



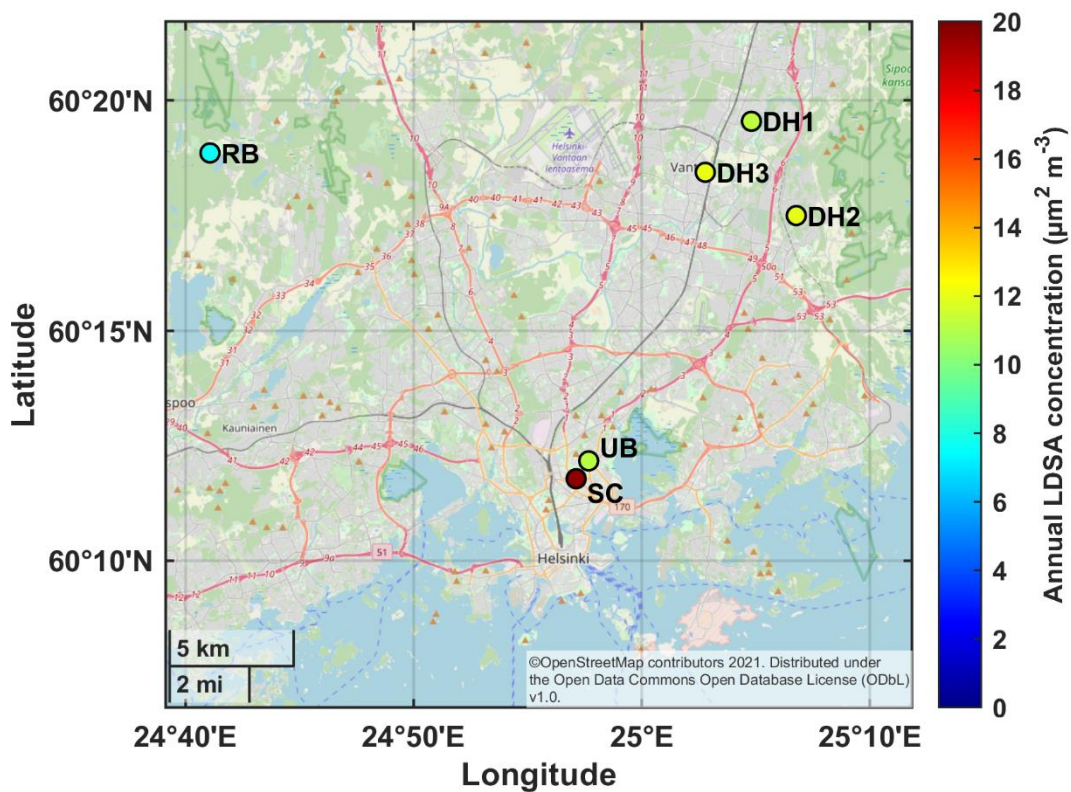
*Supplement of*

## **Input-adaptive linear mixed-effects model for estimating alveolar lung-deposited surface area (LDSA) using multipollutant datasets**

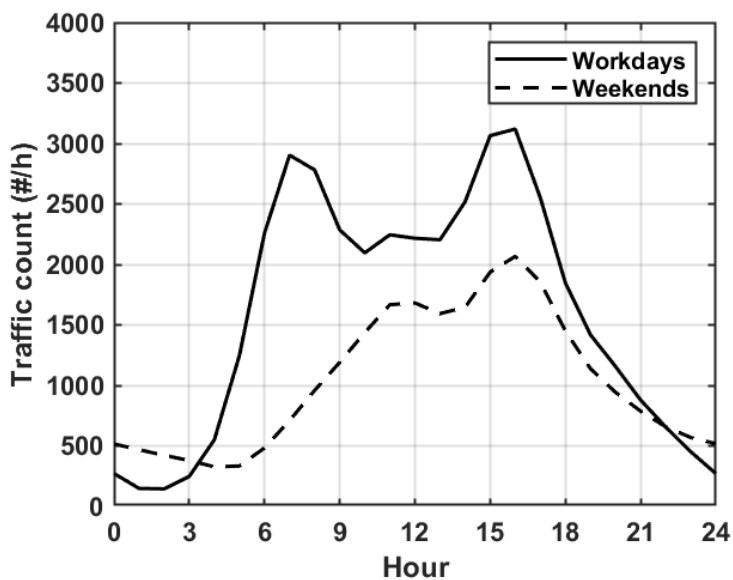
**Pak Lun Fung et al.**

*Correspondence to:* Pak Lun Fung ([pak.fung@helsinki.fi](mailto:pak.fung@helsinki.fi)) and Tareq Hussein ([tareq.hussein@helsinki.fi](mailto:tareq.hussein@helsinki.fi))

