Assessing the spatial-temporal relationships between land use land cover changes and $PM_{2.5}$ concentrations across 38 Indonesian Provinces (2000-2020)

Chandra S. Rahendaputri[1,2*], *Tomohito Matsuo*[1], and *Hikari Shimadera*[1]

[1]The University of Osaka, Graduate School of Engineering, Division of Sustainable Energy and Environmental Engineering, 2-1 Yamadaoka, Suita, Osaka, Japan
[2] Kalimantan Institute of Technology, School of Sustainable Development, Department of Environmental Engineering, Balikpapan 76127, East Kalimantan, Indonesia

Abstract. $PM_{2.5}$ poses significant health risks and has been a global concern, yet systematic evidence linking land use and land cover changes (LULCC) to air quality outcomes across different provinces in Indonesia remains absent. This study addresses this gap by examining spatial-temporal correlations between LULCC and $PM_{2.5}$ concentrations across all 38 provinces of Indonesia in 2000 to 2020, employing satellite-derived $PM_{2.5}$ data validated against 15 ground-based stations with R^2=0.63, Global 30-meter Land Cover (GLC) data and Global Fire Emissions Database Version 5 (GFED5) biomass burning data. Months with major biomass burning episodes were systematically excluded from Pearson correlation analysis. Results revealed substantial national transformations including forest decline from 83.1% to 77.6% and cropland expansion of 9.1% to 13.1%. However, after excluding biomass burning periods, LULCC exhibited predominantly weak correlations with $PM_{2.5}$ concentrations across most provinces, while biomass burning episodes contributed 2.6-7.4 times higher $PM_{2.5}$ concentrations. These findings demonstrate that episodic biomass burning rather than gradual land cover changes, dominates $PM_{2.5}$ pollution in Indonesia, requiring regionally tailored air quality management strategies.

1 Introduction

Air pollution in the form of Particulate Matter with diameter of less than 2.5 μm ($PM_{2.5}$) has been a global concern due to its potential health impact such as cardiorespiratory problems [1].

Indonesia is one of the developing countries in Southeast Asia, which spanning approximately 5,100 km from east to west across the equator. The nation can be grouped into six major islands, namely Sumatera, Java, Bali and Nusa Tenggara, Kalimantan, Sulawesi and Maluku and Papua, with diverse geological and meteorological conditions. As a rapidly developing nation, Indonesia has experienced extensive land use and land cover changes

* Corresponding author: chandra.suryani@ea.see.eng.osaka-u.ac.jp

(LULCC) over the past two decades, driven by agricultural expansion [2,3], urbanization [4], and industrial development [5]. Concurrently with these transformations, air quality has deteriorated significantly, with $PM_{2.5}$ concentrations in major Indonesian cities frequently exceed national standards by two to three times [6].

Existing research has characterized $PM_{2.5}$ sources in specific localities, identifying traffic emissions as the dominant contributors in urban centers, and biomass burning from forest and peatland fires as the primary source in Sumatera and Kalimantan [7]. However, critical gaps remain as existing studies focus predominantly on major cities and fire-prone regions, leaving many of Indonesia's 38 provinces unstudied, particularly in Eastern Indonesia, while temporal analysis is limited to three-to-five-year periods or episodic events, precluding assessment of long-term trends. Most critically, systematic spatial-temporal correlations between LULCC and $PM_{2.5}$ concentrations at provincial scale is absent, as case studies document fire-related $PM_{2.5}$ spikes and deforestation rates separately without integrated analysis quantifying these relationships across Indonesia's administrative provinces. This knowledge gap is particularly problematic for developing evidence-based land management policies, as governments lack quantitative evidence linking land use decisions to air quality outcomes across different geographies and climate contexts.

This study addresses these gaps by examining spatial-temporal correlations between LULCC, biomass burning, and $PM_{2.5}$ concentrations across Indonesia's 38 provinces over two decades. By excluding months with major biomass burning episodes, it isolates chronic LULC-driven $PM_{2.5}$ signals from episodic fire events, offering the first comprehensive provincial-scale nationwide analysis at policy-relevant resolution.

