02786826.2019.1696015

Heatlth Effects Institute | State of Global Air (2020). URL https://www.stateofglobalair.org/report (accessed 26 August 2020).

Holstius, D.M., Pillarisetti, A., Smith, K.R., Seto, E. (2014). Field calibrations of a low-cost aerosol sensor at a regulatory monitoring site in California. Atmos. Meas. Tech. 7, 1121–1131. https://doi.org/10.5194/amt-7-1121-2014

Jacob, D.J., Winner, D. a. (2009). Effect of climate change on air quality. Atmos. Environ. 43, 51–63. https://doi.org/10.1016/j.atmosenv.2008.09.051

Jayaratne, R., Liu, X., Thai, P., Dunbabin, M., Morawska, L. (2018). The influence of humidity on the performance of a low-cost air particle mass sensor and the effect of atmospheric fog. Atmos. Meas. Tech. 11, 4883–4890. https://doi.org/10.5194/amt-11-4883-2018

Jiang, Y., Zhou, L., Raghavendra, A. (2020). Observed changes in fire patterns and possible drivers over Central Africa. Environ. Res. Lett. 15, 0940b8. https://doi.org/10.1088/1748-9326/ab9db2

Jiao, W., Hagler, G., Williams, R., Sharpe, R., Brown, R., Garver, D., Judge, R., Caudill, M., Rickard, J., Davis, M., Weinstock, L., Zimmer-Dauphinee, S., Buckley, K. (2016). Community Air Sensor Network (CAIRSENSE) project: Evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. Atmos. Meas. Tech. 9, 5281–5292. https://doi.org/10.5194/amt-9-5281-2016

Johnson, K.K., Bergin, M.H., Russell, A.G., Hagler, G.S.W. (2016). Using Low Cost Sensors to Measure Ambient Particulate Matter Concentrations and On-Road Emissions Factors. Atmos. Meas. Tech. Discuss. 1–22. https://doi.org/10.5194/amt-2015-331

Kelleher, S., Quinn, C., Miller-Lionberg, D., Volckens, J. (2018). A low-cost particulate matter (PM_{2.5}) monitor for wildland fire smoke. Atmos. Meas. Tech. 11, 1087–1097. https://doi.org/10.5194/amt-11-1087-2018

Kelly, K.E., Whitaker, J., Petty, A., Widmer, C., Dybwad, A., Sleeth, D., Martin, R., Butterfield, A. (2017). Ambient and laboratory evaluation of a low-cost particulate matter sensor. Environ. Pollut. 221, 491–500. https://doi.org/10.1016/j.envpol.2016.12.039

Levy Zamora, M., Xiong, F., Gentner, D., Kerkez, B., Kohrman-Glaser, J., Koehler, K. (2019). Field and Laboratory Evaluations of the Low-Cost Plantower Particulate Matter Sensor.

Environ. Sci. Technol. 53, 838–849. https://doi.org/10.1021/acs.est.8b05174

Li, J., Carlson, B.E., Lacis, A.A. (2015). How well do satellite AOD observations represent the spatial and temporal variability of PM2.5 concentration for the United States? Atmos. Environ. 102, 260–273. https://doi.org/10.1016/j.atmosenv.2014.12.010

Lobo, K.K., Tshiashala, M.D., Lumu, B.M., Matamba, K., Ronneau, C. (1990). Air Quality Status in Kinshasa as Determined by Instrumental Neutron Activation Analysis, Atomic Absorption Spectrometry, and Ion-Exchange Chromatography, in: Nuclear Analytical Methods in the Life Sciences, Humana Press, pp. 295–305. https://doi.org/10.1007/978-1-4612-0473-2_33

Lyapustin, A., Wang, Y., Korkin, S., Huang, D. (2018). MODIS Collection 6 MAIAC algorithm. Atmos. Meas. Tech. 11, 5741–5765. https://doi.org/10.5194/amt-11-5741-2018

Magi, B.I., Cupini, C., Francis, J., Green, M., Hauser, C. (2020). Evaluation of PM2.5 measured in an urban setting using a low-cost optical particle counter and a Federal Equivalent Method Beta Attenuation Monitor. Aerosol Sci. Technol. 54, 147–159. https://doi.org/10.1080/02786826.2019.1619915

Malings, C., Tanzer, R., Hauryliuk, A., Kumar, S.P.N., Zimmerman, N., Kara, L.B., Presto, A.A. (2019). Development of a general calibration model and long-term performance evaluation of low-cost sensors for air pollutant gas monitoring. Atmos. Meas. Tech. 12, 903–920. https://doi.org/10.5194/amt-12-903-2019

Malings, C., Tanzer, R., Hauryliuk, A., Saha, P.K., Robinson, A.L., Presto, A.A., Subramanian, R. (2020a). Fine particle mass monitoring with low-cost sensors: Corrections and long-term performance evaluation. Aerosol Sci. Technol. 54, 160–174. https://doi.org/10.1080/02786826.2019.1623863

Malings, C., Westervelt, D.M., Hauryliuk, A., Presto, A.A., Grieshop, A., Bittner, A., Beekmann,

M. (2020b). Application of low-cost fine particulate mass monitors to convert satellite aerosol optical depth to surface concentrations in North America and Africa. Atmos. Meas. Tech. 13, 3873–3892. https://doi.org/10.5194/amt-13-3873-2020