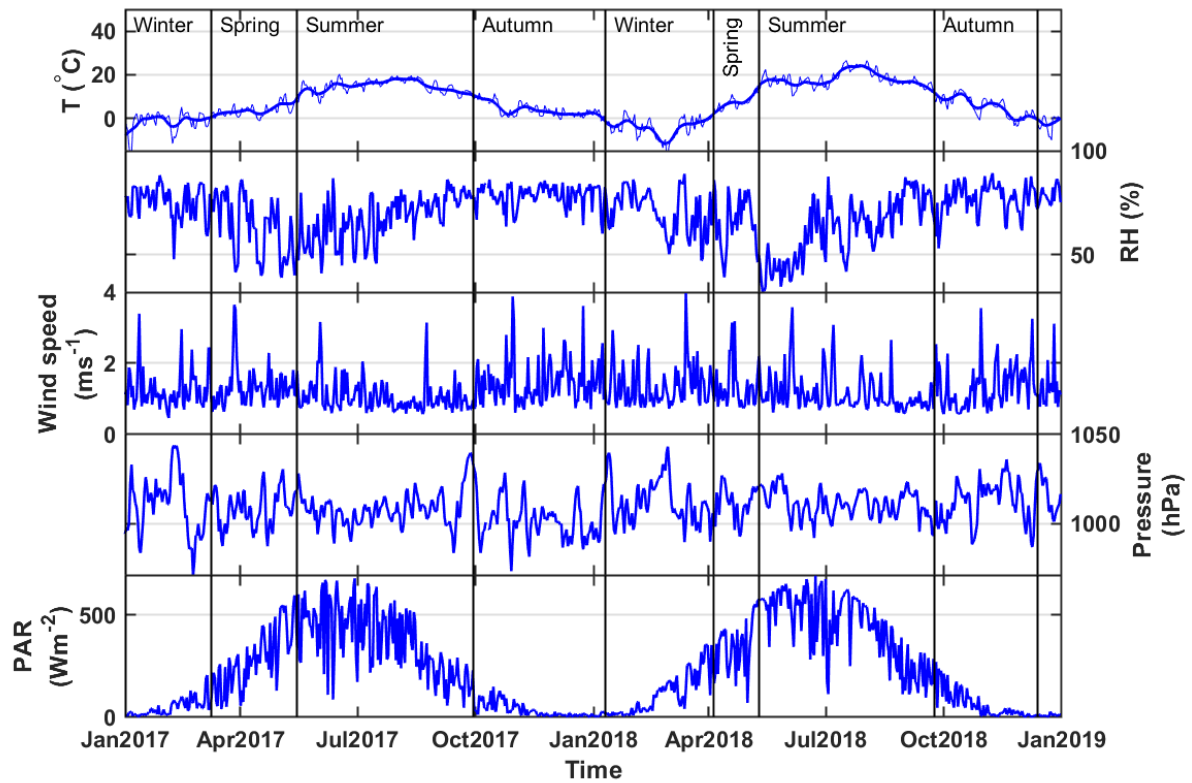
The copyright of individual parts of the supplement might differ from the article licence.



