2020 Advanced Institute on Health Investigation
and Air Sensing for Asian Pollution (AI on Hi-ASAP)
On-line, October 5, 6, 8 & 15, 2020
Academia Sinica, Taipei, Taiwan

Overview of Data Analysis Plan of Hi-ASAP

Shih-Chun Candice LUNG 龍世俊

- Research Center for Environmental Changes, Academia Sinica, Taiwan
- Center for Sustainability Science, Academia Sinica, Taiwan

Objectives of Hi-ASAP

- To apply low-cost sensors
  - (1) to assess $PM_{2.5}$ exposure levels, patterns, behaviors, and source characteristics of short-term or peak exposures
  - (2) to evaluate the changes in health indicators of acute health effects
- in order to
  - (1) assess the short-term $PM_{2.5}$ damage coefficients of exposure-health relationship
  - (2) provide scientific evidences to set criteria or ceiling levels of $PM_{2.5}$ with shorter exposure periods (ex. seasonal, 8-hour or hourly)

Sections of Hi-ASAP Science and Implementation Plan

- 1. Overview and Objectives
- 2. Data policy (data sharing and **quality assurance and quality control, QA/QC**)
- 3. Road map and Timetable

Key scientific focus

- 4. Exposure-Health Evaluation
- 5. Exposure Assessment
- 6. Environmental and Community Monitoring
- 7. Source Characterization

- 8. Stakeholder Engagement
- 9. Working Groups and Conveners
- 10. Path Forward

Research Questions (initial planning)

- What are the **peak $PM_{2.5}$ exposure levels** and patterns of Asian population, especially those **high-exposure or susceptible** populations?
- What are the **sources and activities** causing **peak $PM_{2.5}$** exposures and the **controllable factors** associated with those sources and activities?
- What are **the $PM_{2.5}$ damage coefficients of exposure-health relationship** of peak exposures for **lung and heart conditions**? Are the damage coefficients for the same health outcome **different in different $PM_{2.5}$ concentration ranges**? The huge differences in $PM_{2.5}$ levels in the MANGO region provide a testbed to evaluate this question
- What are the **chemical and toxicological properties** of high-exposure sources, especially **distinctive Asian sources**?
- Should there be **a ceiling value or short-term standard for $PM_{2.5}$** (ex. seasonal, 8-hour or hourly)? **What other considerations** needed to be included to promote the establishment of such a standard?

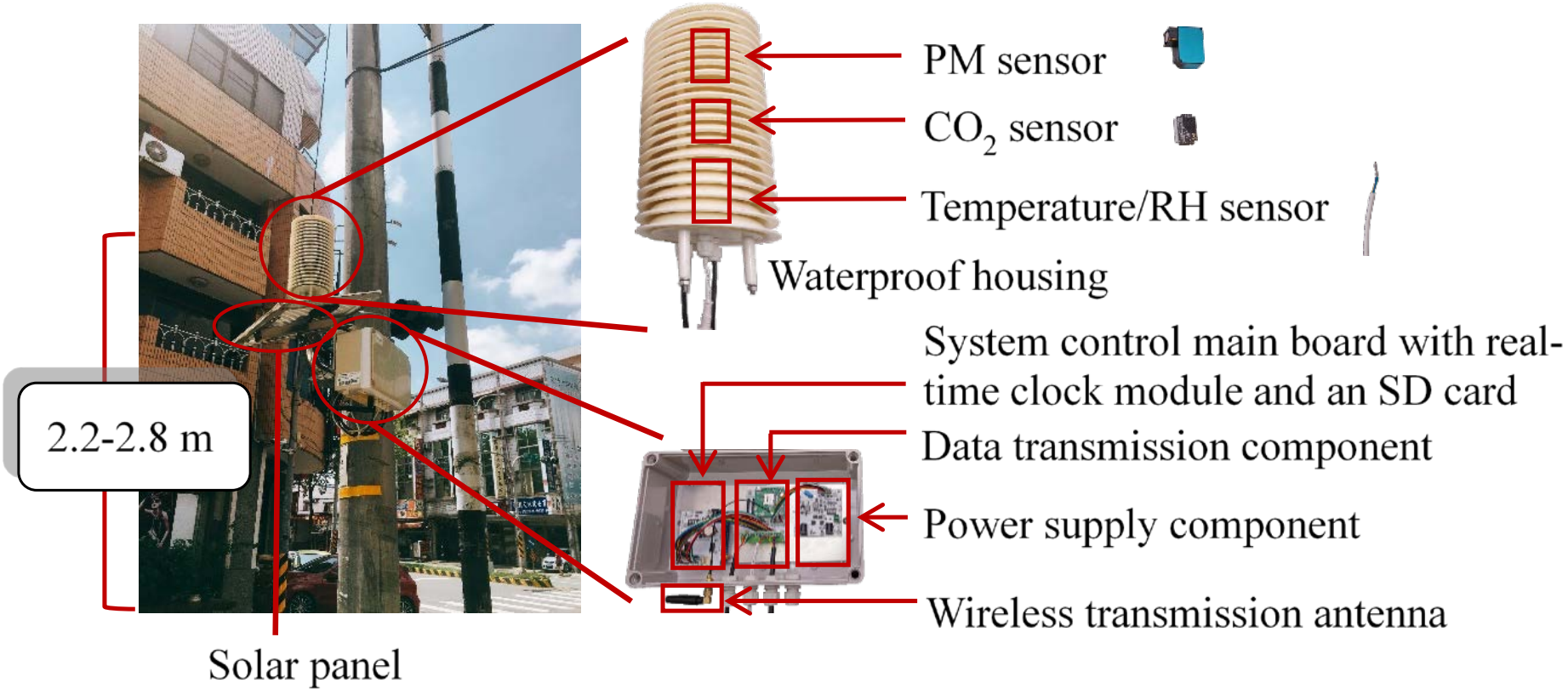
The focus of 2020 AI

- Data Quality
  - Quality Assurance and Quality Control **(QA/QC) of the environmental sensor, survey, and health data**
- What are the **sources and activities** causing **peak $PM_{2.5}$** exposures and the **controllable factors** associated with those sources and activities?
  - **Community Source Evaluation and Exposure Factor Evaluation**
- What are **the $PM_{2.5}$ damage coefficients of exposure-health relationship** of peak exposures for **heart conditions**?
  - **Exposure-health evaluation**

QA/QC of the environmental sensor data

- Improve design of AS-LUNG-O, AS-LUNG-I, and AS-LUNG-P (outdoor, indoor and portable versions, respectively)
- Detect abnormal peaks (ghost peaks) with Python and PyCharm codes
- Evaluate AS-LUNG performance to obtain correction equations for AS-LUNG
  - AS-LUNG side-by-side compared with GRIMM 1.109, a research-grade instrument, which was also compared with an USEPA federal equivalent method (FEM)
  - Wang, W.C.V.; **Lung, S.C.C.***; Liu, C.H.; Shui, C. K. (2020.06) Laboratory evaluations of correction equations with multiple choices for seed low-cost particle sensing devices in sensor networks. *Sensors,* 20(13): 3661. DOI: 10.3390/s20133661. IF: 3.257 3.531 **(sensor evaluation)**

◆ AS-LUNG-O. The left panel shows AS-LUNG-O attached to a light pole in the street; the right panel shows its various components.

