

# AirCloud: A Cloud-based Air-Quality Monitoring System for Everyone

Yun Cheng, Xiucheng Li, Zhijun Li, Shouxu Jiang,  
Yilong Li, Ji Jia, and Xiaofan (Fred) Jiang

Air Scientific Inc.  
Harbin Institute of Technology



# Motivation and Challenges: For Individuals



# Motivation and Challenges: For Individuals



Not the air you  
breathe in

Expensive

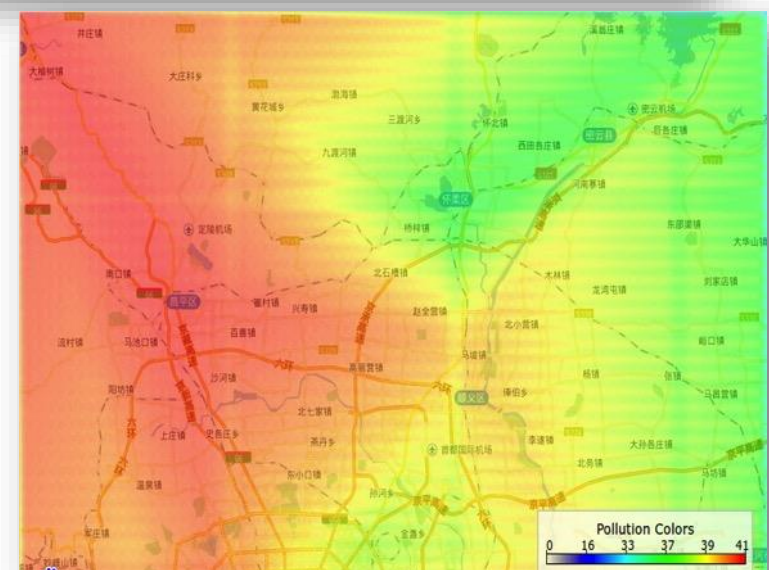
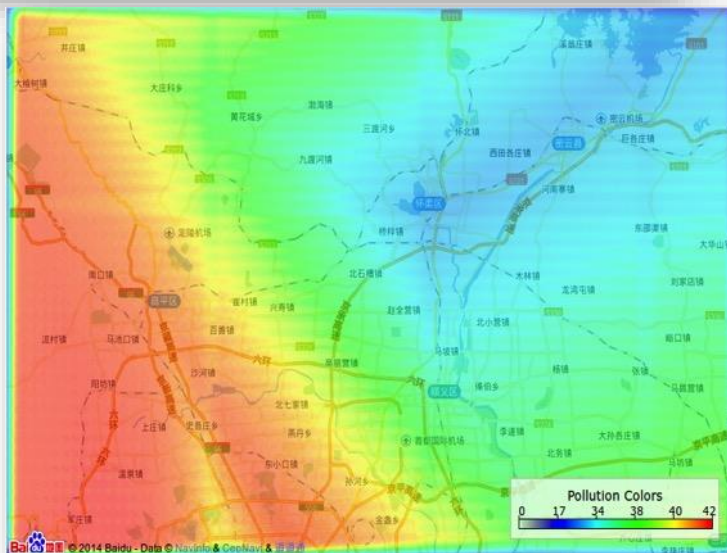
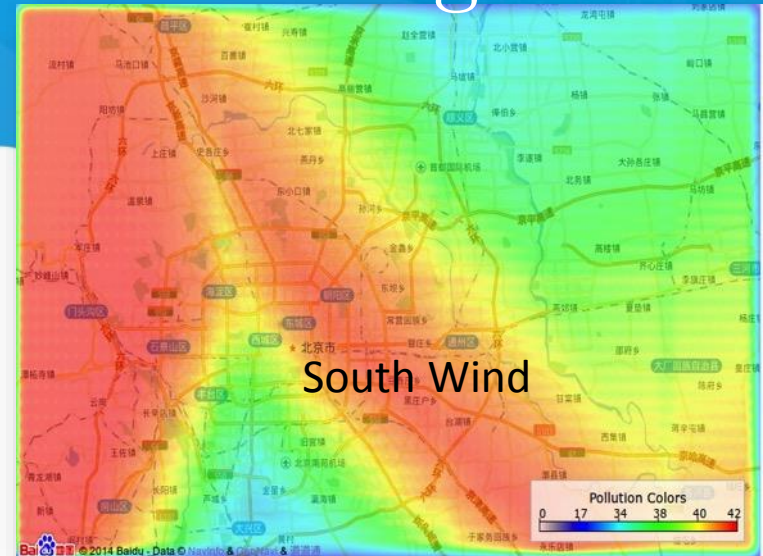
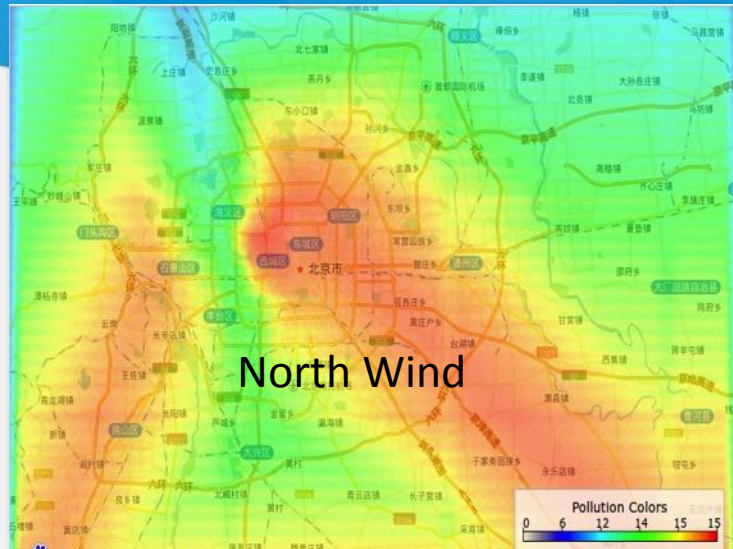


Not accurate





# Motivation and Challenges: For Government and Orgs



# AirCloud System

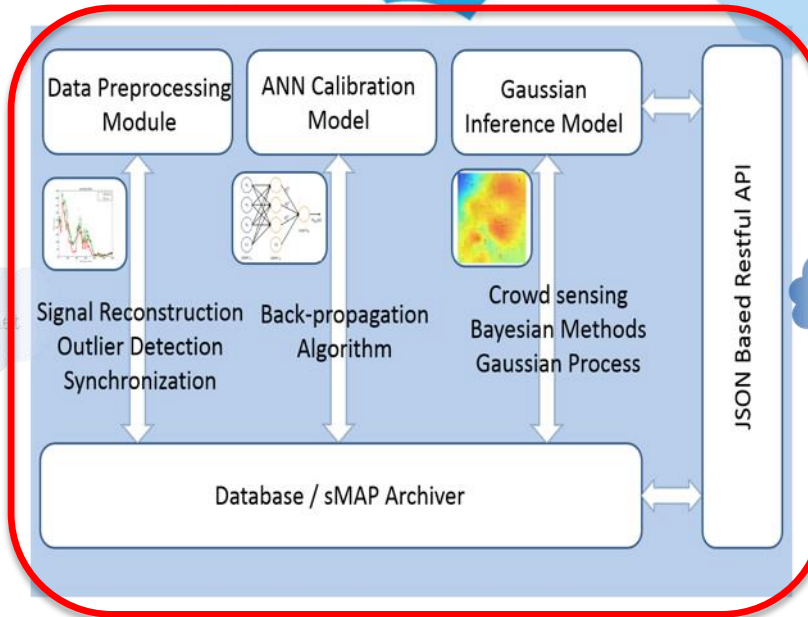
## A Client-Cloud Approach

Data Sources



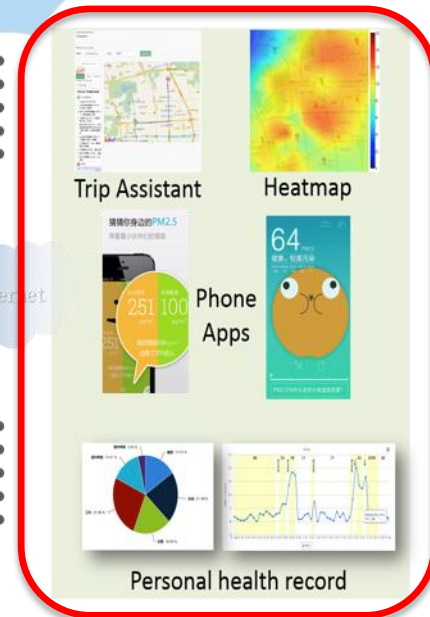
**DATA COLLECTION**  
Low Cost  
Internet-Connected  
Sensor Front-end