Mbelambela, E.P., Hirota, R., Eitoku, M., Muchanga, S.M.J., Kiyosawa, H., Yasumitsu-Lovell, K., Lawanga, O.L., Suganuma, N. (2017). Occupation exposed to road-traffic emissions and respiratory health among Congolese transit workers, particularly bus conductors, in Kinshasa: A cross-sectional study. Environ. Health Prev. Med. 22, 11. https://doi.org/10.1186/s12199-017-0608-9

R Subramanian (2020). Air pollution in Kigali, Rwanda: spatial and temporal variability, source contributions, and the impact of car-free Sundays. Clean Air J. 30. https://doi.org/10.17159/caj/2020/30/2.8023

Shi, X., Brasseur, G.P. (2020). The Response in Air Quality to the Reduction of Chinese Economic Activities During the COVID-19 Outbreak. Geophys. Res. Lett. 47. https://doi.org/10.1029/2020GL088070

Tryner, J., L’Orange, C., Mehaffy, J., Miller-Lionberg, D., Hofstetter, J.C., Wilson, A., Volckens, J. (2020). Laboratory evaluation of low-cost PurpleAir PM monitors and in-field correction using co-located portable filter samplers. Atmos. Environ. 220, 117067. https://doi.org/10.1016/j.atmosenv.2019.117067

Tryner, J., Quinn, C., Windom, B.C., Volckens, J. (2019). Design and evaluation of a portable PM2.5 monitor featuring a low-cost sensor in line with an active filter sampler. Environ. Sci. Process. Impacts 21, 1403–1415. https://doi.org/10.1039/c9em00234k

Van Donkelaar, A., Martin, R. V., Brauer, M., Hsu, N.C., Kahn, R.A., Levy, R.C., Lyapustin, A., Sayer, A.M., Winker, D.M. (2016). Global Estimates of Fine Particulate Matter using a

Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. Environ. Sci. Technol. 50, 3762–3772. https://doi.org/10.1021/acs.est.5b05833

Westervelt, D.M., Horowitz, L.W., Naik, V., Tai, A.P.K., Fiore, A.M., Mauzerall, D.L. (2016). Quantifying PM<inf>2.5</inf>-meteorology sensitivities in a global climate model. Atmos. Environ. 142. https://doi.org/10.1016/j.atmosenv.2016.07.040

TABLES

Table 1: Regression coefficients of the MLR model

	Daily model	**Hourly model**
β_0	64.7	86.9
β_1	0.52	0.55
β_2	-0.23	-1.28
β_3	-0.59	-0.55

Table 2: Statistics of the MLR and RF models

Model	**Averaging time period**	**R^2**	**Mean Absolute Error ($\mu g/m^3$)**
Raw PurpleAir Data	Daily	0.88	14.8
	Hourly	0.88	20.3
Multiple Linear Regression	Daily	0.96	3.4
	Hourly	0.90	7.3

Random Forest	Daily	0.86	5.8
	Hourly	0.91	7.2

Figure Captions

Figure 1. Map of sensor locations in Kinshasa and Brazzaville. Background map © Google 2020.

Figure 2. Performance evaluation and calibration of PurpleAir $PM_{2.5}$ versus Federal Equivalent Method (FEM) $PM_{2.5}$ between September 2019 and February 2020 at the US Embassy in Kampala, Uganda. Raw daily data is shown in purple, FEM data in orange, and corrected low-cost sensor data (using the MLR method) in turquoise.

Figure 3. Weekly mean raw (purple) and corrected (turquoise) $PM_{2.5}$ data, as well as weekly mean temperature (ºC), wind speed (m s^{-1}), and relative humidity (%) (NOAA data) at the Kinshasa US Embassy site between March 2018 and July 2020.

Figure 4. Monthly mean corrected data at 4 sites (excluding Kinshasa Embassy) between November 2019 and July 2020. Shaded areas indicate the 95% confidence interval of the monthly average values.

Figure 5. Diurnal average $PM_{2.5}$ for the entire data record at each of the 5 sites and the site-wide average (black). Shaded areas indicate the 95% confidence interval of the hourly averages. Hour indicates the local time (West Africa Time).

Figure 6. Day-of-the-week averages for $PM_{2.5}$ for the entire dataset at each of the 5 sites, and the site-wide average. Shaded areas indicate the 95% confidence interval of the daily averages.

Figure 7. Violin plot of daily-averaged $PM_{2.5}$ values for the entire dataset at each location and for the site-wide average. Boxes represent median and inter-quartile range.

Figure 8. Analysis of $PM_{2.5}$ changes during COVID19 at the US Embassy Kinshasa location. Averaging time period is April 6 through April 20. (Top) Diurnal $PM_{2.5}$ in 2019 and 2020 by day of the week. (Bottom left) Diurnal mean $PM_{2.5}$ (Bottom middle) Monthly mean $PM_{2.5}$ (Bottom right) Day of the week average. Shaded region represents the 95% confidence interval in the mean.

Figure 9. MAIAC Level 2 550 nm AOD at 1km resolution over the Congo region for a January through July average in (a) 2018, (b) 2019, (c) 2020, and (d) the difference between 2019 and 2020.

Boundaries represent level 2 administrative boundaries (districts and communes) provided by the United Nations Office for the Coordination of Humanitarian Affairs. The black dashed box indicates the Kinshasa-Brazzaville region and the inset in panel (d) shows the broader location within the African continent.