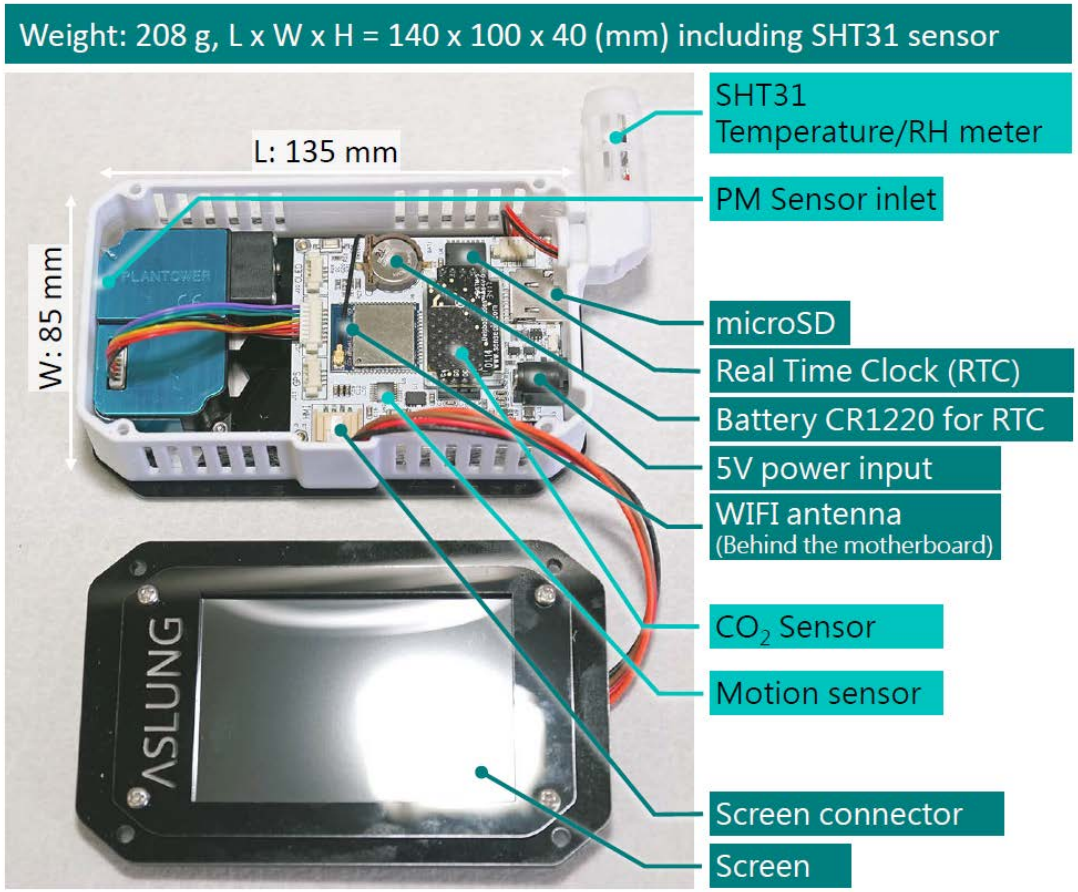


Note: Near the seashore, a sensor was eroded within three months

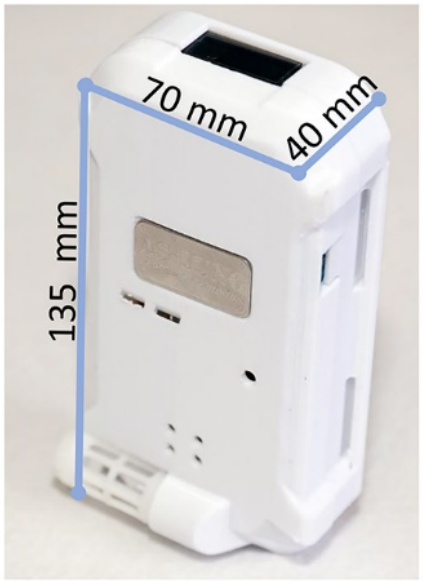
AS-LUNG-I, with various components marked

■Indoor version with a screen

■Assembled AS-LUNG-I



AS-LUNG-P
with components marked



Sensors

Supporting system

- **Weight : 153g**
- **With a mobile battery: 410g**

OLED Display

GPS Locator

Inlet

Outlet

WIFI Antenna

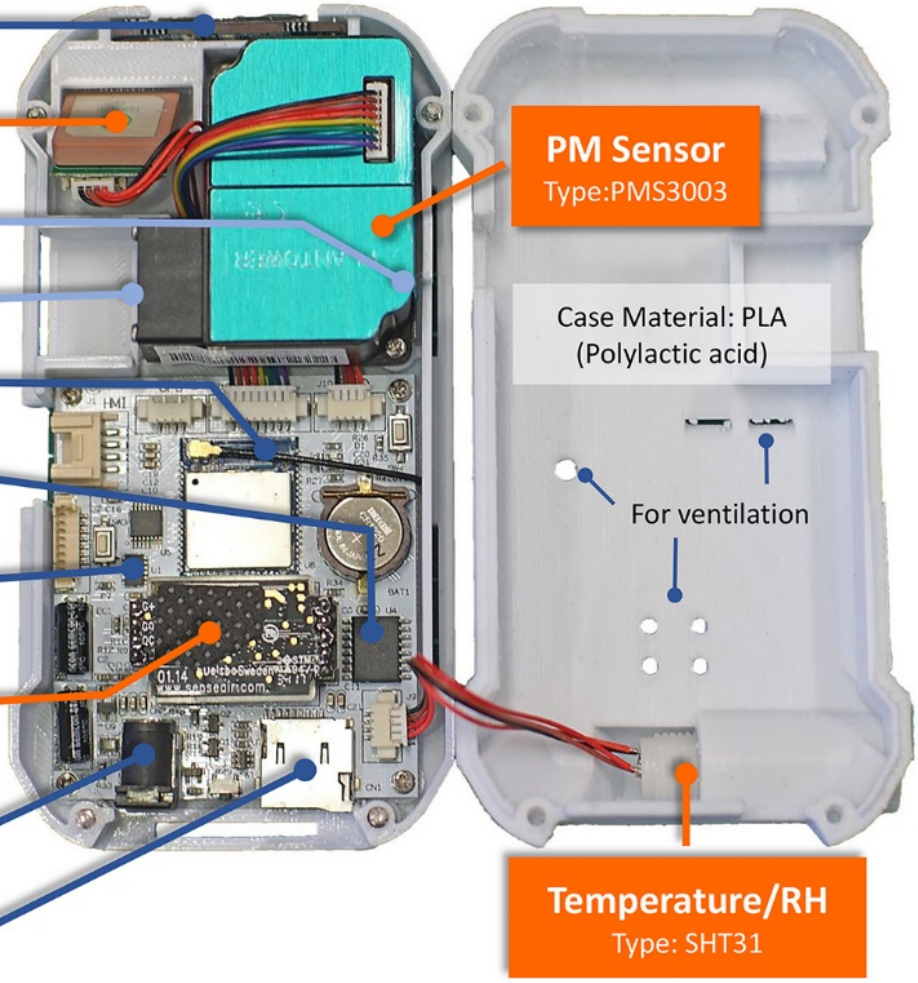
Real-Time clock
Battery Type:CR1220

G-Sensor
type:ADXL345B

CO_2 Sensor
Type : Senseair S8-0053

Power Input
Spec: 5V USB 2.0

Micro SD
Format : CSV



PM Sensor
Type:PMS3003

Case Material: PLA
(Polylactic acid)