AQM Cloud



**ONLINE AIR-QUALITY  
ANALYTICS ENGINE**  
Single-point Calibration  
Multi-point Prediction

Applications



**SERVICES AND  
APPLICATIONS**

# The Client Hardware

custom-designed devices



**AQM**

Stationary Version

Ethernet/GPRS -> backend



**miniAQM**

Portable Version

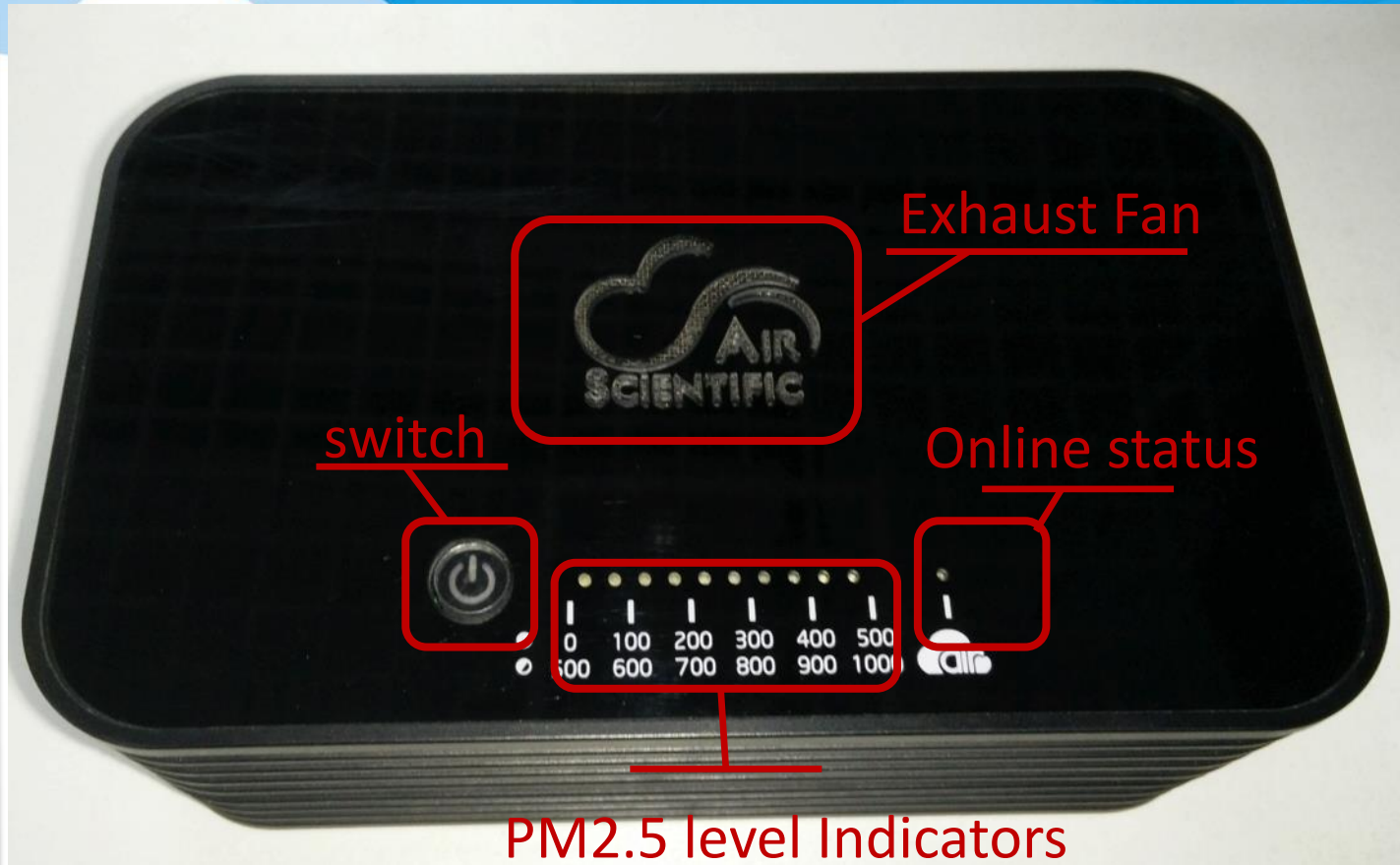
BLE4.0->Phone->backend