FIGURES

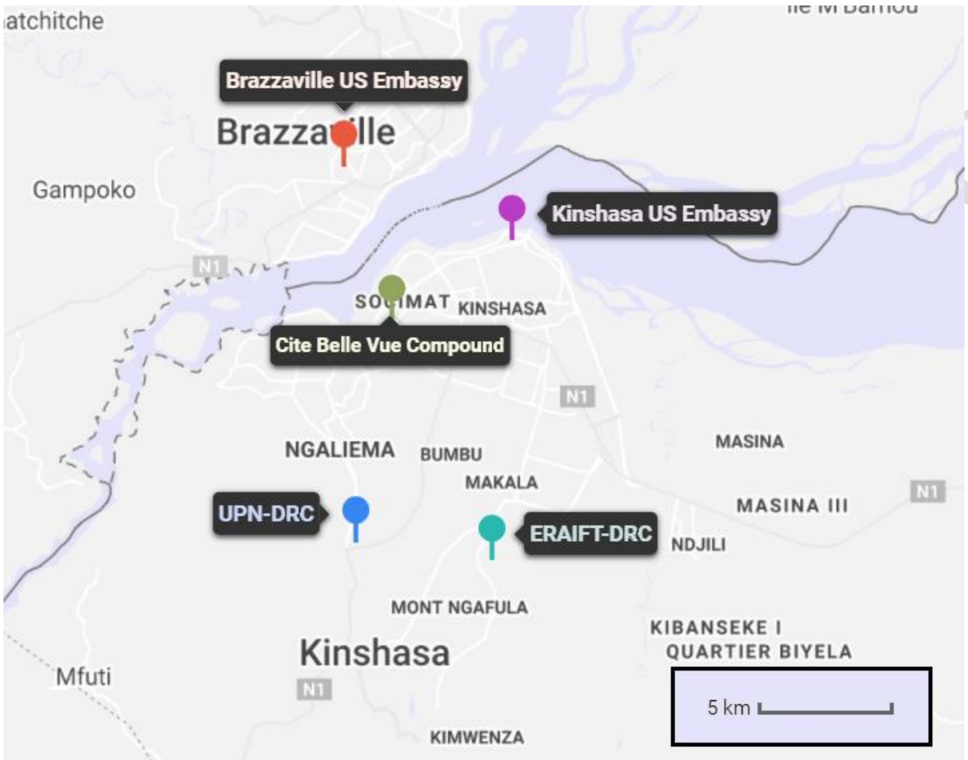


Figure 1. Map of sensor locations in Kinshasa and Brazzaville. Background map © Google 2020.