For ventilation

Temperature/RH
Type: SHT31

◆ **Side-by-side comparisons inside**

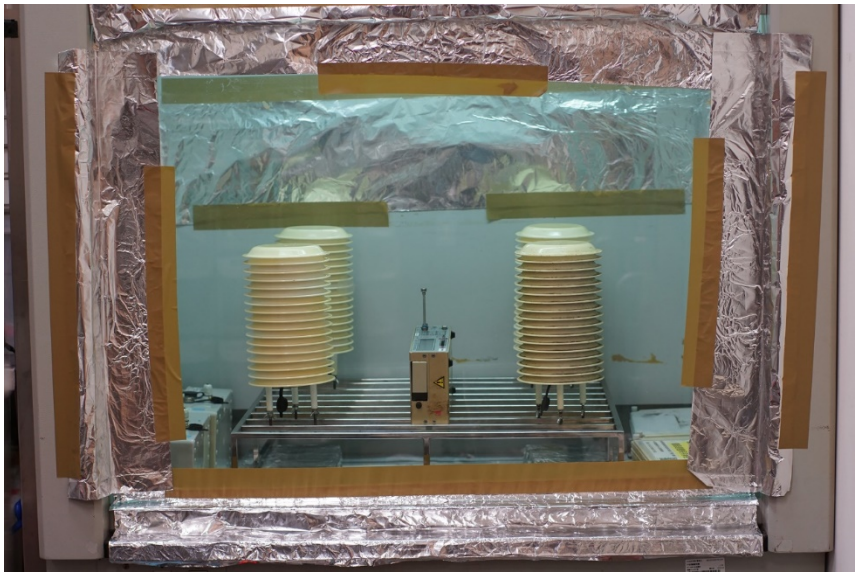
(a) a hood with 6 AS-LUNG-P sets,

(b) a hood with 4 AS-LUNG-O sets, &

(c) a chamber with 9 AS-LUNG-P sets

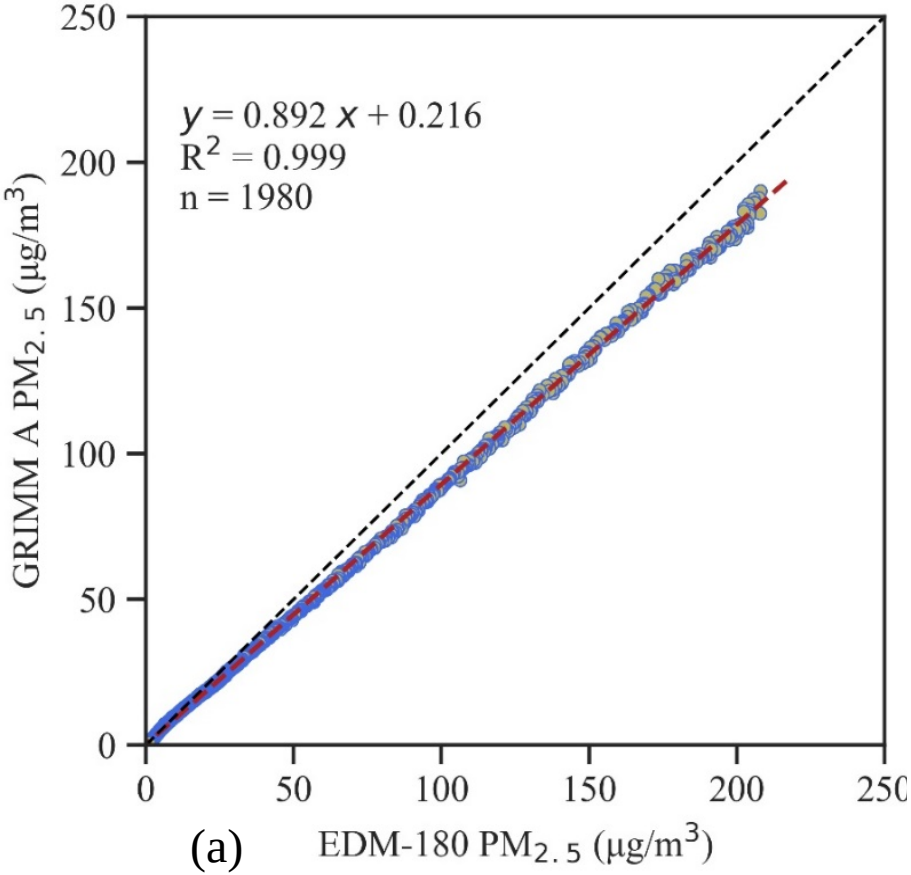
(a)

(b)

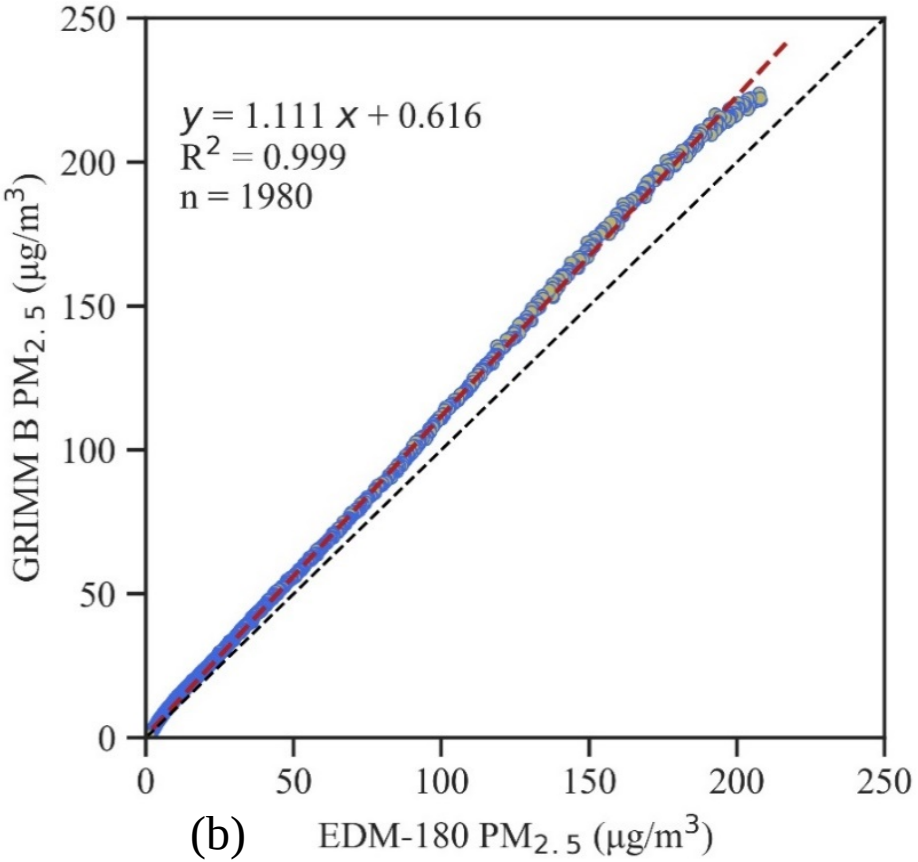


(c)

[Wang et al., 2020]



(a)



(b)

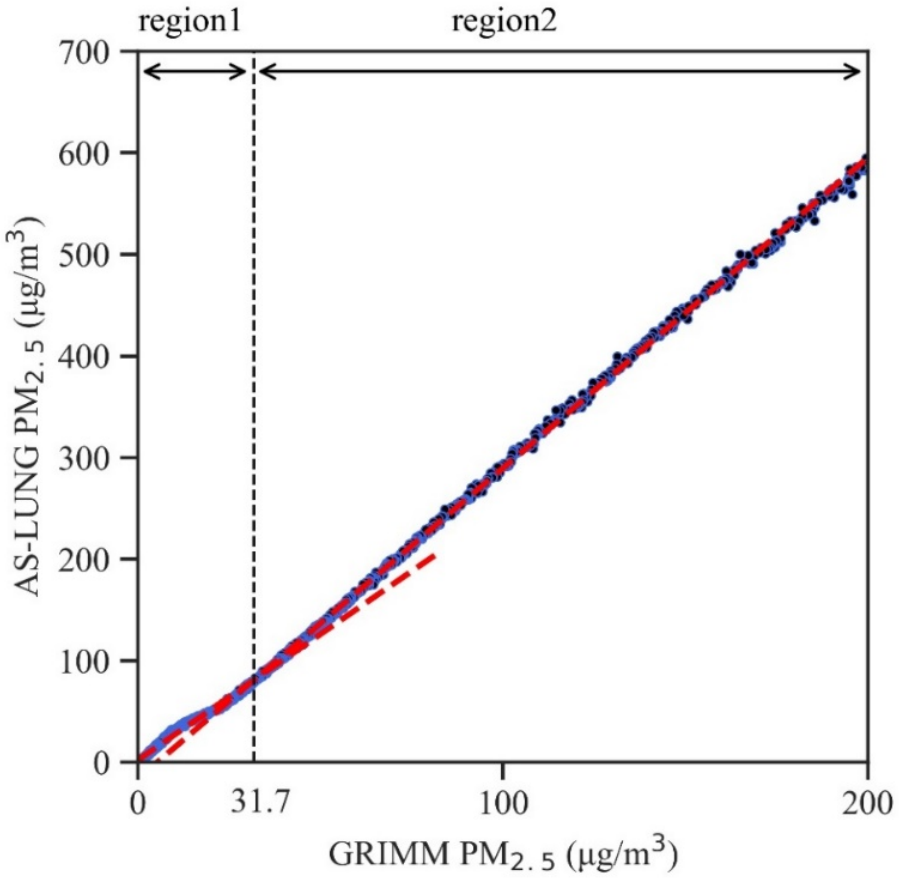
◆ Comparisons for $PM_{2.5}$ with **GRIMM 1.109 and EDM-180 (USEPA FEM)**

(a) GRIMM 1.109 (GRIMM A) under a temperature of 26.2–29.0 °C and humidity of 46–67%.