# The Client Hardware

## mechanical structure

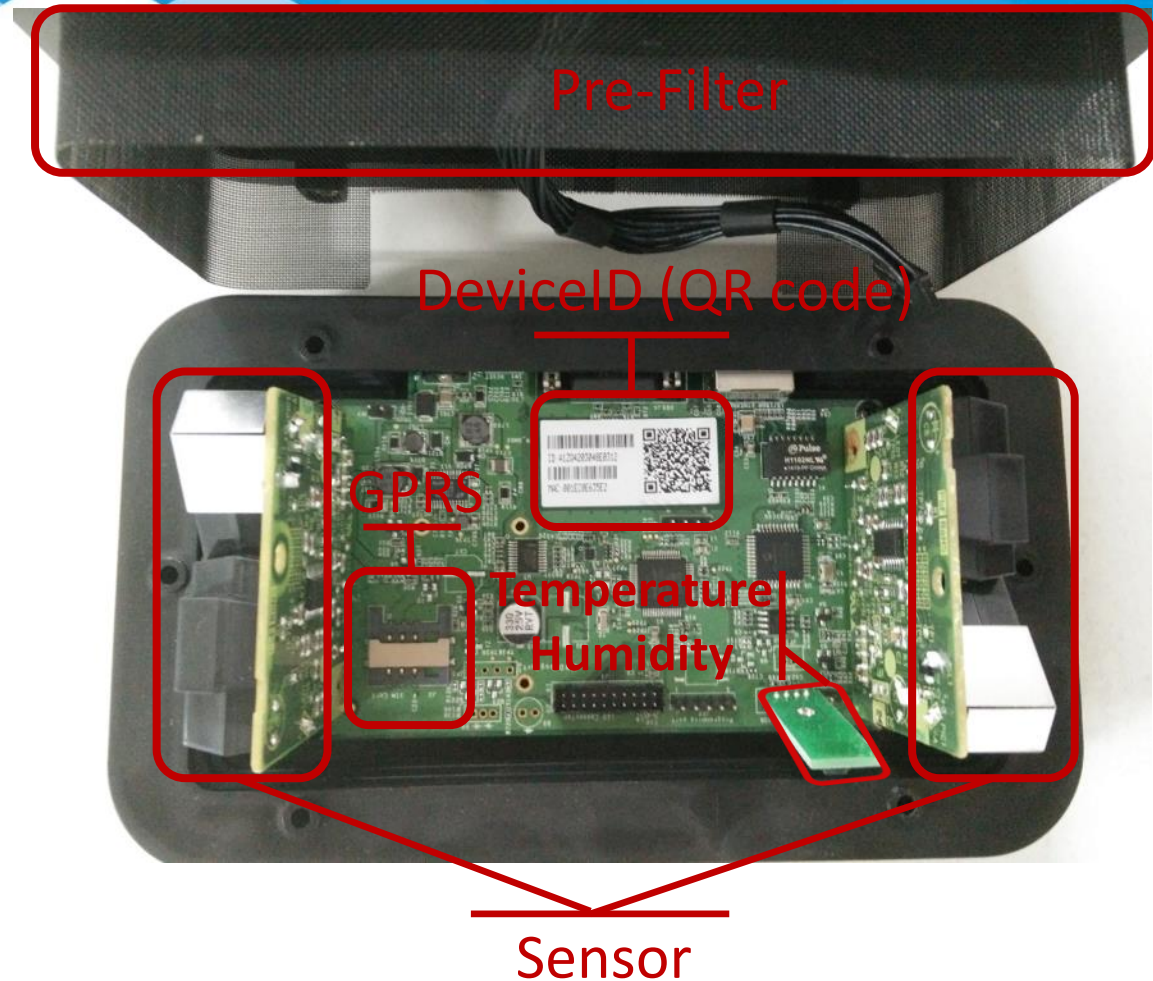
front



# The Client Hardware

## mechanical structure

inside



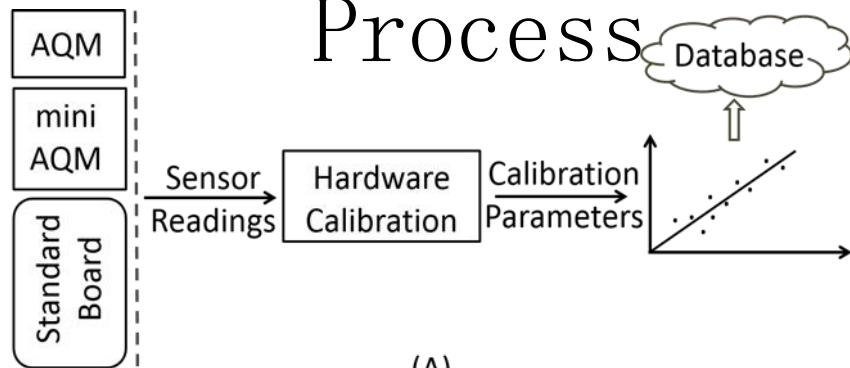


# Hardware Calibration

reduce variations between sensors

## Calibration

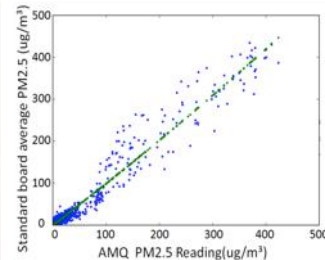
### Process



(A)



(B)



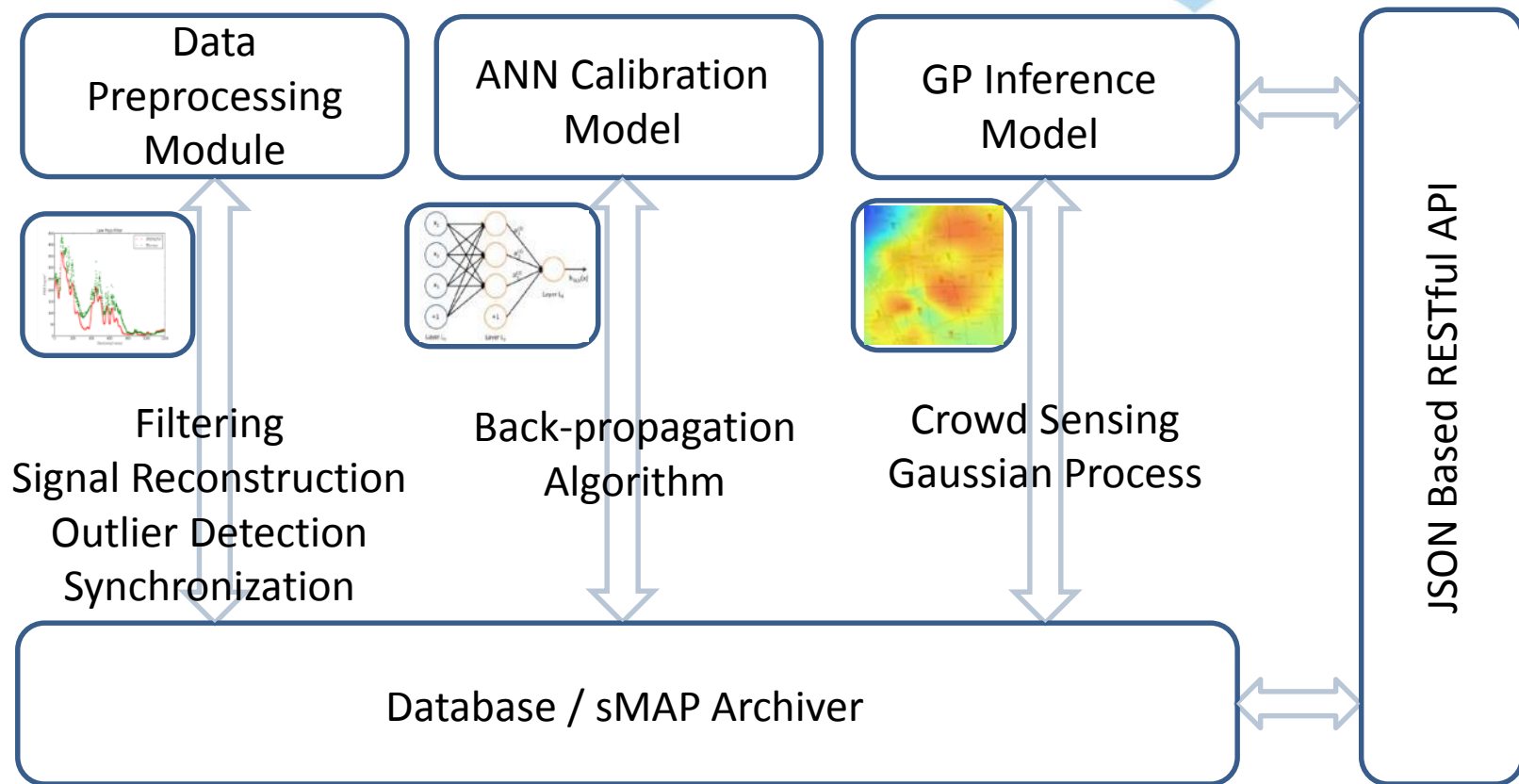
(C)