2 Methods

2.1 Land Cover Data

Land cover data were obtained from the Global 30-meter Land Cover Change (GLC) dataset using Landsat Imagery (1984-2022) processed through continuous change detection algorithms, adaptive classification, and temporal optimization techniques [8]. The original 37 classes were reclassified into 8 categories: Cropland, Forest, Shrubland, Grassland, Wetland, Built-up, Barren Land and Waterbody. Zonal statistics were performed on annual GLC data (2000-2020) using Indonesian provincial boundaries. The percentage of each LULC class was calculated based on pixel proportions obtained from zonal statistics. Temporal changes were subsequently quantified as year-over-year percentage changes and analyzed using statistical methods ad graphical visualization. All spatial analyses were conducted using QGIS 3.40. The study area can be seen in Figure 1.

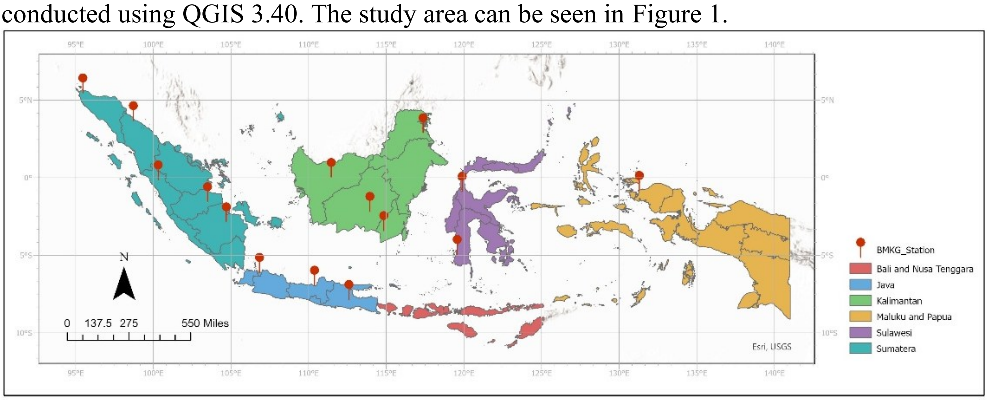

Fig. 1. Study Area and Ground Based Monitoring Map (Island grouped into (a) Sumatera: Aceh,

Sumatera Utara, Sumatera Barat, Riau, Kepulauan Riau, Jambi, Bengkulu, Sumatera Selatan, Kepulauan Bangka Belitung and Lampung, (b) Java: Banten, DKI Jakarta, Jawa Barat, Daerah Istimewa Yogyakarta, Jawa Tengah and Jawa Timur, (c) Bali and Nusa Tenggara: Bali, Nusa Tenggara Barat, and Nusa Tenggara Timur, (d) Kalimantan: Kalimantan Utara, Kalimantan Barat, Kalimantan Tengah, Kalimantan Timur, and Kalimantan Selatan, (e) Sulawesi: Sulawesi Utara, Gorontalo, Sulawesi Tengah, Sulawesi Barat, Sulawesi Tenggara, and Sulawesi Selatan, (f) Maluku and Papua: Maluku Utara, Maluku, Papua Barat Daya, Papua Barat, Papua, Papua Tengah, Papua Pegunungan, and Papua Selatan

2.2 $PM_{2.5}$ Concentration Data

The ground-based monitoring station data of $PM_{2.5}$ was limited. Therefore, this study employed a spatially and temporally continuous $PM_{2.5}$ data from Washington University's Atmospheric Composition Analysis Group (ACAG). The Raster data of $PM_{2.5}$ provided by them is with a spatial resolution of 0.1° × 0.1° and can be accessed via http://satpm25data.net. These raster datasets were obtained by combining GEOS-CHEM chemical transport model with Aerosol Optical Depth (AOD) retrievals which utilize observations from multiple satellite-based NASA instruments, which then being calibrated to global ground-based observations using a residual Convolutional Neural Network (CNN) [9].

To assess the reliability of ACAG satellite-derived $PM_{2.5}$ datasets for national-scale analysis in Indonesia, monthly point values were extracted for 15 ground-based monitoring stations (Figure 1) over 2022-2023 from Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG). Validation metrics included R^2 (acceptable ≥0.60), RMSE (acceptable <30μg/m^3), Pearson correlation (acceptable ≥0.60), linear regression slope (acceptable 0.80-1.20), and relative mean bias (RMB; acceptable ≤60%), following established criteria from national-scale $PM_{2.5}$ modelling studies and benchmarks for air quality model performance [10,11,12].

After validating the simulation data, monthly $PM_{2.5}$ data from ACAG spanning 2000-2020 were processed using zonal statistics to calculate monthly mean concentrations for each province in Indonesia.