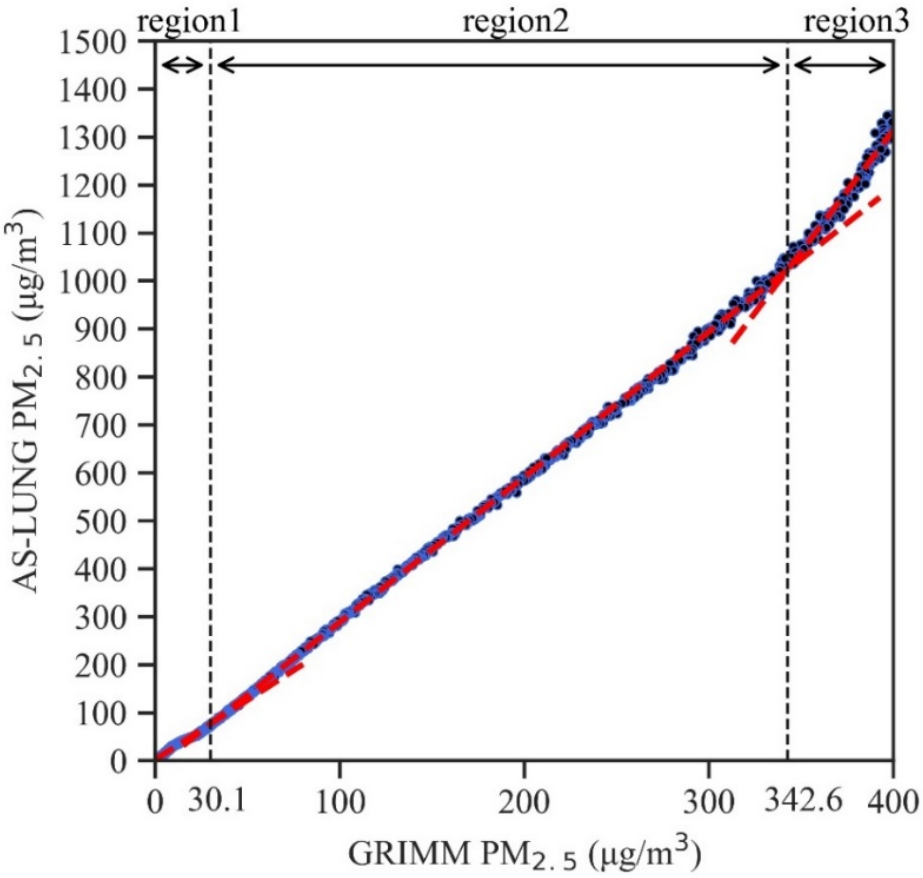
(b) GRIMM 1.109 (GRIMM B) under a temperature of 25.0–29.0 °C and humidity of 50–75%.

[Wang et al., 2020]

- Figure 4. An example of raw $PM_{2.5}$ data in the chamber experiments with incense with segmented regressions for (a) 0.1–200 μg/m³ and (b) 0.1–400 μg/m³; the regions indicated are the concentration ranges for the different regression equations.



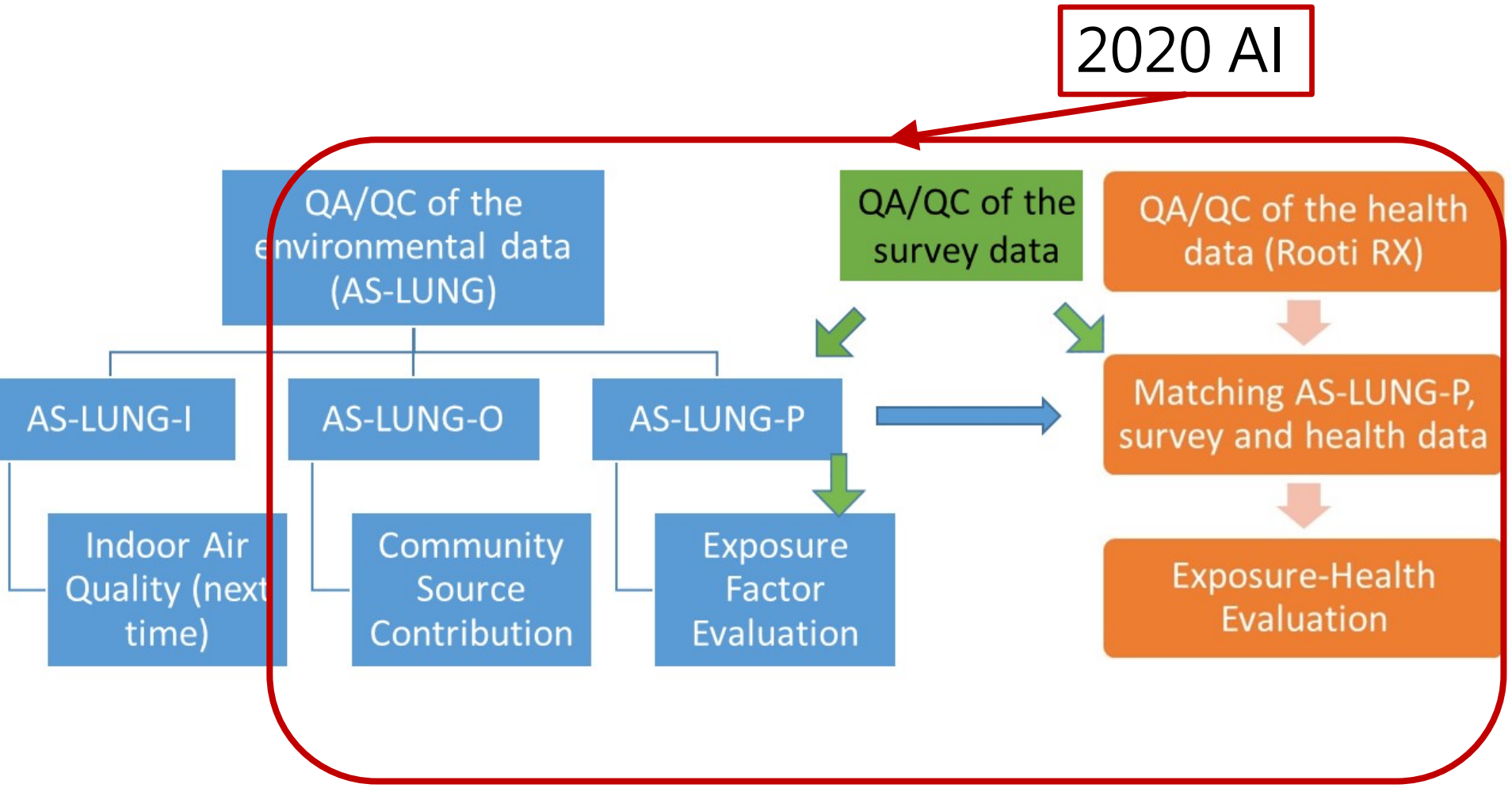
(a)



