Modeling the Air Pollution in China

— Based on Nonparametric methods

Zihan Zhang

December 15, 2021

Motivation

Figure: Air quality in Beijing

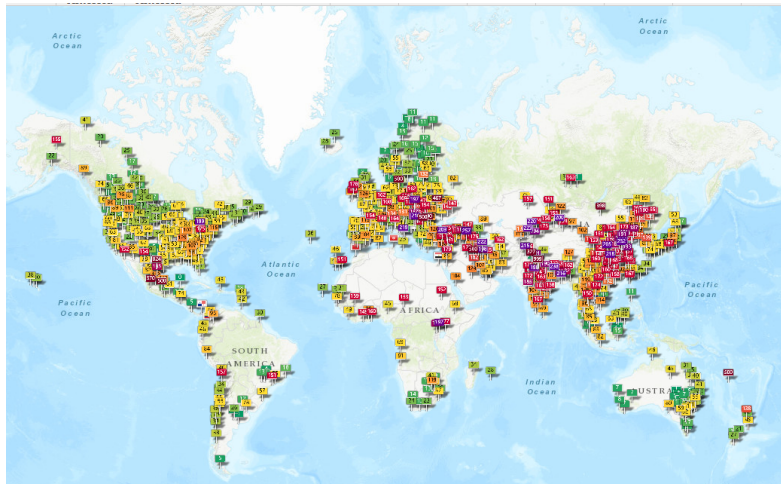


Figure: World map with AQI

DATA

- Hourly city AQI from the National Environmental Monitoring Center (CNEMC).
- City location data.