2.3 $PM_{2.5}$ Emissions Rate from Biomass Burning Episode

In this research, PM2.5 emissions rates which are affected by biomass burning episode were analysed using the dataset from Global Fires Emission Database Version 5 (GFED5). GFED5 is a satellite-derived PM2.5 emissions dataset which is specifically attributed to biomass burning emissions [13]. These datasets isolated the contribution of fire-related activities to air pollution. This GFED5 model quantifies fire emissions in units of grams per square meter per month (g/m²/month). PM2.5 emissions rate related to biomass burning was obtained in the same manner as obtaining PM2.5 concentrations from the ACAG.

To isolate the relationship between LULC and PM2.5 concentrations, in this study a systematic approach was developed to identify and exclude months affected by biomass burning episodes from our annual PM 2.5 calculations.

2.4 Analysis

This research uses three datasets: LULC data from GLC, $PM_{2.5}$ concentrations from ACAG, and GFED5 fire emissions data. Biomass burning episodes identified through GFED5 were systematically excluded from analysis. Pearson correlations between LULCC and $PM_{2.5}$ concentrations, and between biomass burning and $PM_{2.5}$, were computed at annual scales.

Annual LULC pixel percentages were compared with annual $PM_{2.5}$ averages calculated from 9 months of data (excluding biomass burning periods), ensuring temporal alignment.

3 Results and Discussion

3.1 National and Provincial Land Use and Land Cover Change Trend Analysis

Land use and land cover from GLC were classified into eight classes namely cropland, forest, shrubland, grassland, wetland, built-up, barren land and waterbody, and calculated as percentage for each province. Provinces were grouped into Islands and analysed at national, provincial and island level.

Nationally, forest dominated the landscape but declined from 83.1% in 2000 to 77.6% in 2020, with a loss of 5.5 percentage points. Conversely, cropland expanded by 3.94 percentage points, driven by population growth and oil palm expansion (Figure 2).

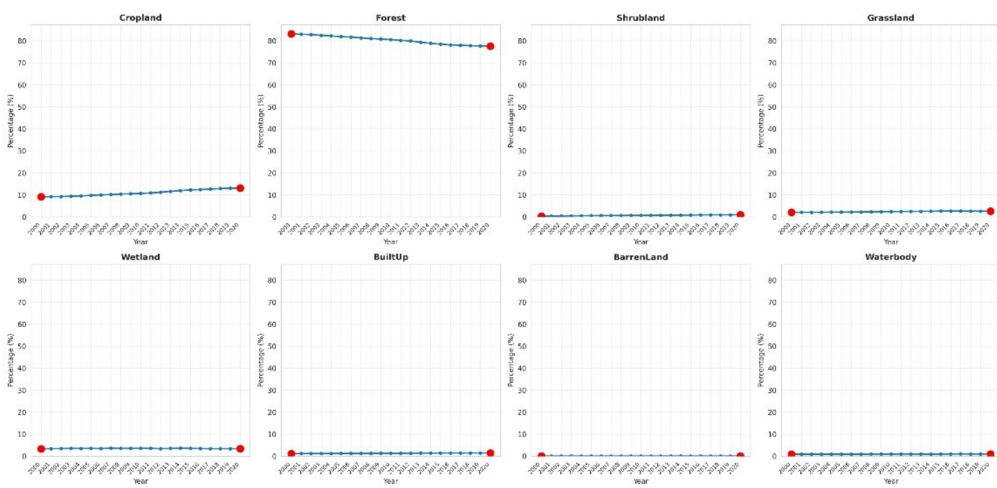

Fig. 2. National Temporal Trends of Eight Types LULC from 2000-2020

Previous studies indicate agricultural demand in tropical countries drives both yield intensity and cropland expansion, with 32% (3.09 Mha) of Indonesia’s forest loss converted to oil palm [14]. Studies investigating Indonesian forest change using 30-meter resolution Global Forest Change (GFC) data identified oil palm and timber plantations as primary deforestation drivers, with lesser contributions from agriculture [4]. Forest loss carries multiple environmental consequences, including biodiversity loss, carbon emissions, and air quality deterioration.

The GLC dataset has an overall classification accuracy of approximately 80-85%, with higher accuracy for forest (85-90%) and lower accuracy for transitional classes like shrubland (70-75%) [8]. This uncertainty could affect our observed changes by approximately 2-3 percentage points. For example, the 5.5 percentage points forest loss could range from 4-7 percentage points when accounting for classification errors. However, the direction and magnitude of changes remain robust, and temporal trends are more reliable than absolute values since classification errors are largely systematic across years.