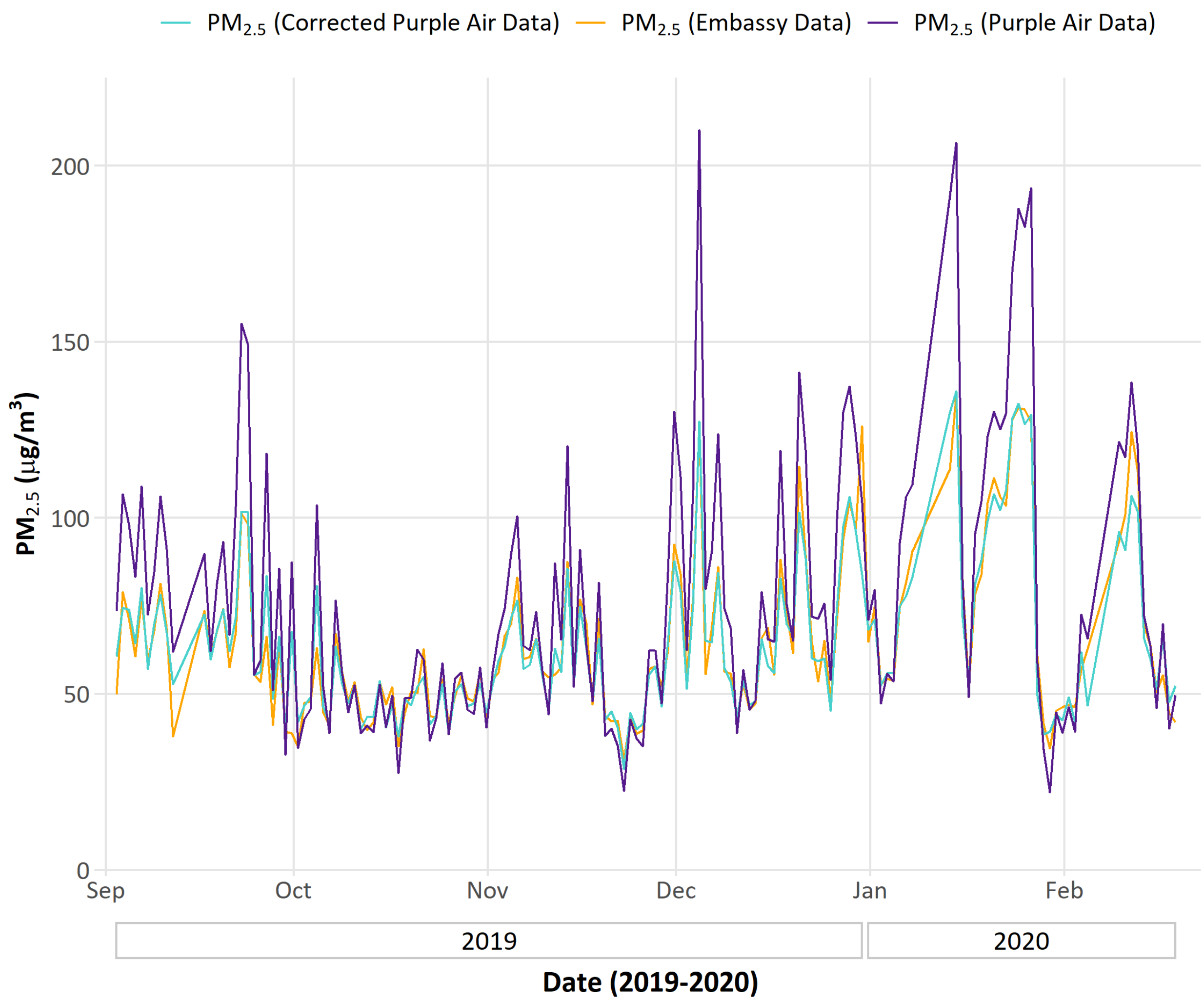


Figure 2. Performance evaluation and calibration of PurpleAir $PM_{2.5}$ versus Federal Equivalent Method (FEM) $PM_{2.5}$ between September 2019 and February 2020 at the US Embassy in Kampala, Uganda. Raw daily data is shown in purple, FEM data in orange, and corrected low-cost sensor data (using the MLR method) in turquoise.

Figure 3. Weekly mean raw (purple) and corrected (turquoise) $PM_{2.5}$ data, as well as weekly mean temperature (°C), wind speed (m s^{-1}), and relative humidity (%) (NOAA data) at the Kinshasa US Embassy site between March 2018 and July 2020.

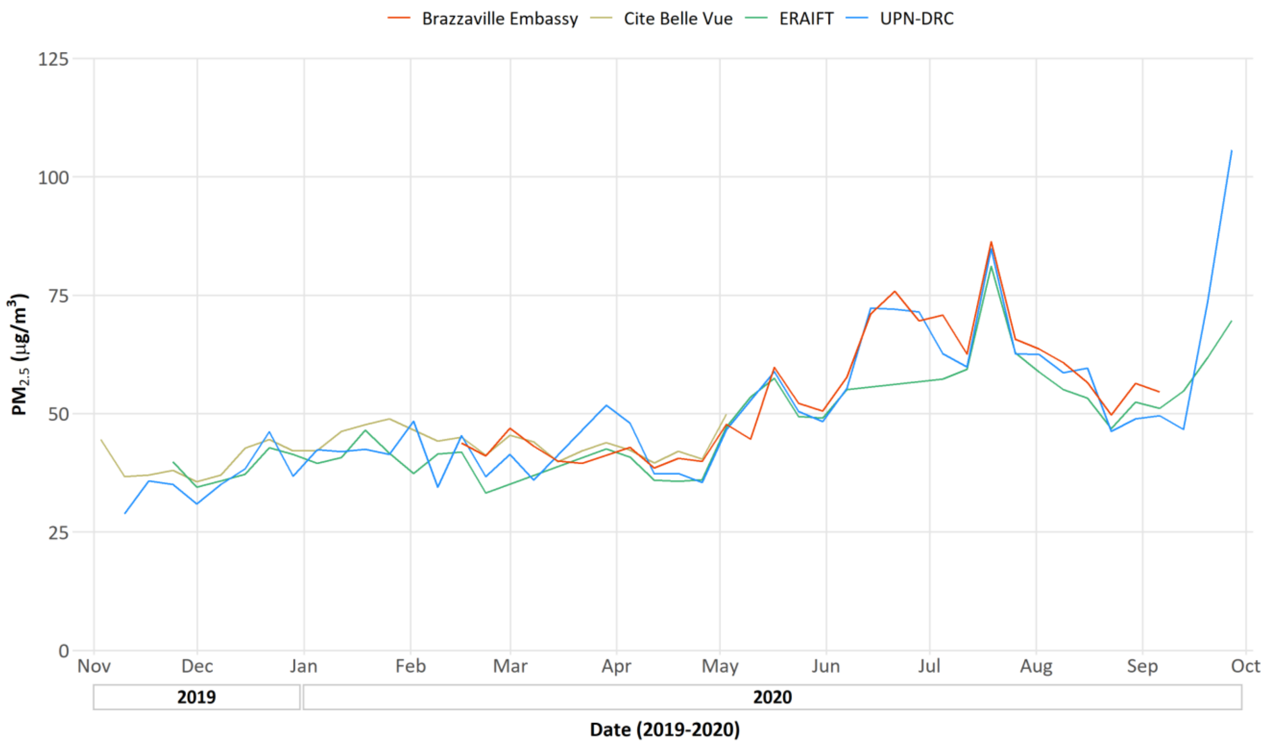


Figure 4. Weekly mean corrected data at 4 sites (excluding Kinshasa Embassy) between November 2019 and September 2020.