## Air Chamber



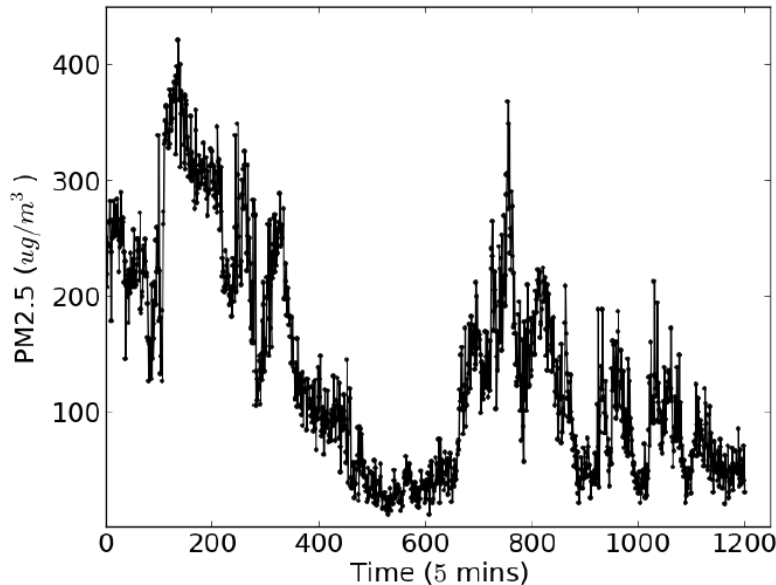
# Online Air-Quality Analytics Engine

## The whole Picture



# Air-Quality Analytics Engine

## Signal Reconstruction Module



$$x_{cor} = x + v$$

$$\text{minimize } \|\hat{x} - x_{cor}\|^2 + \delta \|D\hat{x}\|^2$$

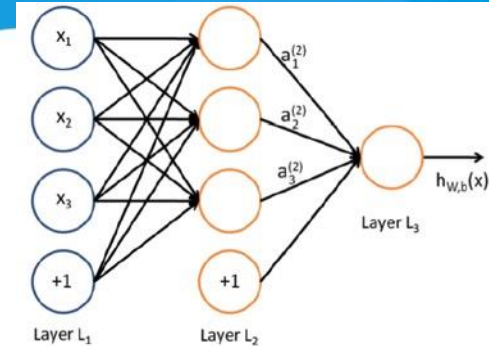
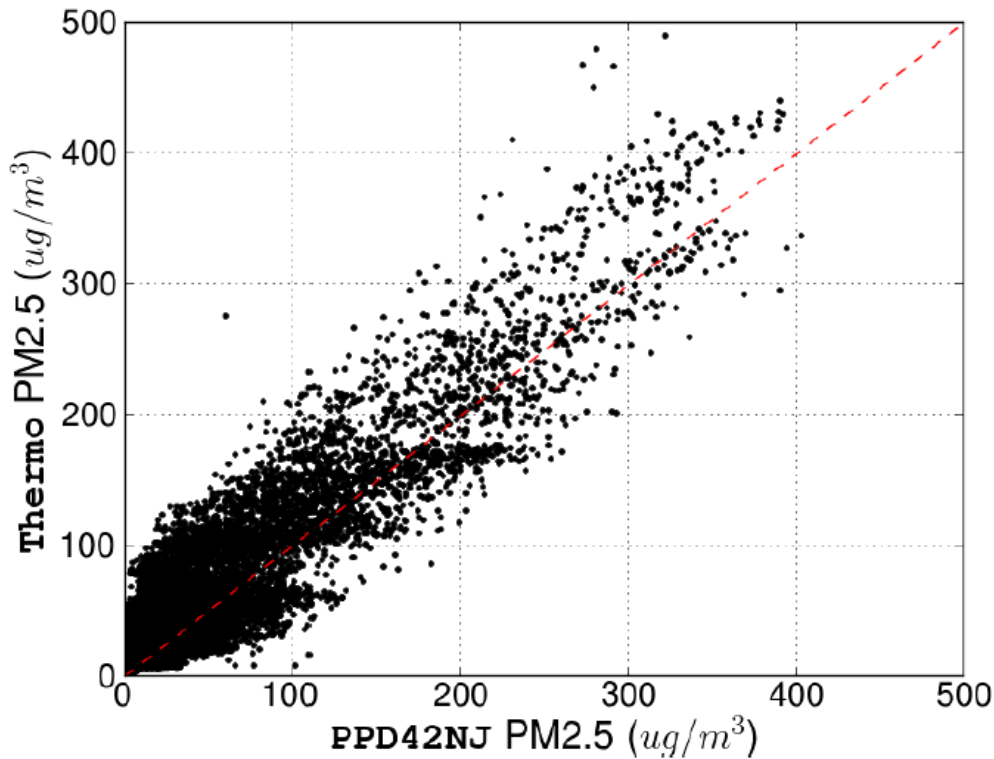
$$\hat{x} = (I + \delta D^T D)^{-1} x_{cor}$$

De-noise and smooth the  
corrupted original signal



# Air-Quality Analytics Engine

## ANN Calibration Module

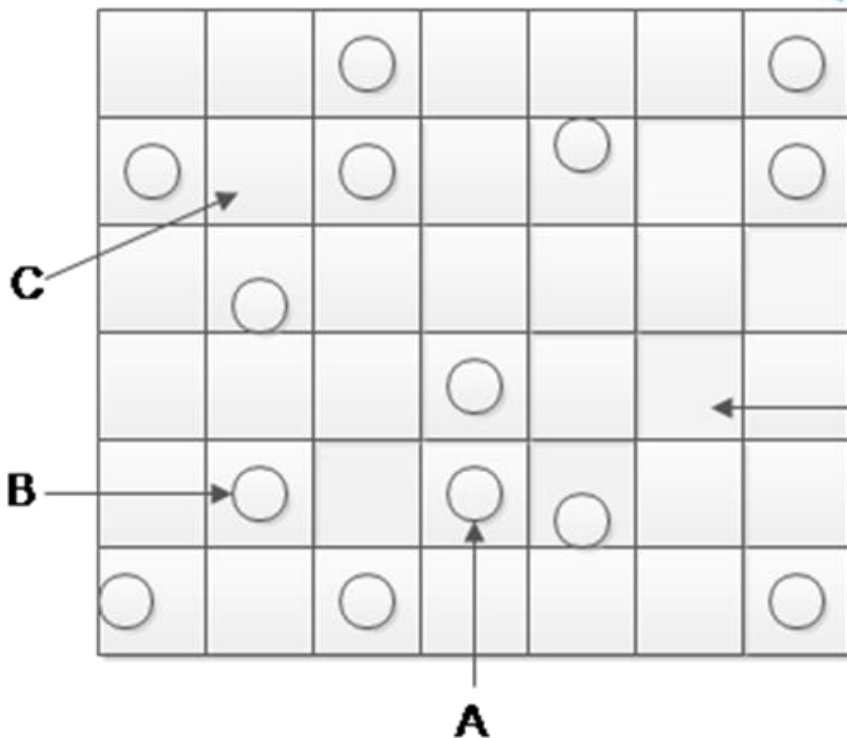


- [PM2.5\_obs, Temp, Humidity]  $\rightarrow$  PM2.5\_true
- Online stochastic gradient descent

Online multi-modal calibration (same location)

# Air-Quality Analytics Engine

## Online Inference Module



Map of Physical Space

Feature vector

$x$ : [distance, {POI}, Temp, Humidity]

Similarity specified by kernel function  $K$ :

$$k_{SE}(x, x') = \exp \left( - \sum_{i=1}^n \frac{(x_i - x'_i)^2}{2w_i^2} \right)$$

$w_i$  specifies importance of the feature  $i$

Compute covariance matrix:

$$K_y = K(X, X) + \text{diag}(\delta_i^2)$$

Smaller  $\delta_i$   $\rightarrow$  higher confidence

Predict PM2.5:

$$\mu_* = K_*^T K_y^{-1} y$$

**Infer/predict PM2.5 based on other data sources (different location)**

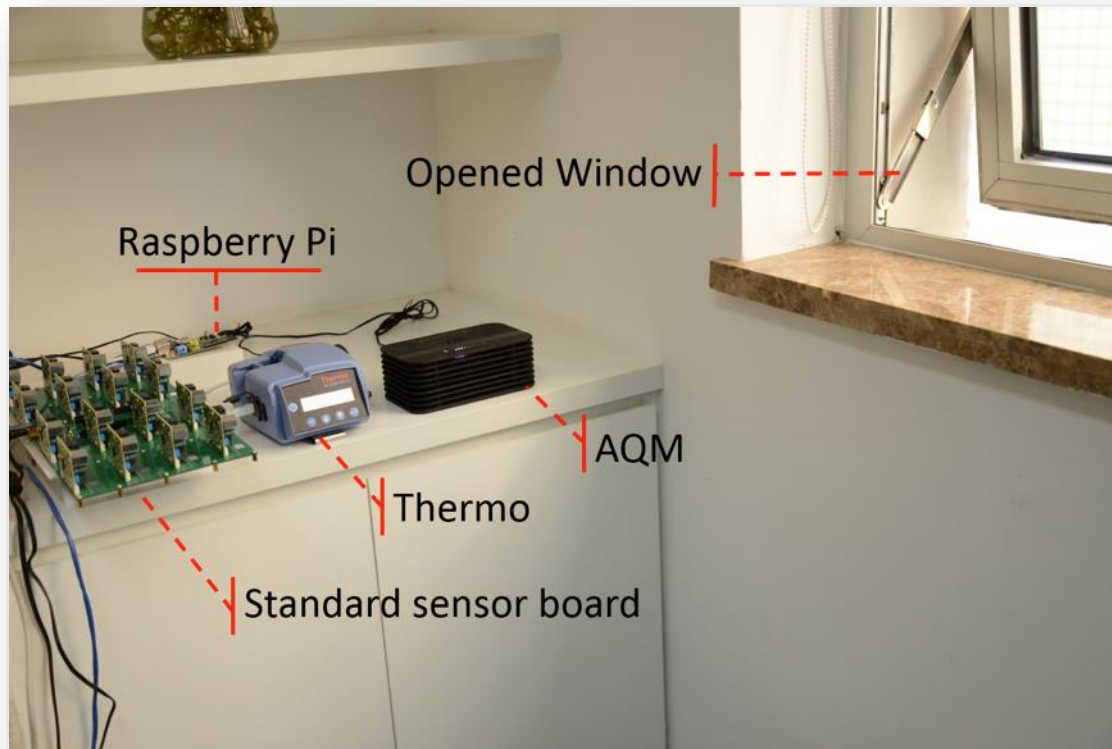
# Evaluation and Results





# Evaluation and Results

## Setup and datasets

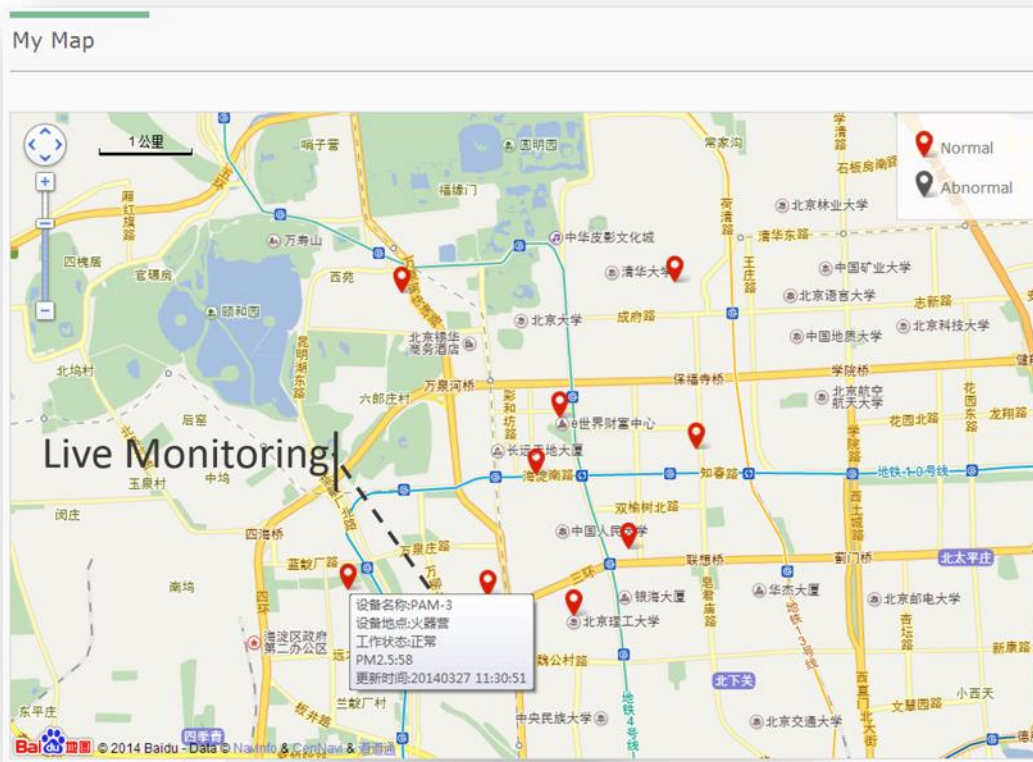


### Historical Dataset:

- Training and Verifying the ANN model
- 5 months data, sample once a minute
- Training dataset (70%)  
testing dataset (30%)

# Evaluation and Results

## Setup and datasets

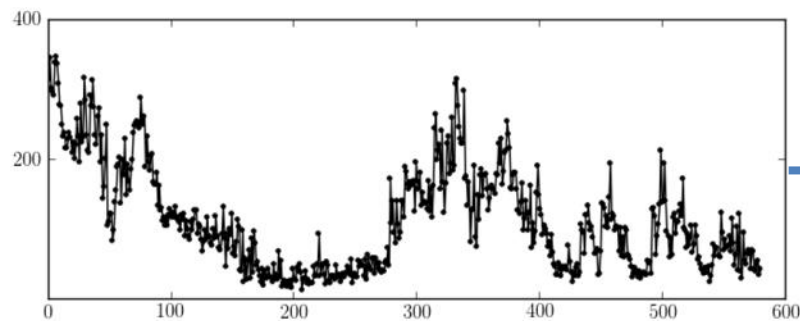


### Deployment Dataset:

- Verifying the ANN and GP inference model
- 2 months data, sample once a minute
- 10 AQMs + 2 miniAQMs + 3 Dylos + 2 Thermo
- 4km \* 4km which covers all kinds of environments (POI)