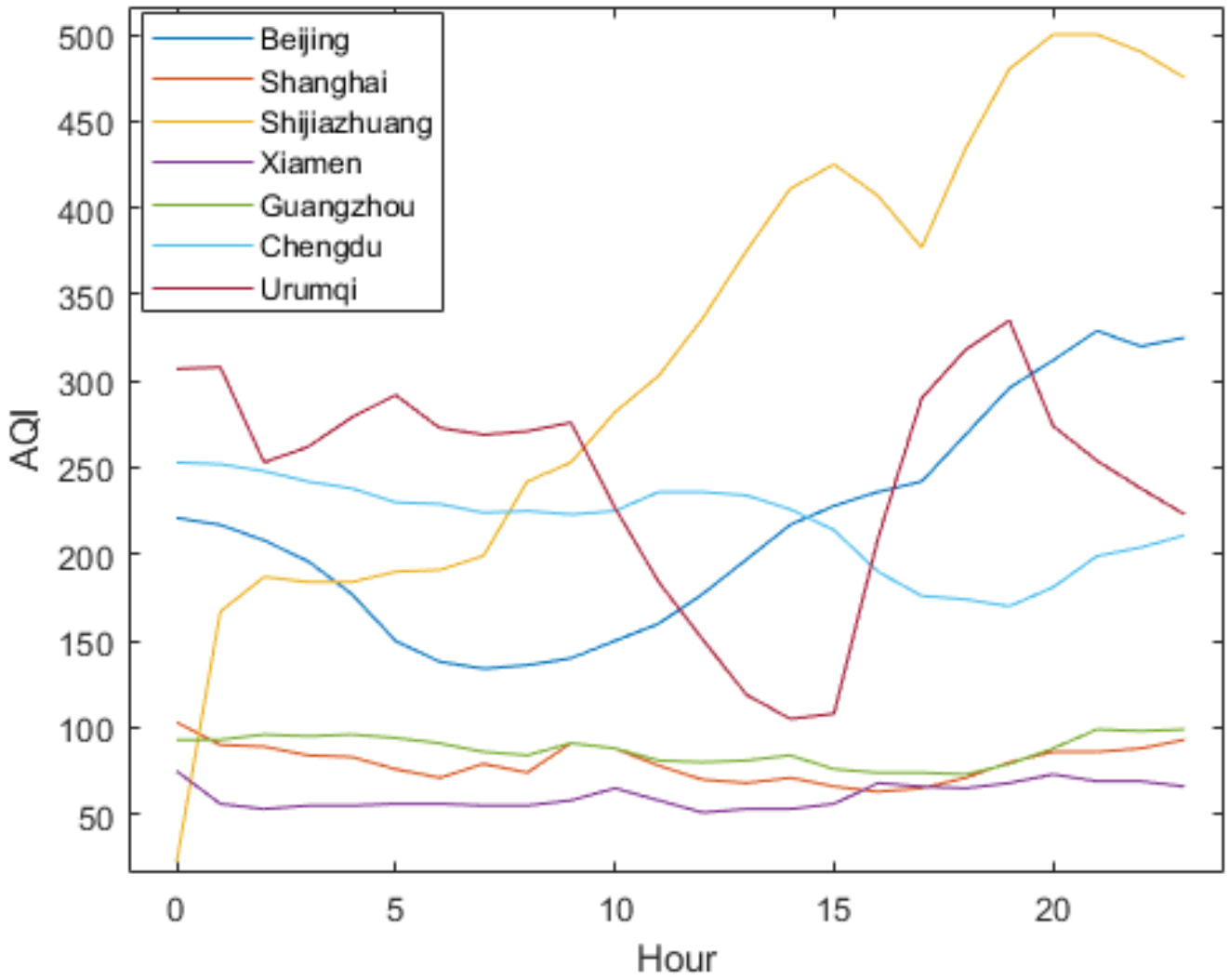


Figure: AQI of seven representative cities in China.

DATA

Table: Descriptive Statistics

	(1) N	(2) mean	(3) sd	(4) min	(5) max
AQI per hour in 7 representative cities					
Beijing	8,497	102.0	82.36	9	500
Shijiazhuang	8,560	132.7	102.2	15	500
Shanghai	8,559	65.88	41.98	9	452
Xiamen	8,558	46.31	23.04	7	221
Guangzhou	8,559	54.63	26.73	9	264
Chengdu	8,559	89.98	48.91	14	310
Urumqi	8,559	107.9	88.55	9	500
Average AQI over year for 285 cities					
AQI	285	74.22	20.42	29.15	132.7