(b)

[Wang et al., 2020]

Data Flow of the Environment, Survey and Health data



Community Source Evaluation

Community Culture-related Air-Pollutant Sources in Asian Cities

$PM_{2.5}$ (WHO classified as a human carcinogen)

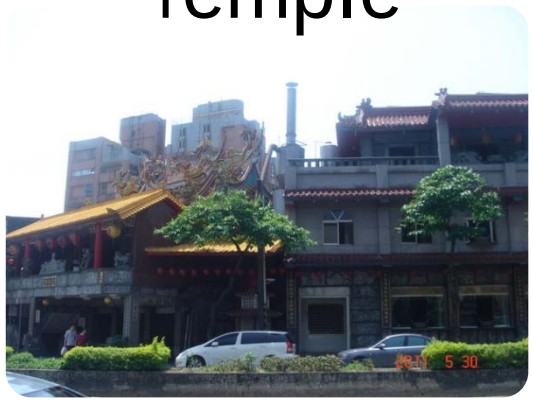
VOCs PAHs

NO_2 SO_2 many others

Asian style restaurant

Traffic

Temple



Night market

Hair salon

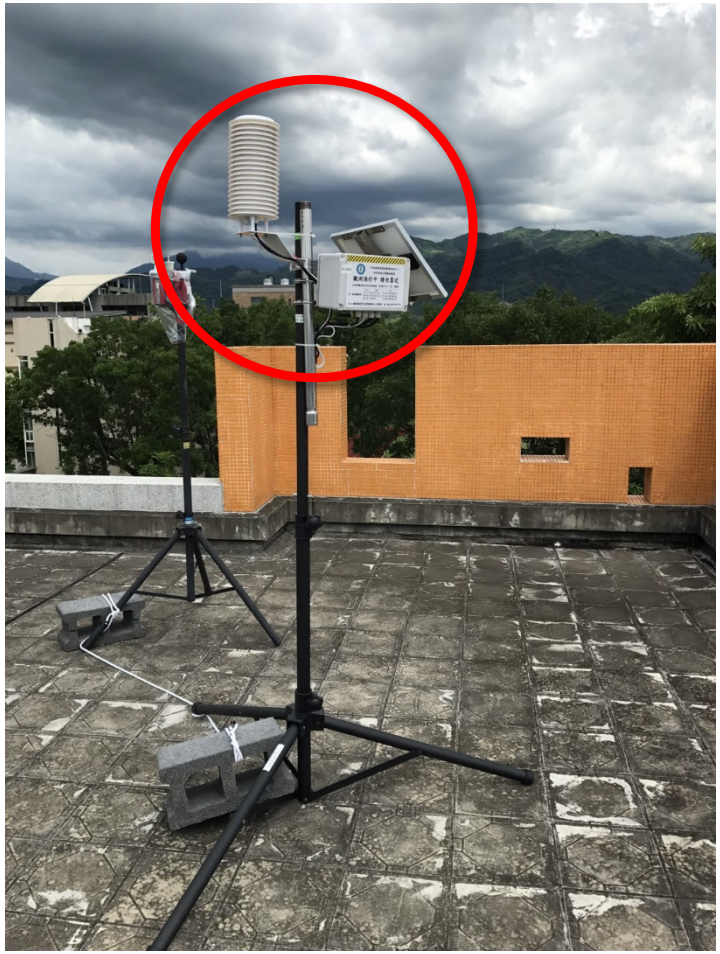
Car salon

Required Tasks in Community Monitoring in Hi-ASAP

- At least one environmental sensor should be set-up to assess $PM_{2.5}$ levels in the ambient environment close to (within 10km radius) the subjects' community
  - set-up at a 10-meter height location or at a ground location without obvious $PM_{2.5}$ source as a reference site for ambient air
- Several environmental sensors (ex. AS-LUNG-O) are set up at street sites near certain community sources (restaurants, temples, traffic within communities, home factories, biomass burning, garbage burning, etc

Application of PM sensors in assessing ambient and community levels

■High-level monitoring (AS-LUNG-O)
■For ambient monitoring



■Street-level monitoring system (AS-LUNG-O)
■For assessing incremental contribution of community sources

- Research-grade
- Water-proof
- Wireless transmission
- High time resolution
- $PM_{2.5}$, Temperature, and RH%
- In the future, other chemical sensors can be deployed

Generic Equation for Assessing Incremental Contribution of Community Sources

- $P_{street\text{-}level} = \beta 0 + \gamma \times P_{ambient} + \beta i \times$ community sources $+ \alpha i \times$ meteorological variables $+ \varepsilon$

  - $P_{street\text{-}level}$ is 5-min $PM_{2.5}$ at the street sites near certain community sources
  - $P_{ambient}$ is 5-min $PM_{2.5}$ in the ambient air (at the high-level site or at street site without any near-by sources)
  - γ, βi, αi are regression coefficients; $\beta 0$ is the intercept; and ε is an error term
  - βi is the incremental contribution of these community sources
  - meteorological variables: such as wind speed, temperature and humidity at the ambient site

One Equation Example for Community Source Incremental Contribution (1)

- $P_{location} = \beta 0 + \gamma 1 P_{high\text{-}level} + \gamma 2 Ws + \gamma 3\ (temperature) + \gamma 4\ (RH) + \sum \beta i Xi + \varepsilon$
- where $P_{location}$ is 5-min $PM_{2.5}$ at the monitoring location and $P_{high\text{-}level}$ is 5-min $PM_{2.5}$ at the high-level site
- $\gamma 1$, $\gamma 2$, $\gamma 3$, $\gamma 4$ and βi are regression coefficients; $\beta 0$ is the intercept
- Ws is the wind speed (from high-level site); and ε is an error term
- Xi is a dummy variable for pollution source. Xi of a monitoring location is assigned as 1 if there is one of those pollution sources within 3-5 m. Multiple Xi at certain locations could be assigned as 1 to account for multiple pollution sources.

[Lung et al., 2020; STE]

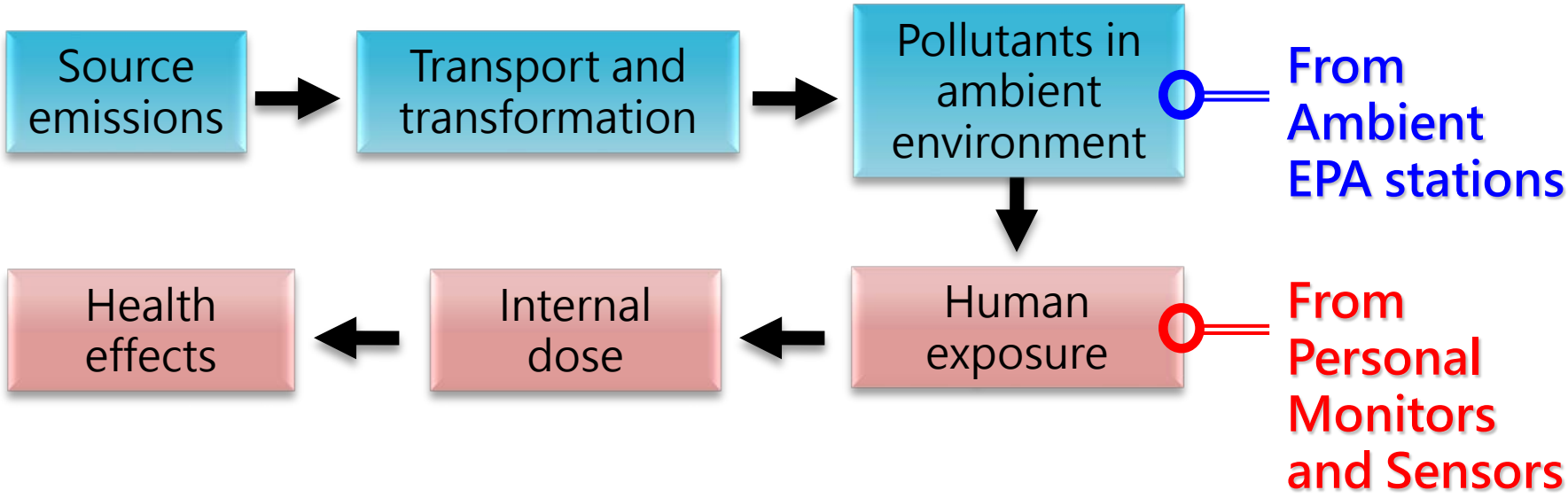
One Equation Example for Community Source Incremental Contribution (2)