For deeper understanding of land use land cover changes trend analysis, provincial land use land cover changes trend analysis was also conducted. Figure 3 displays a heatmap LULCC for all provinces.

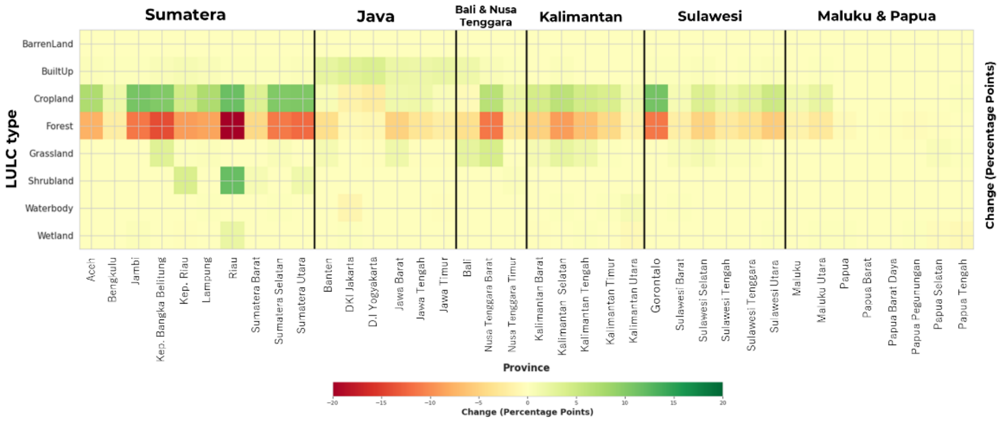

Fig. 3. Provincial LULCC Heatmap 2000-2020 Sorted by Island (Red = decreasing land, Green = increasing land)

At provincial level, Riau suffered the highest percentage of forest loss by 25.87 percentage points, with an increasing percentage point (pp) of cropland (+11.61 pp), shrubland (+11.77 pp) and wetland (+2.13 pp). In built-up areas, highest increase was found in Daerah Istimewa Yogyakarta (+3.65 pp) followed by DKI Jakarta, the capital city, in the second. Increasing built-up areas mostly happened in Java Island which being centre of development in Indonesia. At the island level, Sumatera experienced the greatest forest loss due to its extensive initial forest cover, high fire susceptibility, and intensive oil palm and pulpwood plantation. This trend is followed by Kalimantan Island with the same criteria.

3.2 Assessment of Reliability of $PM_{2.5}$

To assess the reliability of raster datasets from Washington University's Atmospheric Composition Analysis Group, point value from the monthly raster dataset in timespan of 2022-2023 was extracted by using point of 15 ground-based monitoring stations in Indonesia. The scatterplot result of observed and simulated value can be seen in Figure 4.

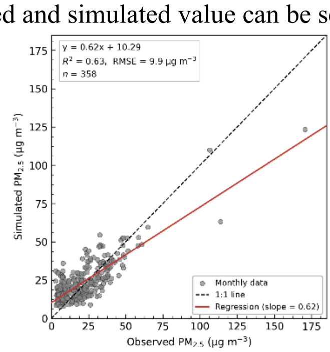

Fig. 4. Scatterplot of overall Observed and Simulated $PM_{2.5}$

Figure 4 shows moderate correlation between observed and simulated $PM_{2.5}$(R^2=0.63, r=0.82, RMSE=9.94 µg/m^3, RMB=32.91%). The model underestimates higher concentrations (slope=0.62) with systematic offset (intercept=10.29). These metrics meet established criteria for national-scale $PM_{2.5}$ modelling as has been discussed in section 2.2. This accuracy is acceptable for capturing spatiotemporal variability at national scale, though systematic biases remain that warrants further refinement.

3.3 $PM_{2.5}$ Concentration Trend and Biomass Burning Episode Analysis

Figure 5 shows the aggregated $PM_{2.5}$ concentration trend for six major island groups in Indonesia.