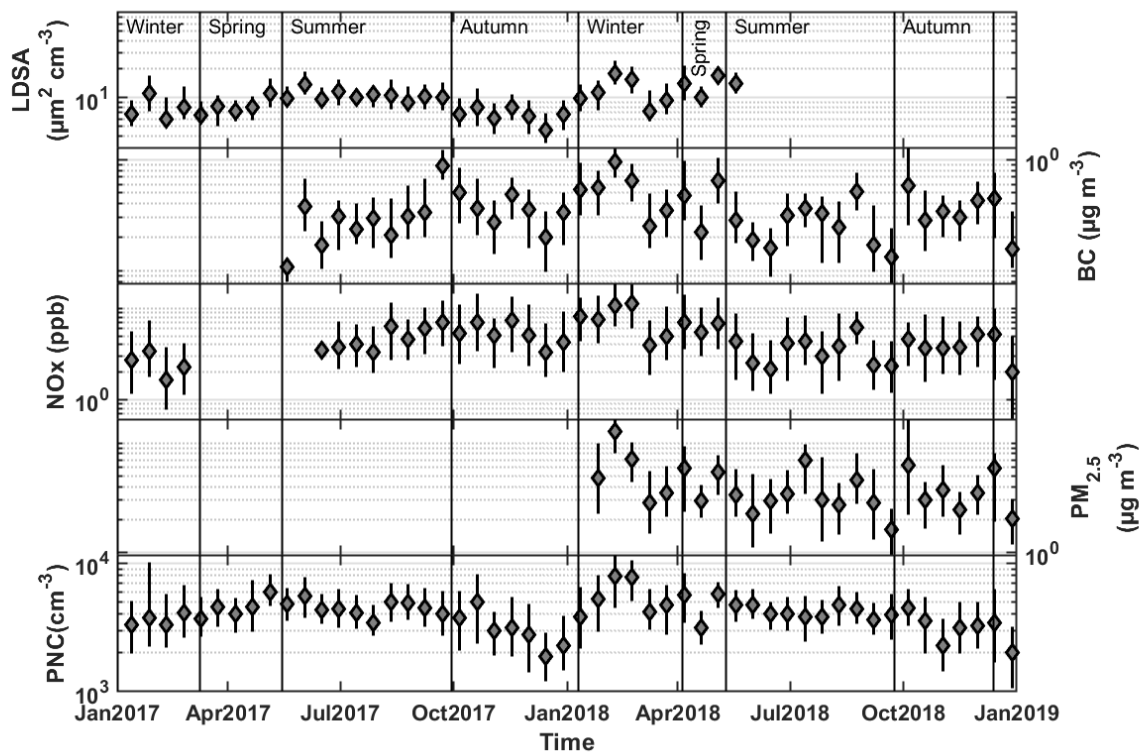
**Figure S1.** Location of measurement stations involved. SC, UB, DH1–3 and RB represent street canyon, urban background, detached houses and regional background, respectively. The corresponding annual LDSA concentration at each site is indicated by different colours. The copyright of the map belongs to OpenStreetMap contributors 2021, distributed under the Open Data Commons Open Database License (ODbL) v1.0.