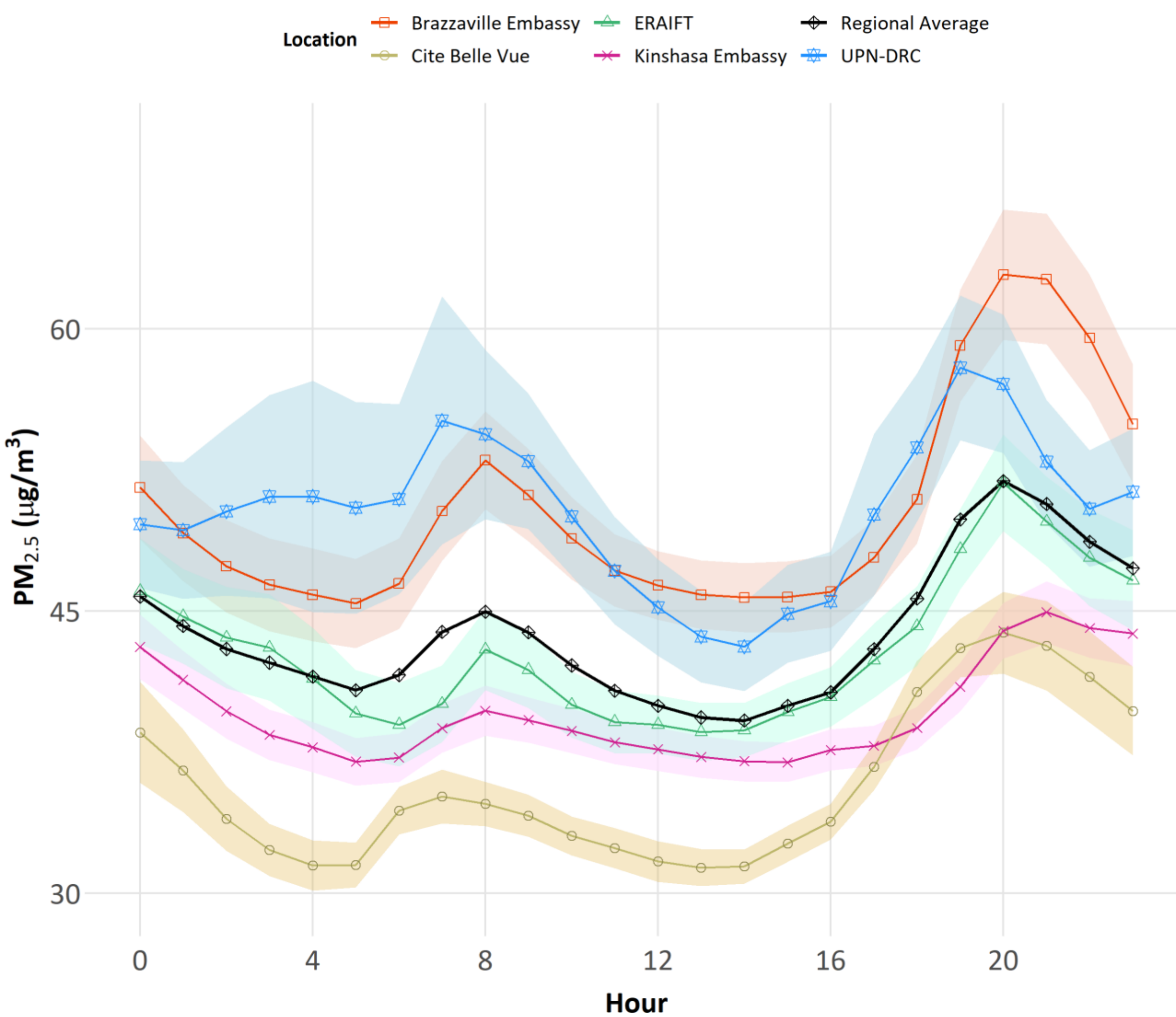


Figure 5. Diurnal average corrected $PM_{2.5}$ for the entire data record at each of the 5 sites and the site-wide average (black). Shaded areas indicate the 95% confidence interval of the hourly averages. Hour indicates the local time (West Africa Time).