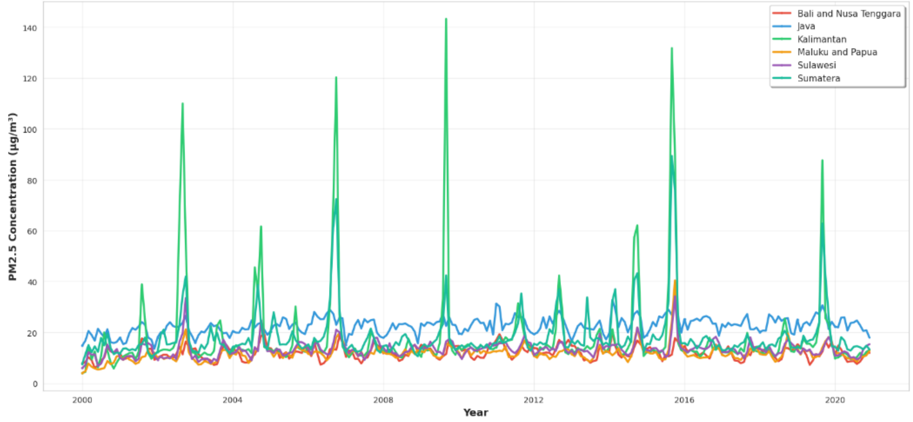

Fig. 5. $PM_{2.5}$ Trends by Major Island Groups (2000-2020)

There was an observable dramatic peak in Sumatera and Kalimantan each year. Given that Kalimantan and Sumatera are dominated by forest, and have suffered substantial forest loss, the dramatic peaks suggest that biomass burning episodes occurred regularly in Indonesia and significantly affected $PM_{2.5}$ concentration. For deeper understanding, Figure 6 shows the seasonal pattern of both GFED $PM_{2.5}$ emissions and the pattern of $PM_{2.5}$ concentration. There are an observable biomass burning episode pattern which happened mostly in around August to October in each island. Thus, those months will be excluded from correlation analysis to produce an annual mean $PM_{2.5}$ concentration with excluding month with biomass burning episodes.

Fig. 6. Biomass Burning Emissions vs $PM_{2.5}$ Concentration by Island Group (2000-2020)

Excluding biomass burning months underestimates annual exposure by 10.2% nationally (up to 36% in Kalimantan), as August to October concentrations are 128-210% of annual means. However, for LULCC and $PM_{2.5}$ correlation analysis, exclusion is methodologically justified because biomass burning represents episodic, event-driven pollution that can mask gradual LULCC signals. This separation enables detection of persistent LULCC-driven trends rather than conflating them with seasonal burning variability.

3.4 Correlation Analysis between LULCC and $PM_{2.5}$ Concentration

To assess the correlation between LULCC and $PM_{2.5}$ concentration, Pearson correlation analysis was executed. The data used was annual mean $PM_{2.5}$ concentration with excluding month with biomass burning episodes as have been discussed in previous sections. Figure 7 presents the provincial Pearson correlation heatmap between LULCC and annual mean $PM_{2.5}$ (BB excluded) concentration across all 38 provinces.

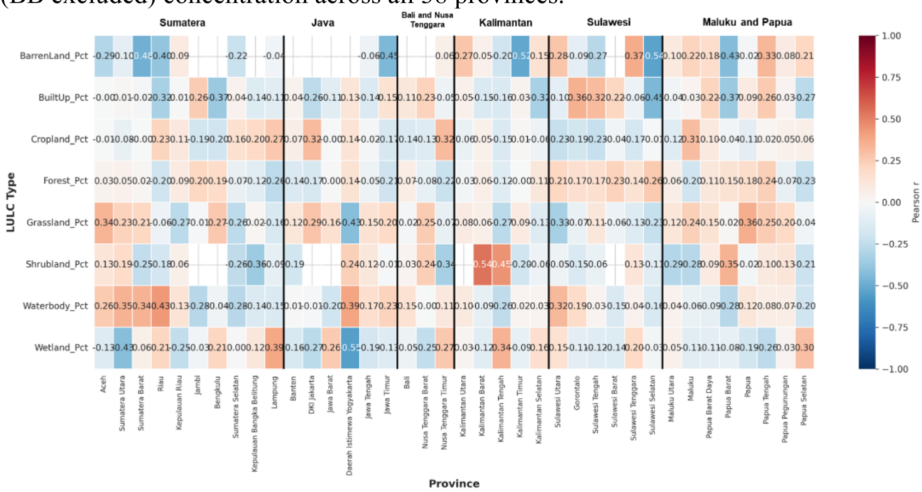

Fig. 7. Provincial Pearson Correlation of LULCC and Annual Mean $PM_{2.5}$ BB Excluded Concentration

At provincial level, each LULCC correlate in various manner to the $PM_{2.5}$ concentration. The highest correlation can be found in Shrubland percentages changes in Kalimantan Barat and Kalimantan Tengah. This might be driven by frequent peatland fires [15]. In Riau, where the deforestation occurred massively and cropland was increasing in tremendous amounts, cropland have positive moderate correlation. It means that when the cropland area is changing, $PM_{2.5}$ concentration will follow to the same direction. Daerah Istimewa Yogyakarta, city with highest increasing area of built up exhibit a weak positive correlation with $PM_{2.5}$ concentration.