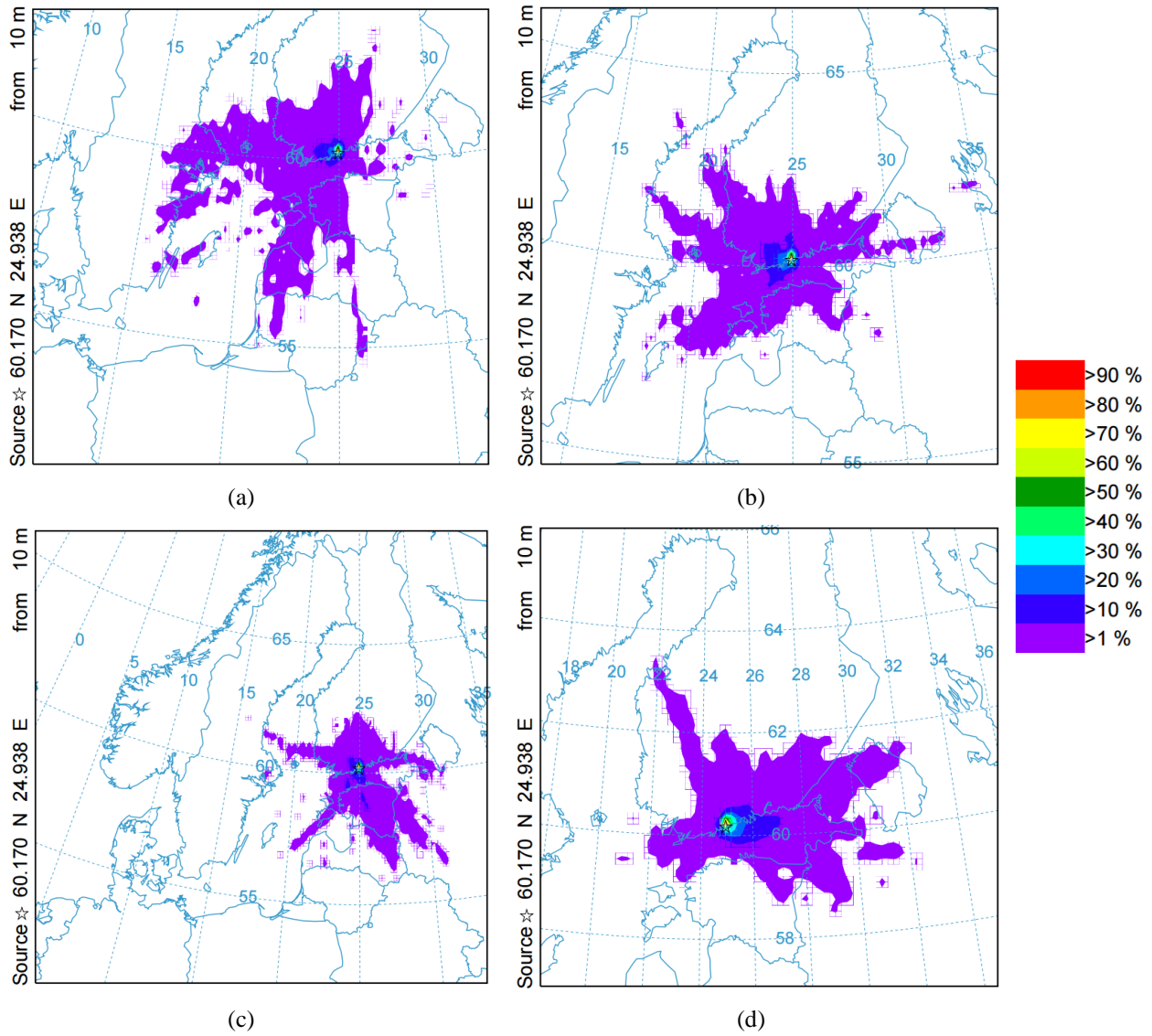
**Figure S2.** Hourly traffic counts on workdays (solid) and weekends (dashed) site in year 2017–2018. The vehicle counting is operated by the City of Helsinki and the counting site is located 600 m north of the Mäkelänkatu SC air quality monitoring site.