- Traffic emission was further categorized as passing-by vehicles or stop-and-go. For example, if a monitoring site is near traffic emissions of stop-and-go, then dummy variable "stop-and-go" is assigned 1 and other variables, zero
- only observations between 6 am to 10 pm were included in the analysis since almost all emission sources dwindled in nighttime
  - Cooking emission was assigned 1 while the vendor/restaurant was in operation (4 pm-10 pm)
  - Traffic near school was assigned 1 only on school days at hours when students enter/leave school
- The street-background location was considered but not included in the model to avoid collinearity
- Multiple regression or Stepwise regression was applied excluding raining hours

[Lung et al., 2020; STE]

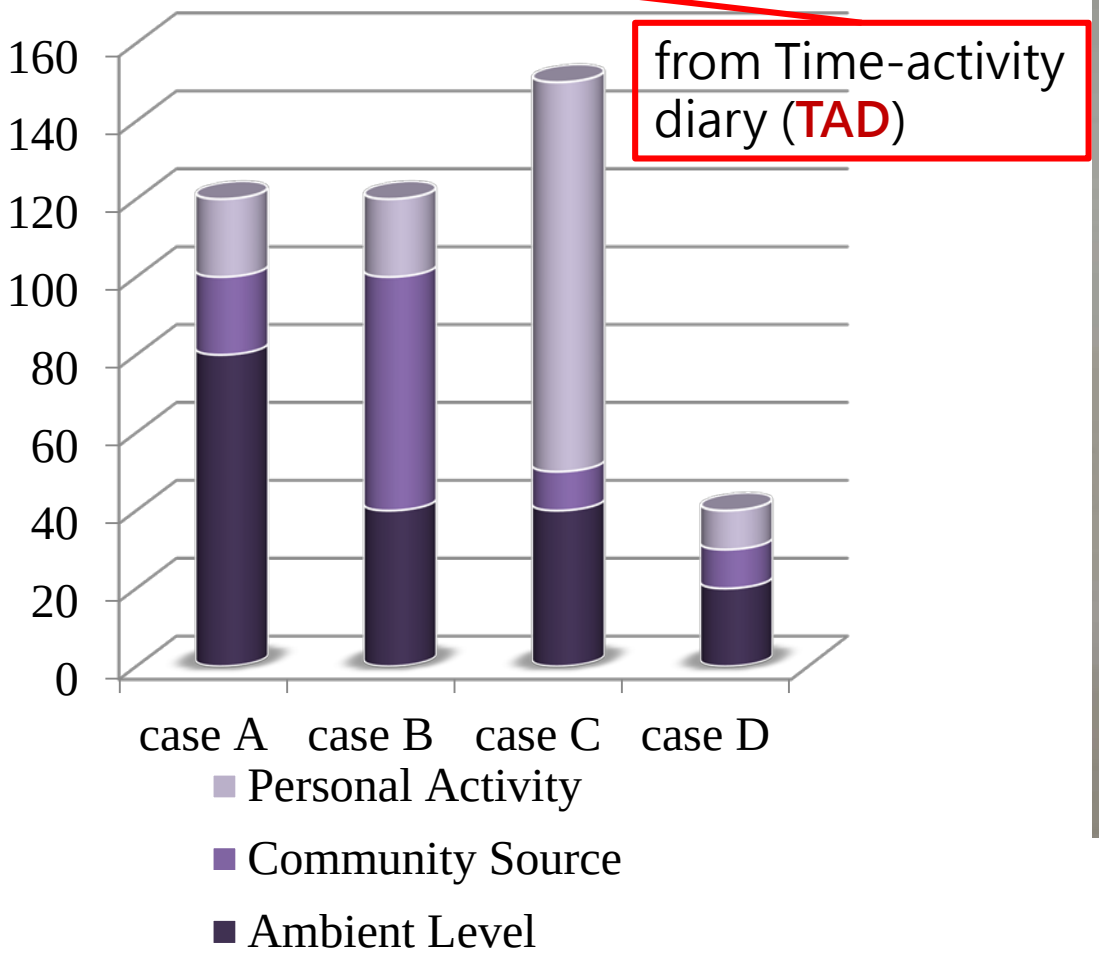
Exposure Factor Evaluation & QA/QC on survey data

Progression of Pollutants from Emissions to Health Effects



Personal Exposure Levels

Exposure Level =**Ambient** level
+ **Community source** contribution
+ **Personal activity** contribution



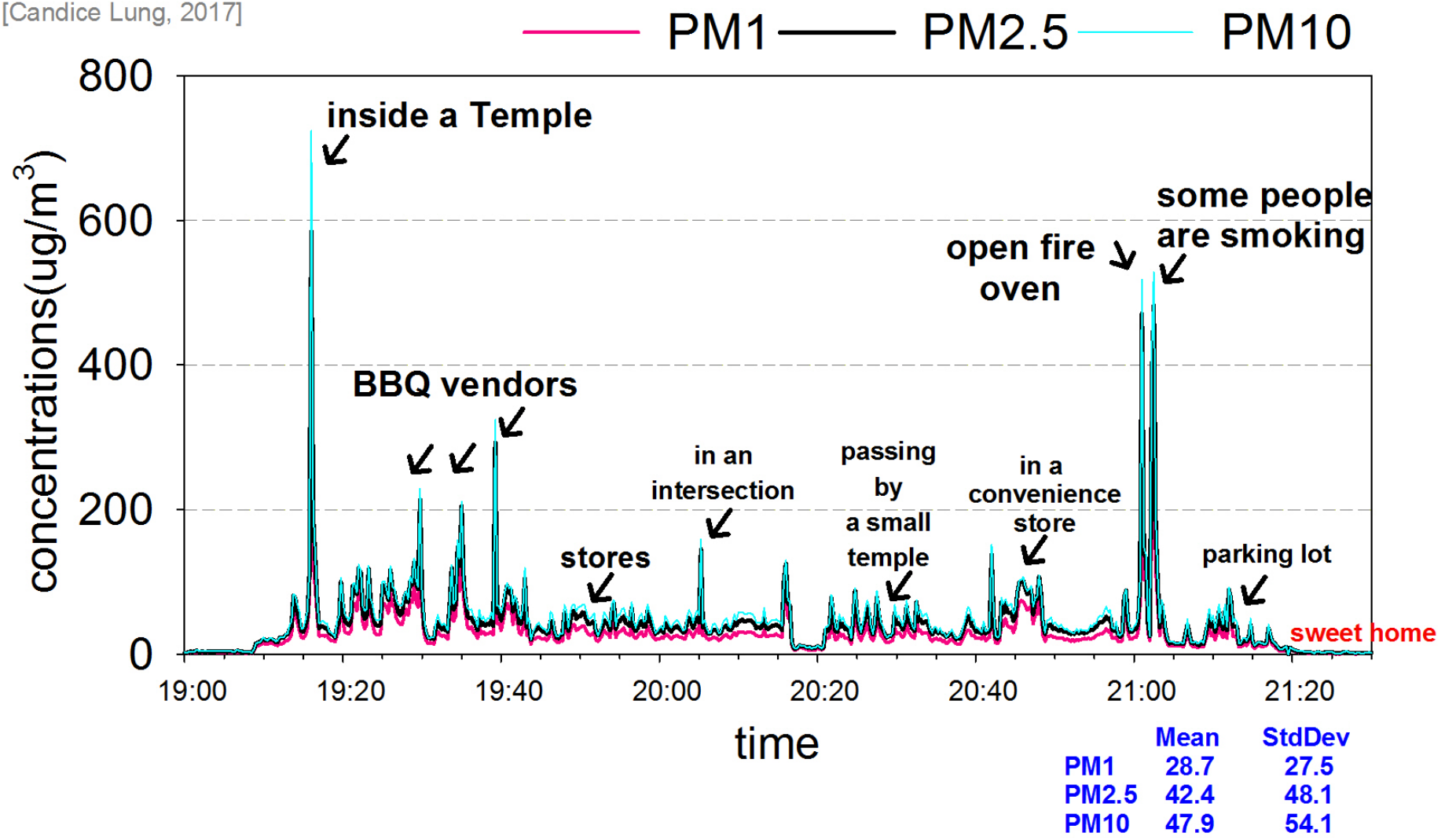
AS-LUNG-P with GPS and wifi

[Candice LUNG, 2014]