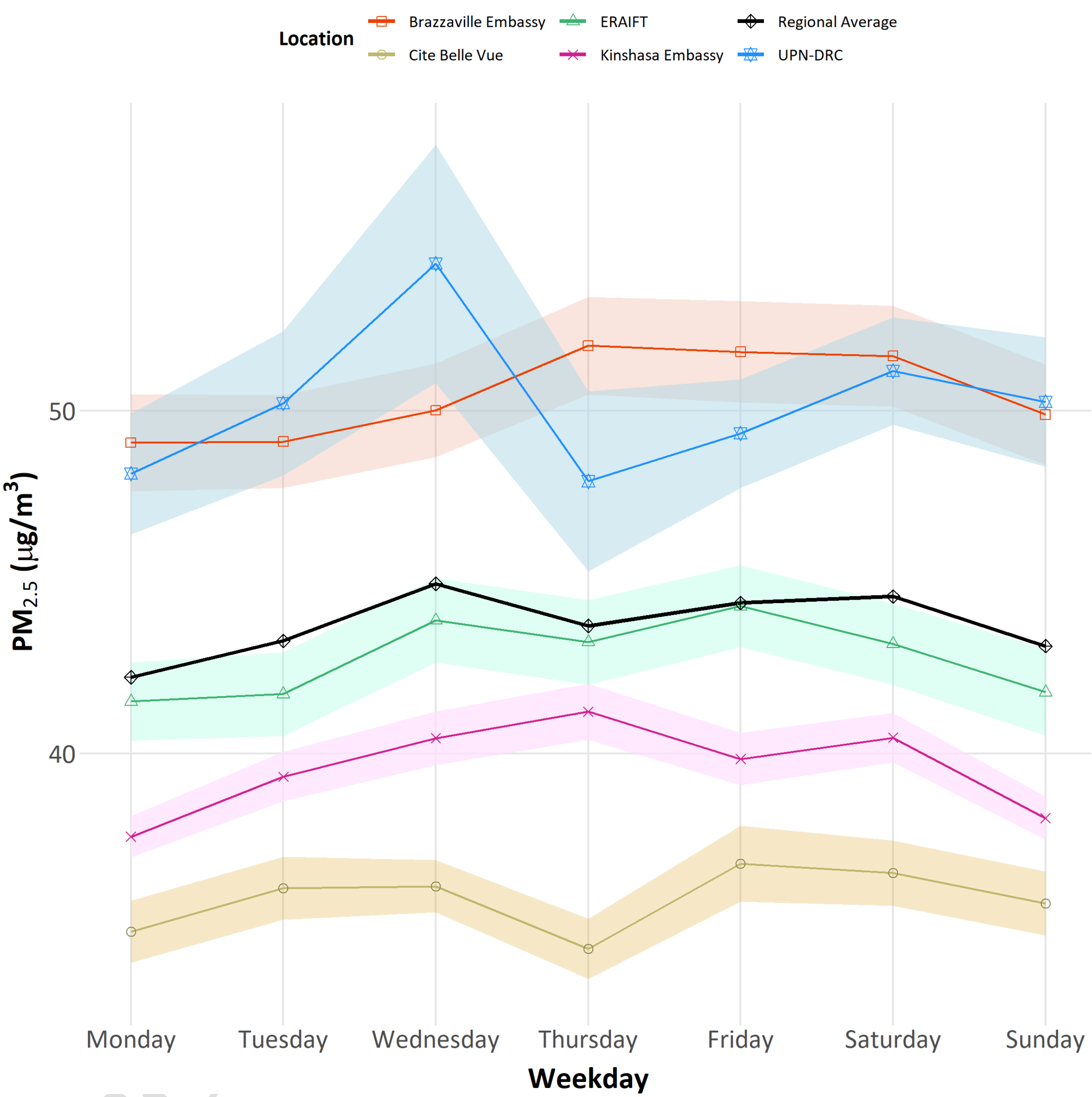


Figure 6. Day-of-the-week averages for corrected $PM_{2.5}$ for the entire dataset at each of the 5 sites, and the site-wide average. Shaded areas indicate the 95% confidence interval of the daily averages (corrected data).