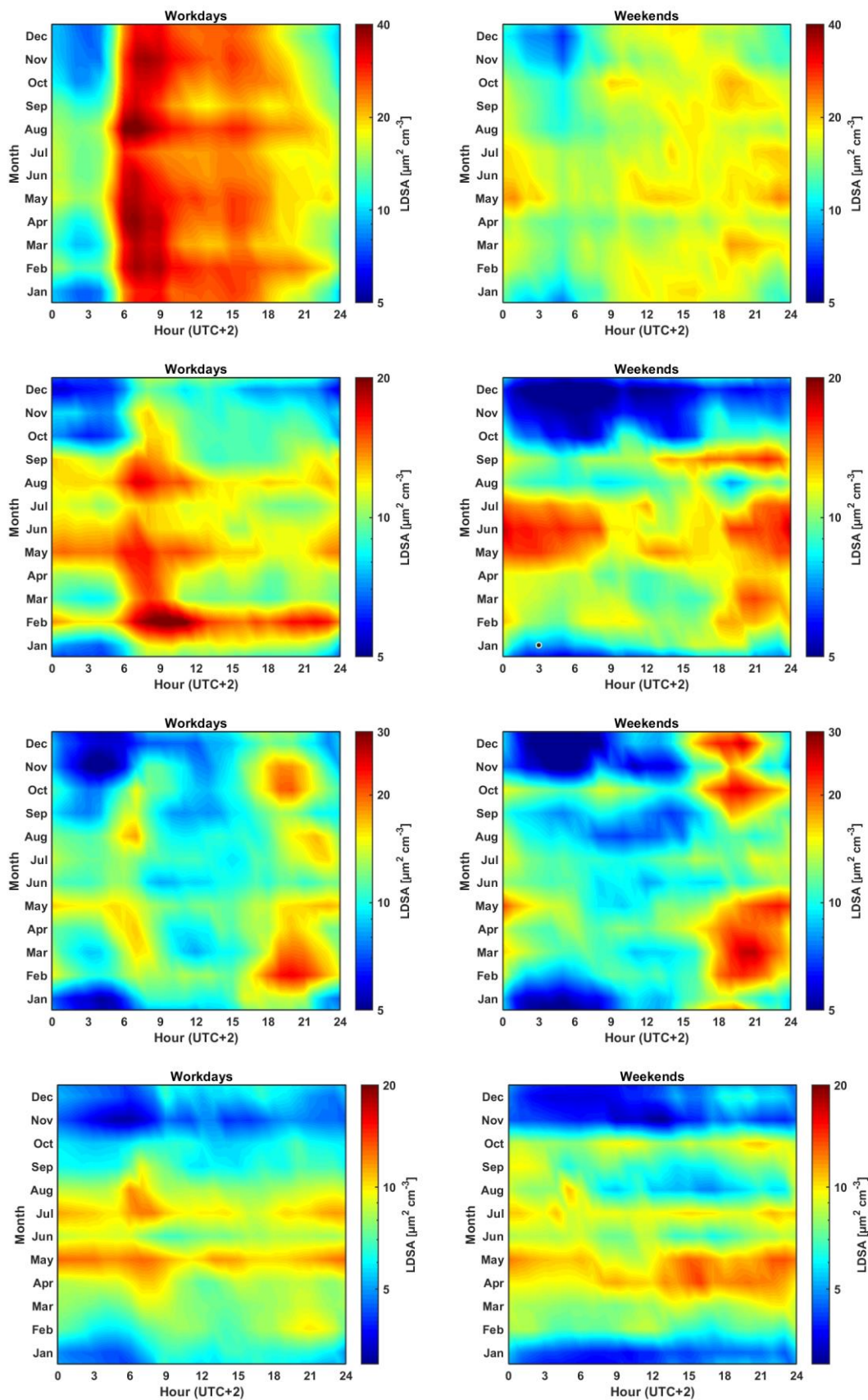
**Figure S3.** Time series of daily meteorological conditions (First to end row: temperature ( $T$ ,  $^{\circ}\text{C}$ ), relative humidity (RH, %), wind speed ( $\text{m s}^{-1}$ ), pressure (hPa) and photosynthetically active radiation (PAR,  $\text{Wm}^{-2}$ )) at the SC site during the measurement period 1 Jan 2017–31 Dec 2018. Moving average of temperature (bold line) is also calculated and used for separating seasons.



**Figure S4.** Time series of the selected air pollutant parameters (First to end row: LDSA ( $\mu\text{g m}^{-3}$ ), BC ( $\mu\text{g m}^{-3}$ ), NO<sub>x</sub> (ppb), PM<sub>2.5</sub> ( $\mu\text{g m}^{-3}$ ) and PNC ( $\text{cm}^{-3}$ )) at Kumpula UB site during the measurement period 1 Jan 2017–31 Dec 2018 (LDSA measurements until May 2018). Each bar represents a period of two weeks where the shaded diamond marker is the median and the vertical error bars are the 25<sup>th</sup> and 75<sup>th</sup> percentiles. Seasons are thermally separated.