AS-LUNG
Academia Sinica.TAIWAN

[Candice Lung, 2017]

July 4, 2017, Taipei night market
AS-LUNG sampling rate: every 15 seconds



	Mean	StdDev
PM1	28.7	27.5
PM2.5	42.4	48.1
PM10	47.9	54.1

Questionnaire B – Time-Activity Diary

Date: [continental time]:20 / / (yyyy/mm/dd)

Name:

Weather: □Sunny □Cloudy □Rainy □Windy

Items / Time (hhmin)	A. Location	B. Ventilation status	C. Activity type		D. Air quality		
			Activity 1	Activity 2	Source 1	Source 2	Weather
1800-1830	18	9	8	6	3	sticky tofu	S, W
Please set your watch first and then write down your time-activity information below !							

【Part A:selection card】

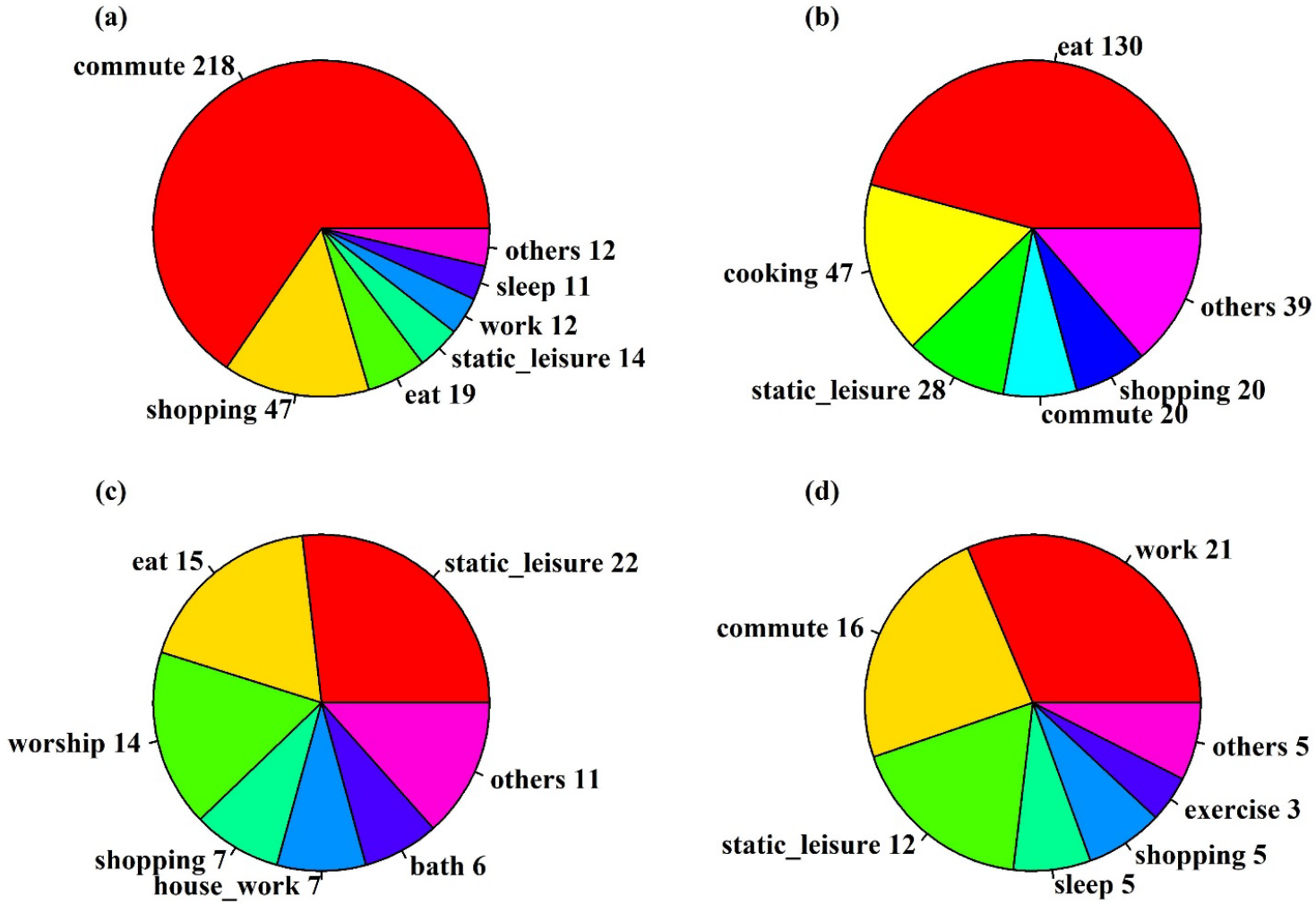
[Vehicles]

9. Bike, electric scooter
10. Motorcycle, electric locomotive
11. Car, truck, taxi
12. Bus
13. Underground; tube; subway
14. Train
15. Farm, livestock oriented machine car
16. Walking, wheelchair
62. Other vehicles, please write down

[Outdoor] hall, road, outdoors, open ground

16. On the road(walking on the road, waiting for bus or traffic light)
17. Traditional market
18. Night market, fairgrounds, school carnival
19. Small market under ten vendors
20. Sports field (playgrounds, golf course, etc.)
21. Parks, Scenery spot, farm and other outdoor places
22. Home exteriors (balcony or loft)
23. Outdoor workplace
63. None of the above, please write down other outdoor locatio

Activities with exposure to (a) vehicle emission (n = 333), (b) cooking (n = 284), (c) incense burning (n = 82), and (d) environmental tobacco smoke (n = 67) according to TAD responses at 30-min intervals from 33 subjects



[Lung et al., 2020; JESEE]

Generic Equation for Assessing Incremental Contribution of Exposure Sources

- $\mathrm{PM}_{\mathrm{personal}} = \alpha_0 + \sum \beta i Xi + \gamma_2 \mathrm{PM}_{\mathrm{outdoor}} + \sum \zeta i Vi$

  - $\mathrm{PM}_{\mathrm{personal}}$ and $\mathrm{PM}_{\mathrm{outdoor}}$ are personal and outdoor $\mathrm{PM}_{2.5}$ levels, respectively
  - α_0 is the intercept, and βi is the regression coefficient of Xi, which is a dummy variable representing different sources recorded in TADs, with no-recorded-source as the base case.
    - Sources encountered less than 100 times (sample size of the total valid 30-min records of 26 subjects is 9350) were not incorporated into the model in order to focus on significant ones
  - γ_2 are regression coefficients of outdoor $\mathrm{PM}_{2.5}$ levels
  - ζi is the regression coefficient of Vi, a dummy variable of ventilation statuses
    - Typical ventilation statuses in the Taiwanese households included windows open without air-conditioning (AC) (hereinafter window-open, the base case), windows closed without AC (hereinafter window-closed), and windows closed with AC on (hereinafter AC-on)
  - Stepwise regression was used