City location

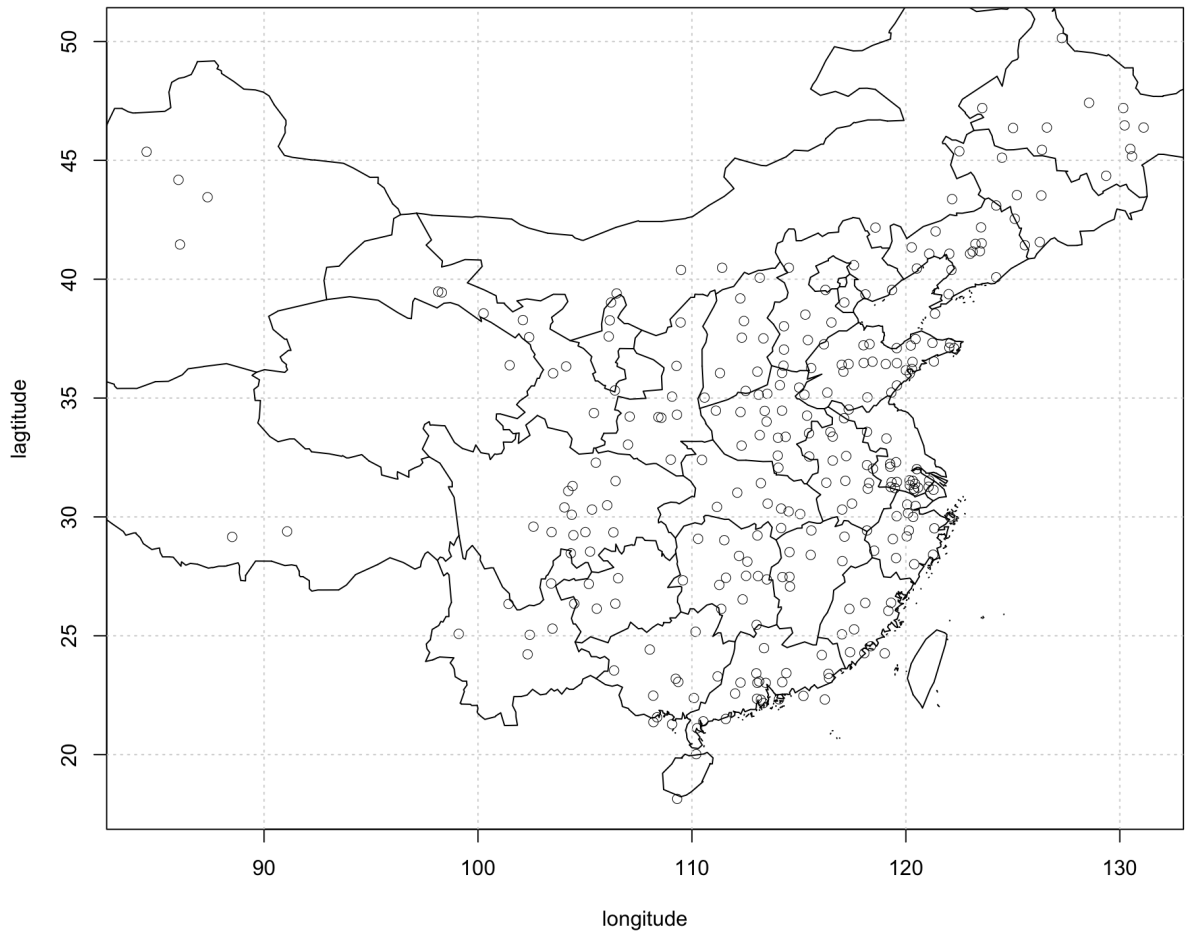


Figure: City location

METHODS

- Gaussian process and Local constant kernel regression

$$y = f(x) + \epsilon$$

$$\begin{bmatrix} y \\ f(x^*) \end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix} k(x,x) + \sigma_n^2 I & k(x,x^*) \\ k(x^*,x) & k(x^*,x^*) \end{bmatrix}\right)$$

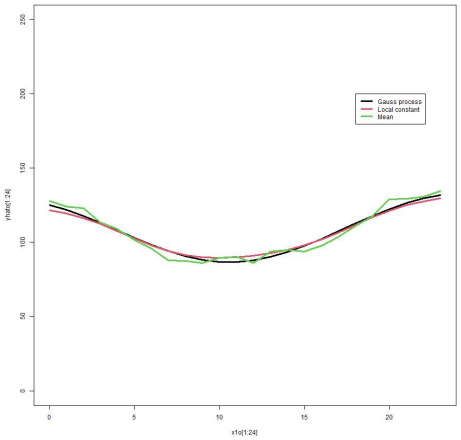
$$f(x^*) \mid x^*, x, y \sim \mathcal{N}\Big(k(x^*,x)\left[k(x,x) + \sigma_u^2 I\right]^{-1} y,$$
$$k(x^*,x^*) - k(x^*,x)\,k(x,x)^{-1}k(x,x^*)\Big)$$

, where $k(x,x') = \sigma_f^2 \exp\left(-(x-x')^2 / (2l^2)\right)$

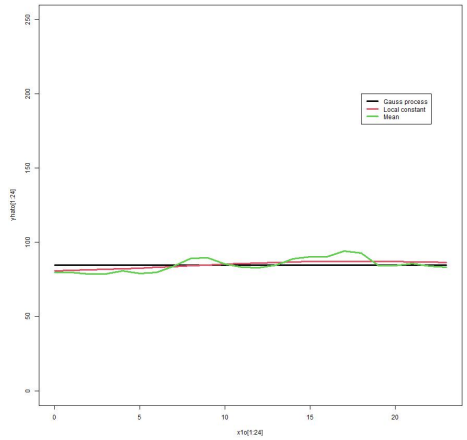
Hyper-parameters $\hat{\theta} = (\sigma_f, \sigma_n, l)$ are chosen by,

$$\hat{\theta} = \arg\max_{\theta} \log p(y \mid x, \theta) = -\frac{1}{2} y^T K_y(\theta)^{-1} y - \frac{1}{2}\log\left|K_y(\theta)\right| - \frac{n}{2}\log 2\pi$$