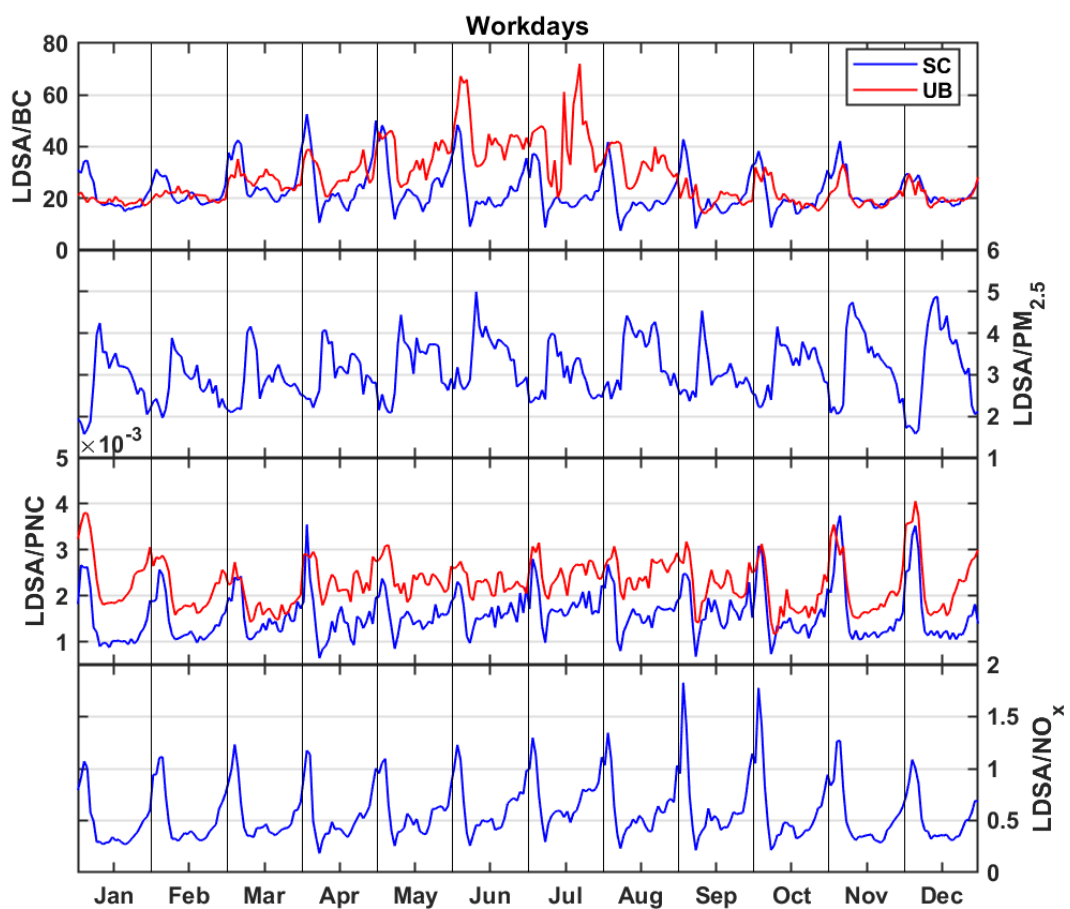


**Figure S5.** The monthly frequencies of backward trajectory by NOAA HYSPLIT Trajectory Model (<https://www.ready.noaa.gov/HYSPLIT.php>; Last assessed: 8 November 2021) with the starting point in Helsinki for (a) January 2017, (b) July 2017, (c) January 2018 and (d) July 2018. The spatial trajectory frequency was 0.25 degree. The starting interval was 3 hours and the level height was set to be 10 m. The total run time was 24 hours. The frequencies were calculated by the number of endpoints per grid divided by the maximum number of endpoints in any grid in percentage. Cold and warm colors represent low and high percentages, respectively, as shown on the right in the color legend.