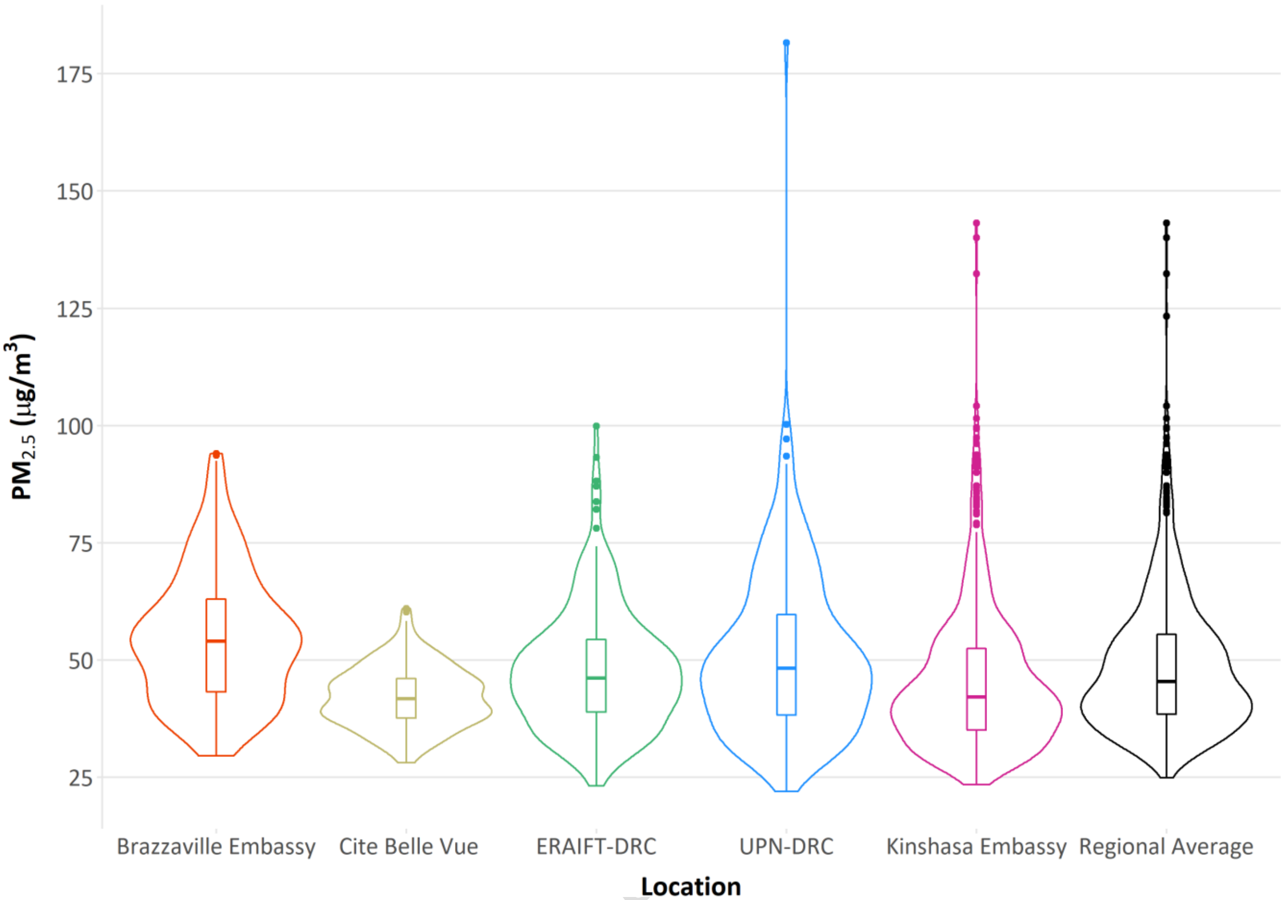


Figure 7. Violin plot of daily-averaged corrected $PM_{2.5}$ values for the entire dataset at each location and for the site-wide average. Boxes represent median and inter-quartile range.

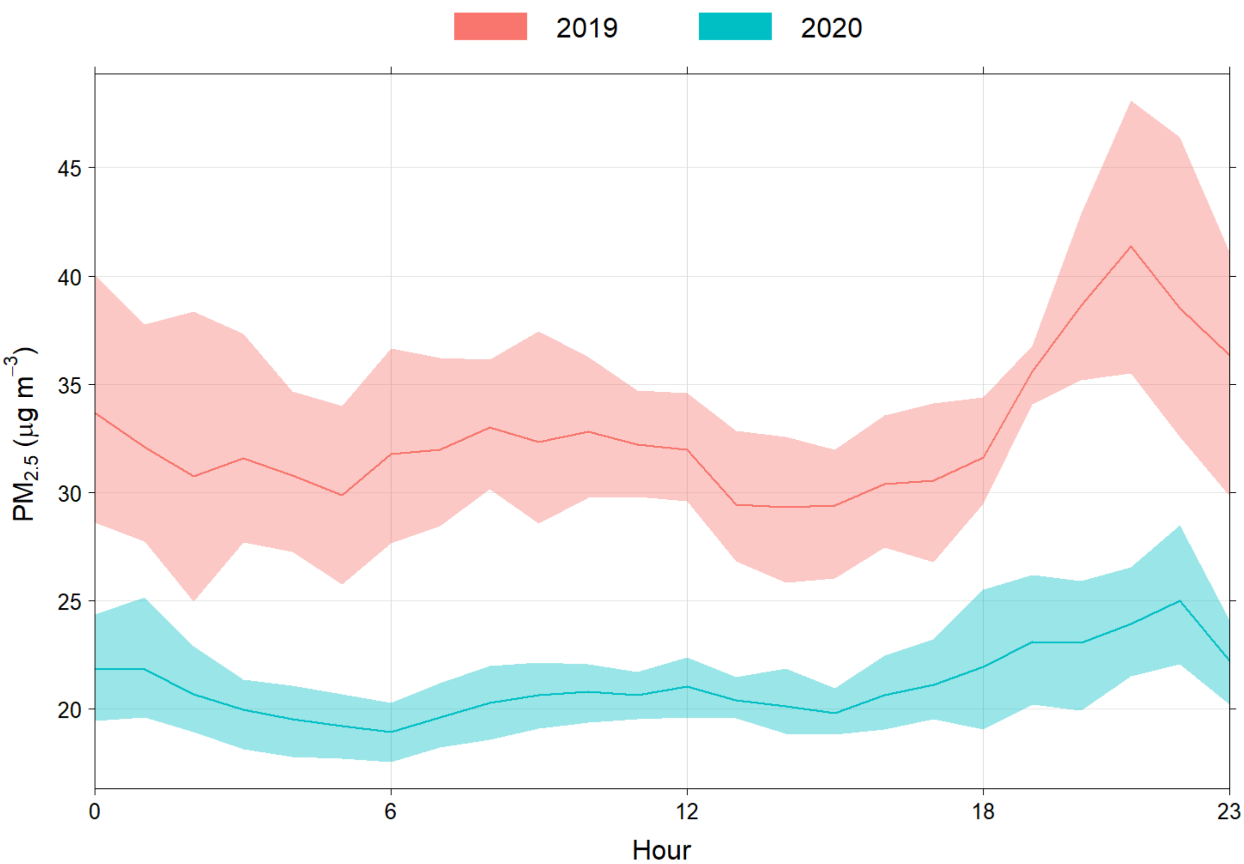


Fig 8. Analysis of corrected $PM_{2.5}$ changes during COVID19 at the US Embassy Kinshasa location. Diurnal $PM_{2.5}$ in 2019 and 2020 by hour of the day. Averaging time period is the month of April. Shaded region represents the 95% confidence interval in the mean.