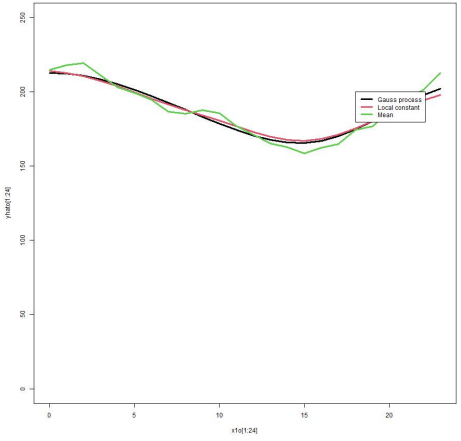
Result: Hourly trend



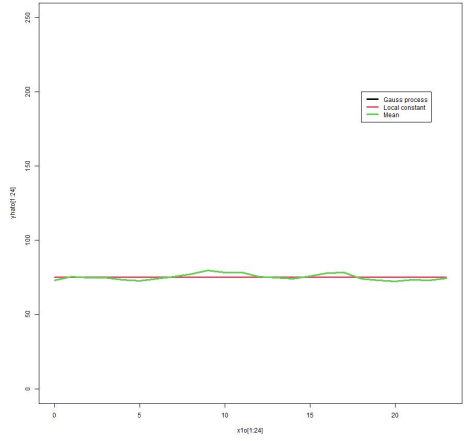
(a) Beijing Winter



(b) Beijing Summer



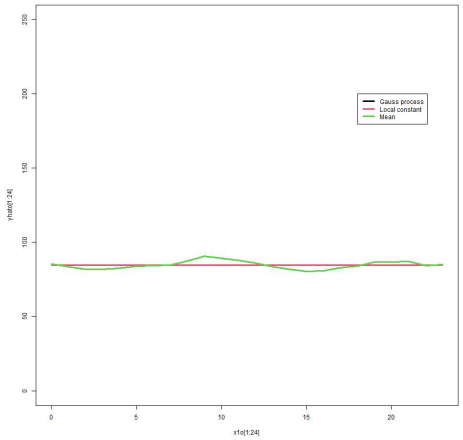
(c) Shijiazhuang Winter



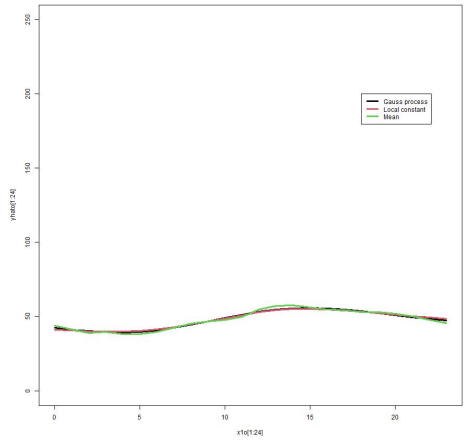
(d) Shijiazhuang Summer

Figure: AQI hourly trend: Northern cities