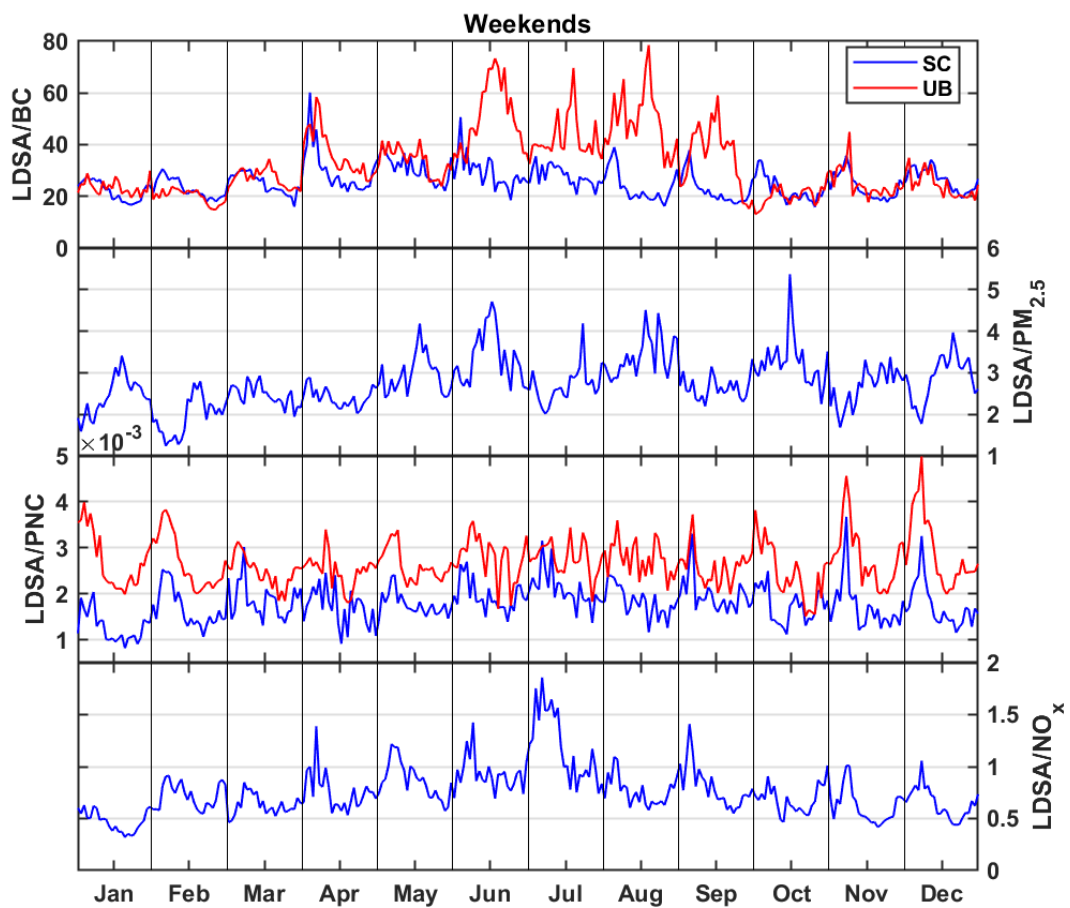


**Figure S6.** Heatmaps showing the average measured LDSA in different months and different hours at the four sites for workdays and weekends. First to end row represent: SC (2017–2018), UB (2017–May 2018), DH (2018) and RB (2018). The colour scale represents the LDSA concentrations. Note that the colour scales vary from site to site for better illustration.