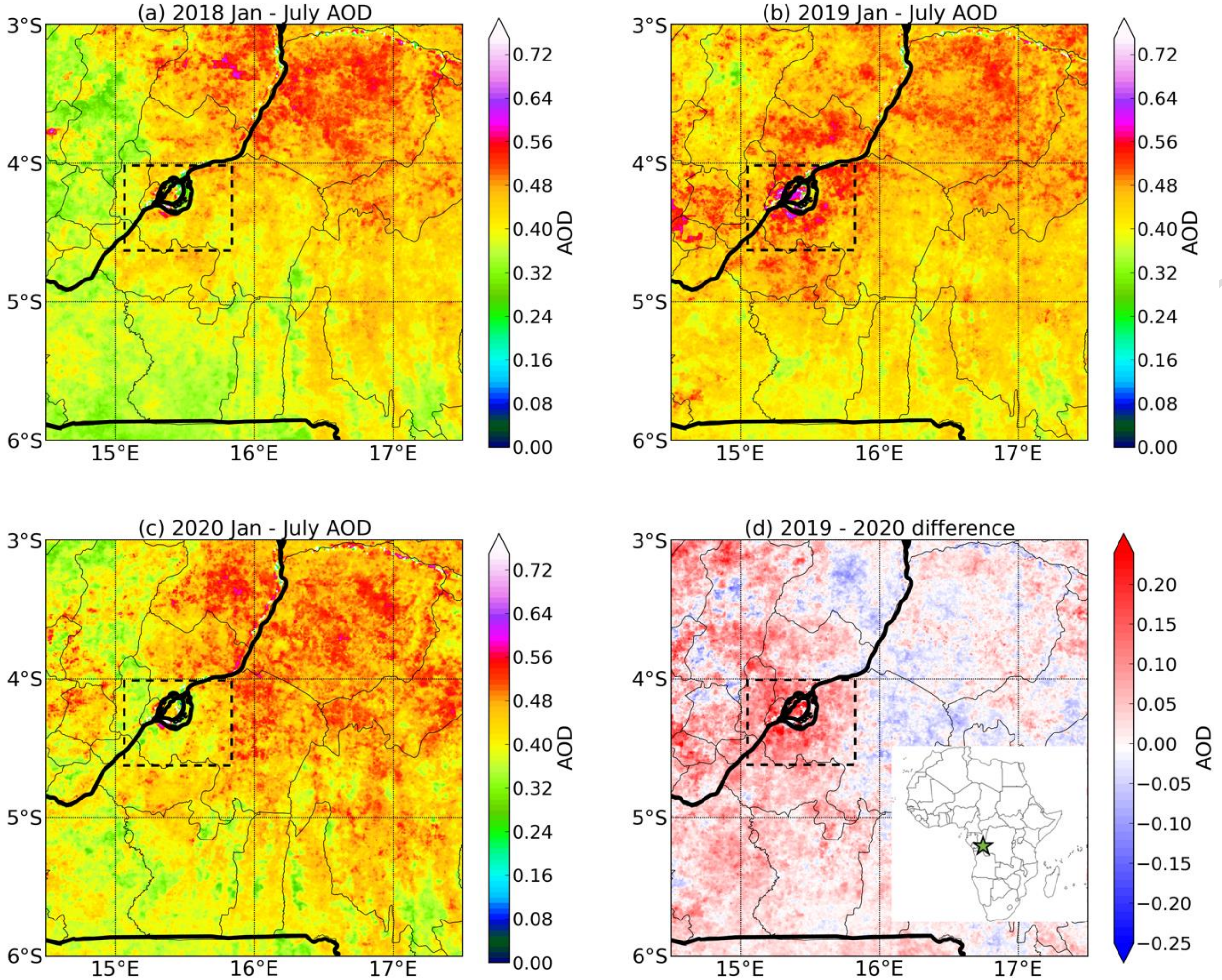
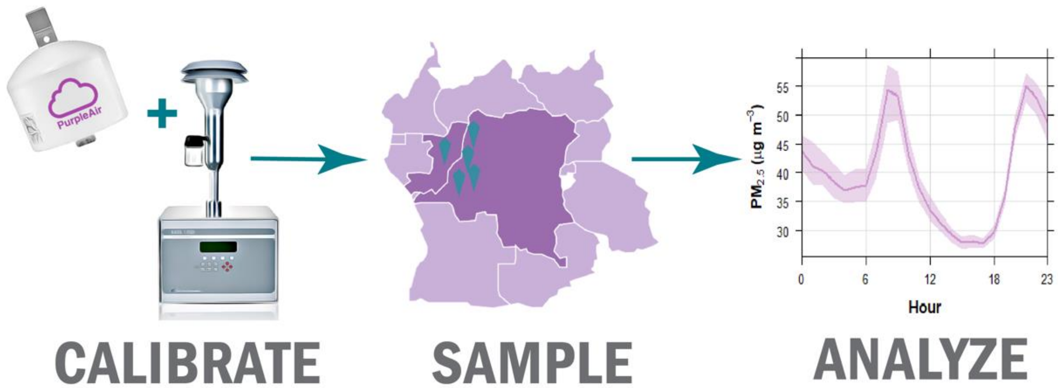


Figure 9. MAIAC Level 2 550 nm AOD at 1km resolution over the Congo region for a January through July average in (a) 2018, (b) 2019, (c) 2020, and (d) the difference between 2019 and 2020. Boundaries represent level 2 administrative boundaries (districts and communes) provided by the United Nations Office for the Coordination of Humanitarian Affairs. The black dashed box indicates the Kinshasa-Brazzaville region and the inset in panel (d) shows the broader location within the African continent.

Graphical Abstract:



Caption: Methodology and example data output from data collection campaign in Kinshasa, DRC, and Brazzaville, ROC.