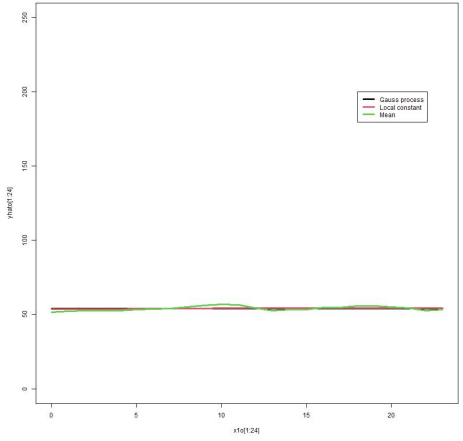
Result: Hourly trend



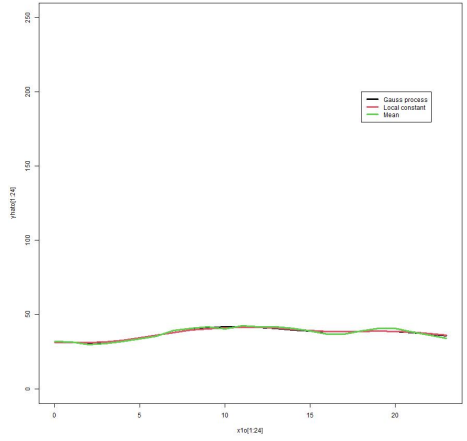
(a) Shanghai Winter



(b) Shanghai Summer



(c) Xiamen Winter



(d) Xiamen Summer

Figure: AQI hourly trend: Southern cities

RESULT: SPATIAL DISTRIBUTION

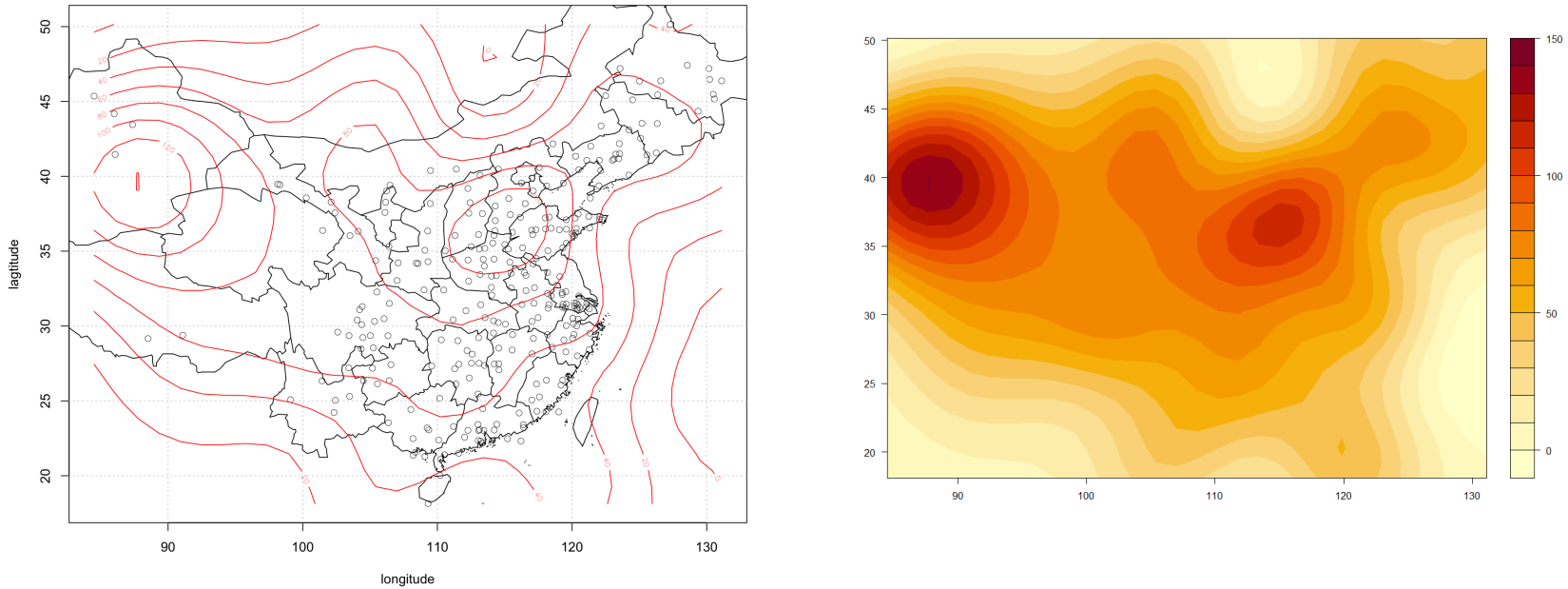


Figure: Spatial distribution of air pollution

– Thanks –