4 Conclusions

This study examined spatial-temporal correlations between LULCC-$PM_{2.5}$ concentrations across Indonesia's 38 provinces from 2000 to 2020. Despite a 5.5 pp forest loss and 3.9 pp cropland expansion, LULCC showed weak correlation with $PM_{2.5}$ after excluding biomass burning periods. Biomass burning dominated $PM_{2.5}$ variability, contributing 2.6-7.4 times higher concentrations than background levels and generating acute episodic spikes of 128-210% above annual means. Quantifying precise source contribution ratios would require controlled scenario modelling (e.g., WRF-CHEM) with independent emission source toggling. These results underscore the need for region-tailored air quality control strategies in Indonesia: forest fire controlled in fire-prone region (e.g., Sumatera) , and urban traffic control in Java. Future work should expand ground-based monitoring and develop regionally calibrated models to strengthen attribution.

Author would like to acknowledge the Indonesia Endowment Fund for Education (LPDP) for PhD scholarship.

References

1. Weichenthal, et al., $PM_{2.5}$, oxidant defence and cardiorespiratory health: a review, Environ. Health 12, 40-40 (2013)
2. Austin, et.al., What causes deforestation in Indonesia?, Environ. Res. Lett. **14**. (2019)
3. D.L.A. Gaveau, et al., Slowing deforestation in Indonesia follows declining oil palm expansion and lower oil prices, PLoS ONE 17, e0266178 (2022)
4. F. Rachman, et.al., Insights from 30 Years of Land Use/Land Cover Transitions in Jakarta, Indonesia, via Intensity Analysis, Land **13**, 545 (2024)
5. M. Mandala, et al., Land Use and Land Cover Change in East Java Indonesia from 1972 to 2021: Learning from Landsat, J. Environ. Res. Eng. Manag. **80**, 57-69 (2024)
6. M. Santoso, et al., Assessment of PM2.5 and PM10 concentrations in Indonesian urban environments and their chemical characterization, Pollut. Res. 11, 2215-2225 (2020)
7. S. Siregar, et al., Chemical Composition, Source Appointment and Health Risk of $PM_{2.5}$ and $PM_{2.5\text{-}10}$ during Forest and Peatland Fires in Riau, Indonesia, Aerosol Air Qual. Res. **22**, 220015 (2022)
8. X. Zhang, et al., GLC_FCS30D: the first global 30 m land-cover dynamic monitoring product with a fine classification system from 1985 to 2022 generated using dense-time-series Landsat imagery and the continuous change-detection method, Earth Syst. Sci. Data **16**, 1353-1381 (2024)
9. A. van Donkelaar, et al., Enhancing Global Estimation of Fine Particulate Matter Concentrations by Including Geophysical a Priori Information in Deep Learning, ACS ES&T Air (2024)
10. G. Shaddick, et al., Data Integration for the Assessment of Population Exposure to Ambient Air Pollution for Global Burden of Disease Assessment, Environ. Sci. Technol. 52, 9069 (2018)
11. M.S. Hammer, A. van Donkelaar, et al., Global Estimates and Long-Term Trends of Fine Particulate Matter Concentrations (1998-2018), Environ. Sci. Technol. 54, 7879 (2020)
12. J.W. Boylan, A.G. Russell, PM and light extinction model performance metrics, goals, and criteria for three-dimensional air quality models, Atmos. Environ. 40, 4946 (2006)
13. G.R. Van der Werf, et al., Global fire emissions estimates during 1997-2016, Earth Syst. Sci. Data **9**, 697-720 (2017)
14. K. Winkler, et al., Six decades of global crop yield increase and cropland expansion from 1960 to 2020, Environ. Res. Commun. **7**, (2025)
15. A. Schmidt, et al., Fire frequency, intensity, and burn severity in Kalimantan's threatened Peatland areas over two Decades , Front. For. Glob. Change **7**, 1221797 (2024)