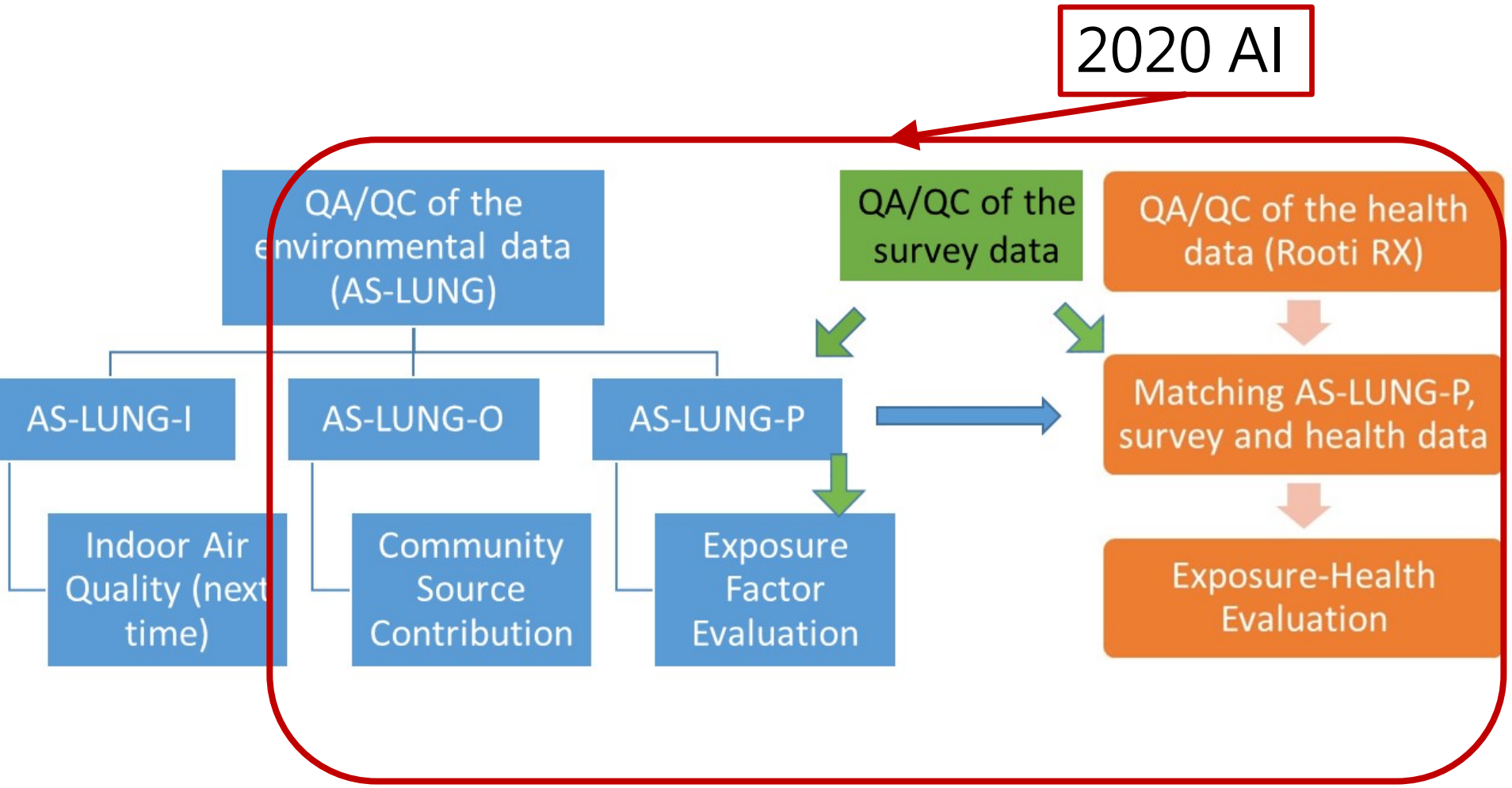
[Lung et al., 2020; Indoor Air]

Required Questionnaire for Hi-ASAP

- For international comparison, core questions should be the same for all subjects across different areas. Core questions will be provided by our group which should be translated into local languages. Please translate the core questions based on faithfulness, expressiveness, and elegance.
- You may design more questions based on knowledge of local $PM_{2.5}$ sources or customs. But do keep the questionnaire not too long. Otherwise, your subjects will lose patience and give your random answers
- When designing questions, always think how I analyze these questions with environment data to fit our research purposes; be sure to use plain languages for your subjects
- Think about the possibility to collaborate with scientists in other fields using the same questions or combining other datasets
- Keep in mind the time-resolution of questionnaire and environment data

[from 2019 AI on Hi-ASAP]

Data Flow of the Environment, Survey and Health data



QA/QC of the health data & Exposure-health Evaluation

(panel studies assessing heart rate variability (HRV) with Rooti RX)

Exposure-Health Evaluation

- The association between $PM_{2.5}$ and log10-transformed HRV indicators with 5-min resolutions was analyzed using the general additive mixed model (GAMM, R Version 3.5.0)
  - **Lung, S. C. C.***; Chen, N.; Hwang, J. S. Hu, S.C.; Wang, W.C.V.; Wen, T.Y.J.; Liu, C.H. (2020.08) Panel study using novel sensing devices to assess associations of $PM_{2.5}$ with heart rate variability and exposure sources. *Journal of Exposure Science and Environmental Epidemiology.* DOI: 10.1038/s41370-020-0254-y. IF: 3.531 **(exposure-health evaluation)**
- **Basic concept: HRV = $PM_{2.5}$ + other factors** (some with linear relationship with HRV and some with non-linear relationship with HRV)
  - GAMM can adjust for individual differences and autocorrelation of HRV signals

The focus of 2020 AI

- Data Quality
  - Quality Assurance and Quality Control **(QA/QC) of the environmental sensor, survey, and health data**
- What are the **sources and activities** causing **peak $PM_{2.5}$** exposures and the **controllable factors** associated with those sources and activities?
  - **Community Source Evaluation and Exposure Factor Evaluation**
- What are **the $PM_{2.5}$ damage coefficients of exposure-health relationship** of peak exposures for **heart conditions**?
  - **Exposure-health evaluation**

Any question?

Shih-Chun Candice LUNG 龍世俊

- Research Center for Environmental Changes, Academia Sinica, Taiwan
- Center for Sustainability Science, Academia Sinica, Taiwan

Resources

- Wang, W.C.V.; **Lung, S.C.C.***; Liu, C.H.; Shui, C. K. (2020.06) Laboratory evaluations of correction equations with multiple choices for seed low-cost particle sensing devices in sensor networks. *Sensors,* 20(13): 3661. DOI: 10.3390/s20133661. IF: 3.257 3.531 **(sensor evaluation)**
- **Lung, S. C. C.***; Wang, W.C.V.; Wen, T.Y.J.; Liu, C.H.; Hu, S.C. (2020.05) A versatile low-cost sensing device for assessing $PM_{2.5}$ spatiotemporal variation and quantifying source contribution. *Science of the Total Environment*,716. DOI: 10.1016/j.scitotenv.2020.137145. IF: 6.551 3.531 **(community source evaluation)**
- **Lung, S. C. C.***; Chen, N.; Hwang, J. S. Hu, S.C.; Wang, W.C.V.; Wen, T.Y.J.; Liu, C.H. (2020.08) Panel study using novel sensing devices to assess associations of $PM_{2.5}$ with heart rate variability and exposure sources. *Journal of Exposure Science and Environmental Epidemiology.* DOI: 10.1038/s41370-020-0254-y. IF: 3.531 **(exposure-health evaluation)**
- Sinaga, D.; Setyawati, W.; Cheng, F.Y.; **Lung, S. C. C.*** (2020.08) Investigation on daily exposure to $PM_{2.5}$ in Bandung City, Indonesia using low-cost sensor. *Journal of Exposure Science and Environmental Epidemiology*. DOI: 10.1038/s41370-020-0256-9. IF: 3.531 **(exposure factor evaluation)**