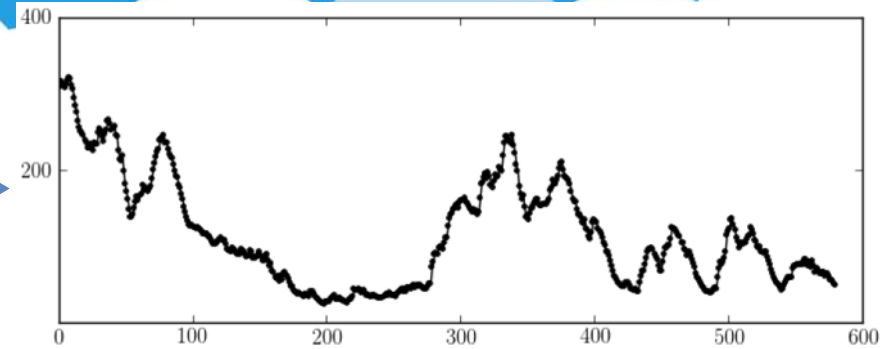
# Evaluation and Results

## Signature Reconstruction

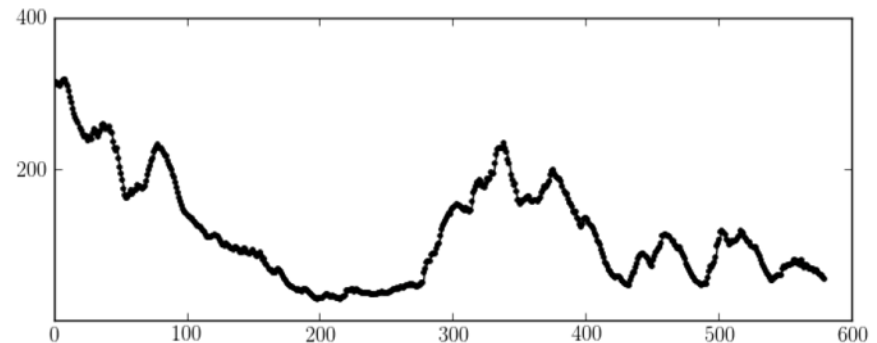


$\delta = 0$   $RMSE = 208.59$

Signal Reconstruction



$\delta = 25$   $RMSE = 96.98$



$\delta = 50$   $RMSE = 96.69$



# Evaluation and Results

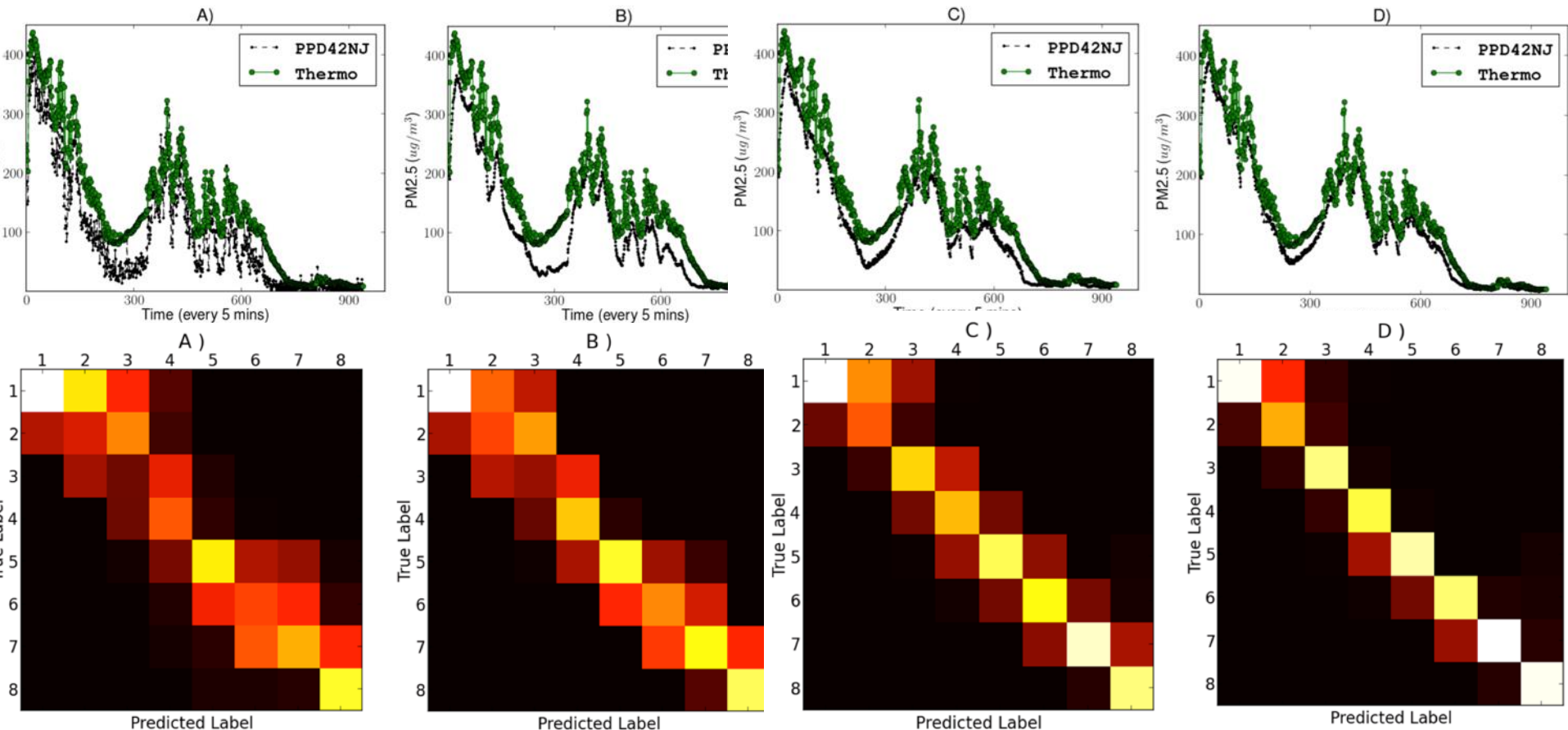
## definition of PM2.5 levels

$PM_{2.5}(ug/m^3)$	Levels	Values Levels of Health Concern
0-35	1	Good
35-75	2	Moderate
75-115	3	Unhealthy for sensitive groups
115-150	4	Unhealthy
150-200	5	Highly unhealthy
200-250	6	Very unhealthy
250-350	7	Hazardous
>350	8	Severely Hazardous

- Level 1-4 is the same as the official one
- Level 5 and 6 equal to official 5
- Level 6 and 7 equal to official 6

# Evaluation and Results

## SR + ANN + GP Online Interference (deployment dataset)



Raw Data

Signal  
Reconstruction

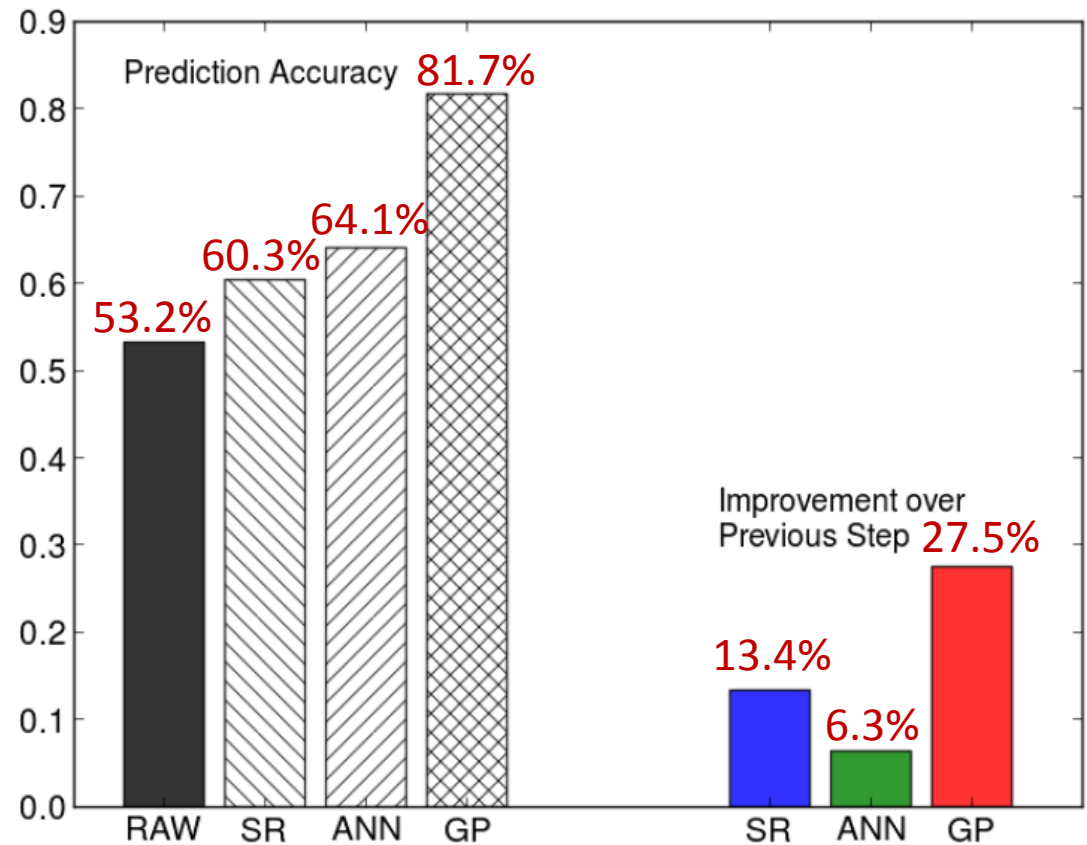
Neural  
Network

Inference

# Evaluation and Results on the deployment dataset

Overall Results:

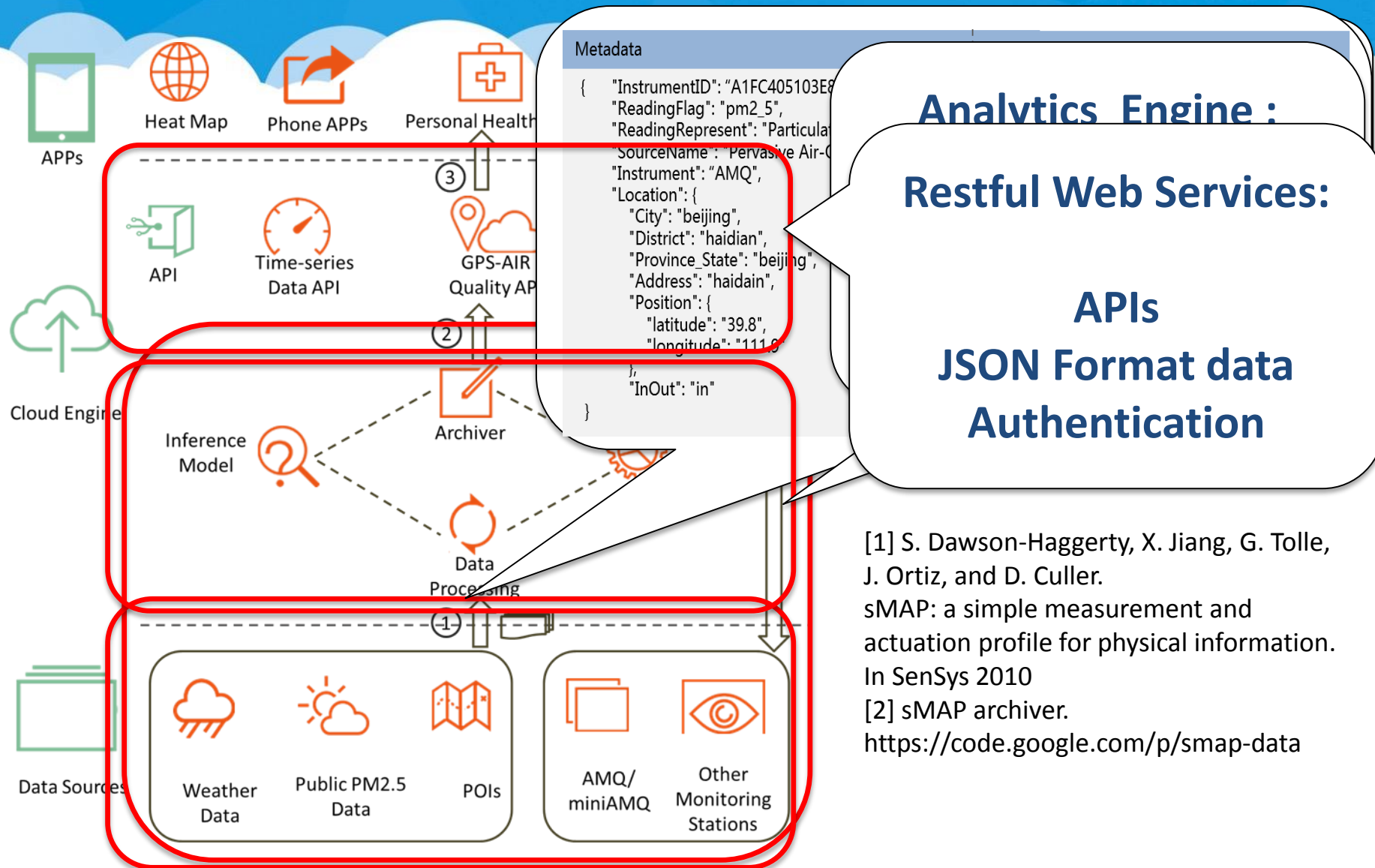
- 53.6% increase





# Cloud Architecture

## open data exchange platform

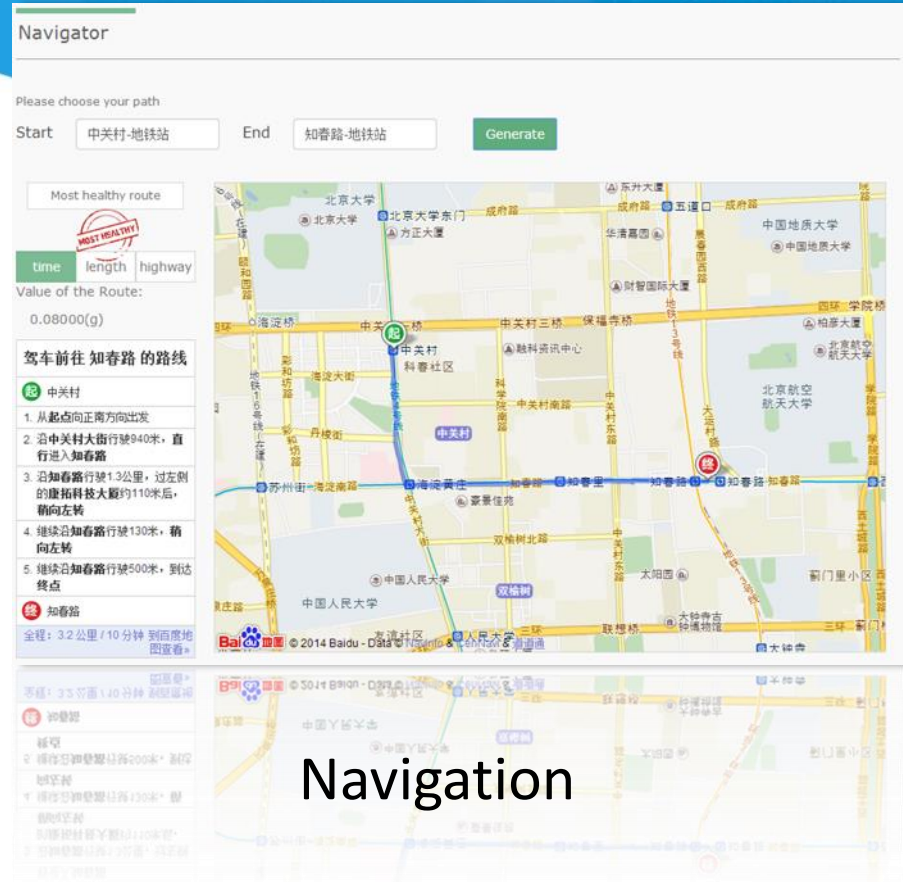


# APIs and Applications

## web applications



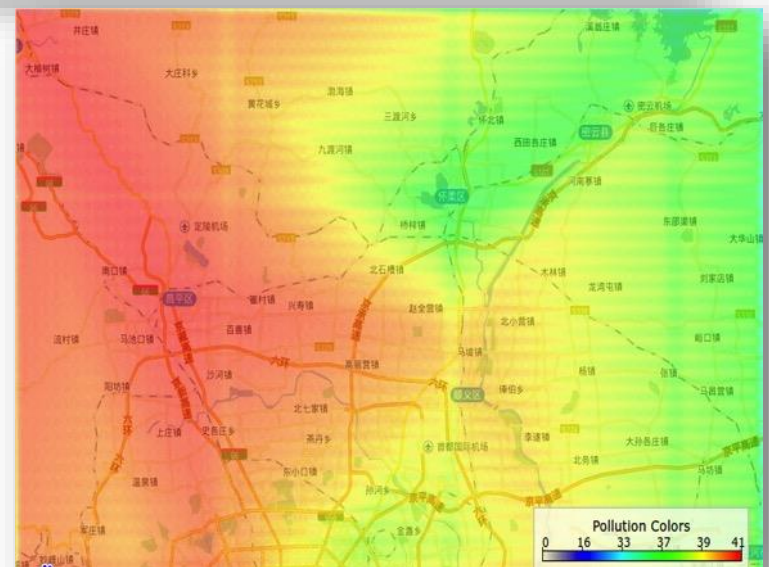
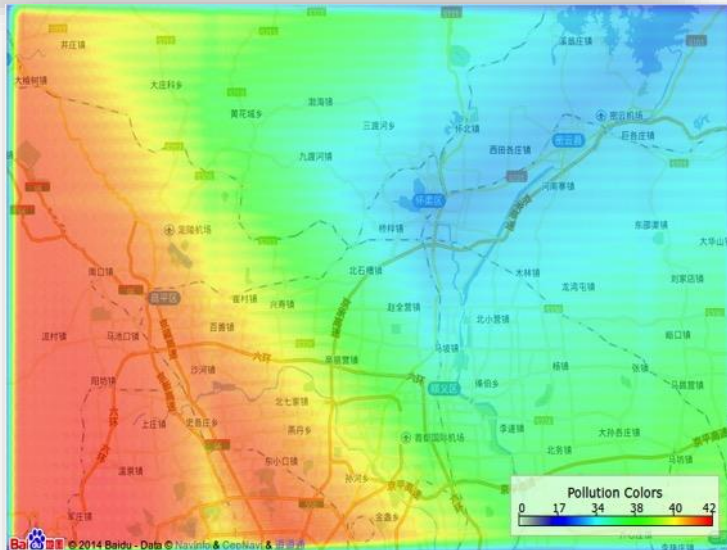
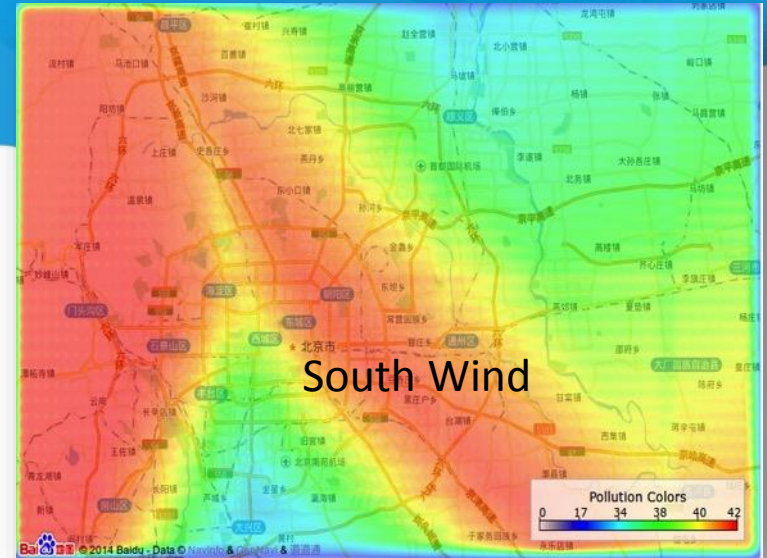
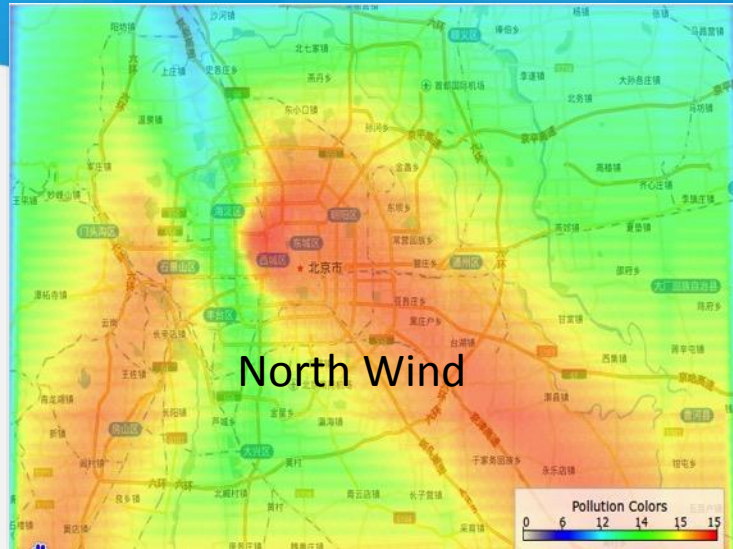
## Visualization



Find more on our website:  
<http://www.air-scientific.com>

# APIs and Applications

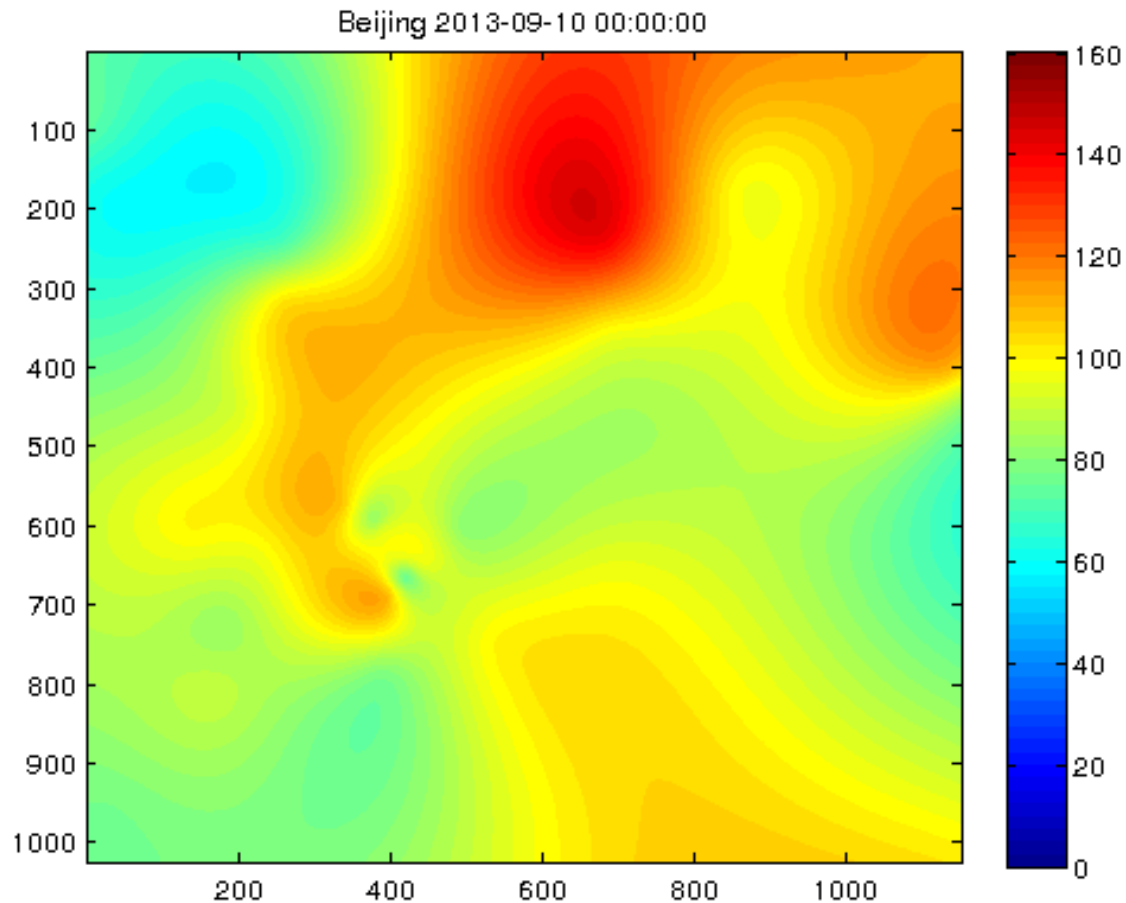
## heat map





# APIs and Applications

## heat map



Heatmap of the Spatial Inference



# APIs and Applications

## mobile applications



A:AirPet



B:AirFace



C:Aerial Drone



F:WeChat



D:AQM



E:MyAir

Find more info on our website:  
<http://www.air-scientific.com>

# Conclusion

- Stationary and mobile Internet-connected PM2.5 monitors
- Cloud-based data analytics engine to increase the accuracy of AQM by 53.6% in real deployment
- An open platform to provide easily accessible air quality data and web services
- Several useful web and mobile applications for individuals and government

Thank you for  
listening!

Questions?