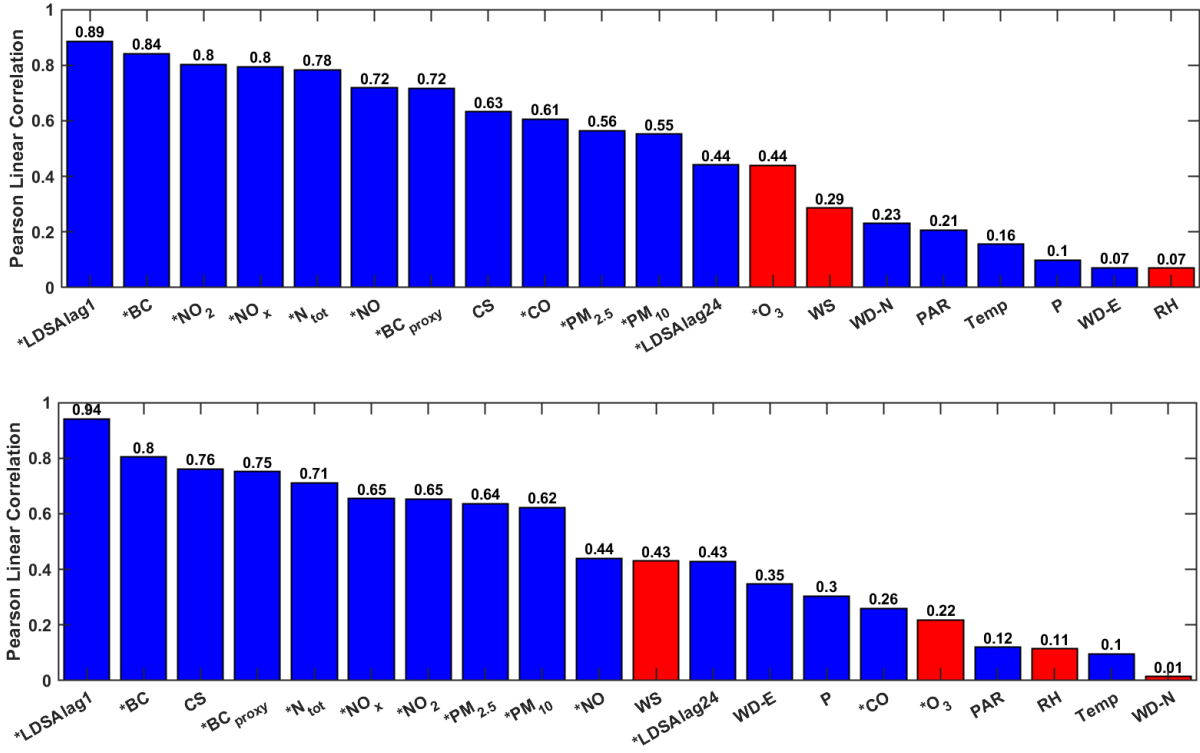


**Figure S7.** Hourly ratio of LDSA and other air pollutants including BC, PM<sub>2.5</sub>, PNC and NO<sub>x</sub> over workdays in different months at Mäkelänkatu SC site (blue, 2017–2018) and at Kumpula UB site (red, 2017–May 2018).



**Figure S8.** Hourly ratio of LDSA and other air pollutants including BC, PM<sub>2.5</sub>, PNC and NO<sub>x</sub> over weekends in different months at Mäkelänkatu SC site (blue, 2017–2018) and at Kumpula UB site (red, 2017–May 2018).





**Figure S9.** Pearson correlation coefficients ( $r$ ) as bar chart (a) at SC and (b) at UB. Blue and red bars represent positive and negative correlation, respectively. Parameters with asterisk are transformed in a logarithm scale. LDSAlag1 and LDSAlag24 represent the LDSA measurements with lag of 1 and 24 hours, respectively. CS and BC<sub>proxy</sub> stand for condensation sink and black carbon proxy based on Fung et al (2020). These four variables are not included in the data analysis.

**Table S1.** Measurement sites specification

Station name and code	Latitude	Longitude	Sampling height (from ground level)	Nearby street and distance	Number of vehicles per workday	Retrieval period
Mäkelänkatu, SC	60°11'47''N	24°57'07''E	4 m	Mäkelänkatu, 0.5 m	28 100	2017–2018
Kumpula, UB	60°12'10''N	24°57'40''E	4 m (gases and aerosols, including LDSA) 31 m (meteorological variables)	Erik Palménin aukio, 40 m Hämeentie, 150 m	- 45 000	2017–2018
Rekola, DH1	60°19'33''N	25°4'48''E	4 m	Laurantie, 5 m Peijaksentie, 260 m Hanabölenie, 280 m	<300 5 600 2 400	2018
Itä-Hakkila, DH2	60°17'30'' N	25°6'46''E	4 m	Koulutie, 6 m Palttinatie, 5m	2 700 1 200	2018
Hiekkaharju, DH3	60°18'27''N	25°2'47''E	2 m	Metsätähdentie, 20 m Talkootie, 370 m	<300 4 000	2018
Luukki, RB	60°18'52''N	24°41'05''E	4 m	Vihdintie, 800 m	4 300	2018

**Table S2.** List of variables and the corresponding instruments used in various sites.

		SC	UB	DH1-3	RB	
Aerosol variables	LDSA	x	x	x	x	Pegasor AQ Urban (Pegasor Ltd.)
	PSD	x	x			SC: DMPS (Vienna DMA and Airmodus A20 CPC) UB: Twin DMPS (Hauke-type DMA and TSI Model 3025 CPC + Hauke-type DMA and TSI Model 3010 CPC)
	PM <sub>2.5</sub> /PM <sub>10</sub>	x	x			SC: TEOM 1405 UB: TEOM 1405-D
	BC	x	x			MAAP Thermo Scientific 5012
Gaseous variables	NO <sub>x</sub> /NO <sub>2</sub> /NO	x	x			Chemiluminescence analyzers SC: APNA-370 (Horiba) UB: Thermo TEI42S
	O <sub>3</sub>	x	x			SC: UV photometric analyzers (Horiba APOA-370 and Thermo Model 49i) UB: IR-absorption photometer TEI49
	CO	x	x			Non-dispersive IR-absorption analyser SC: APMA-360 (Horiba) UB: APMA-370 (Horiba)
Weather variables	Temp/RH/WS/WD/P/PAR	x	x			Respective instruments

