

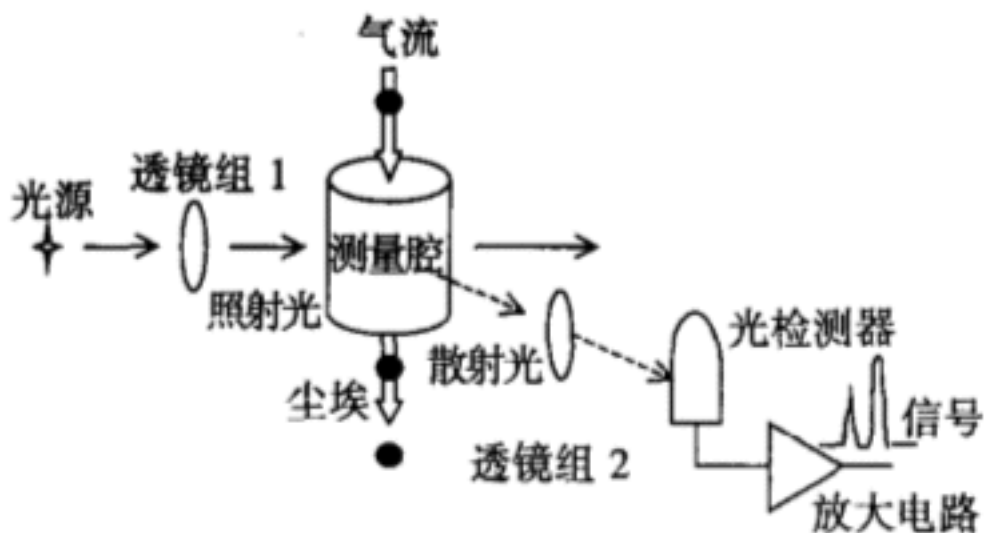
网站小组成果

- 数据组

- 北京12个国控监测点的数据, 以及35个省控点的数据, 大致包括PM2.5,PM10等数据。
- 利用爬虫即时收集雾霾相关的数据
- 对于已有数据的解释, 数据中的项目意义, 数据的误差, 数据的缺失的原因。
- 小量污染源信息, 北京企事业单位环境信息。
- 利用交通的数据和雾霾的数据, 分析出交通对于空气质量的影响。

- 设备组

- 研究光散射粒子计数器 设备原理
 - 光散射粒子计数器通过气泵、激光器、光学聚焦及采光设备、光电感应设备对气流中的粒子进行计数, 设备的示意图如 (1). 在激光器波长给定, 散射光接收角确定时, mie 散射理论指出对一定大小范围内的粒子, 其散射光强随粒子大小递增, 经由光电感应器及脉冲放大电路, 可以对每一个通过光敏区的粒子得到一个脉冲信号. 通过统计采样气流中各大小脉冲信号的个数, 我们就能获得采样气流的粒子数浓度.



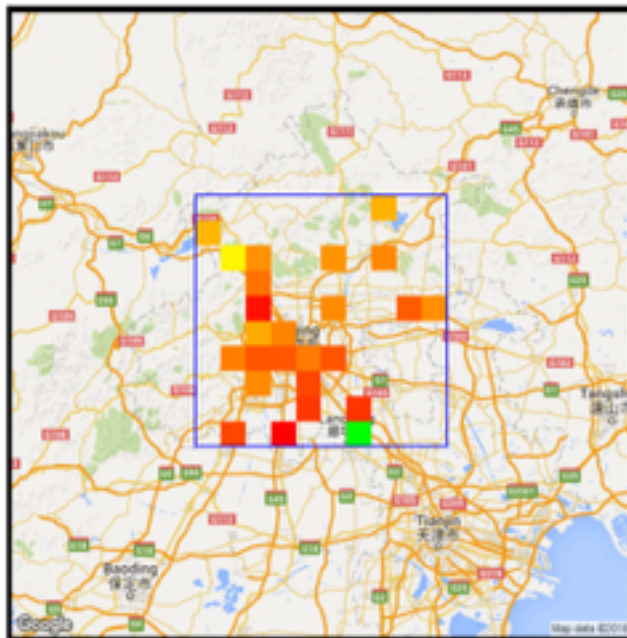
- “粒径-脉冲”标定及不同设备的一致性

- 对一台一般的基于光散射原理的粒子计数器而言, 虽然 mie 散射理论指出了粒径 - 散射 光强的关系, 但工程上仍然采用标定的方法来确定“粒径 - 脉冲”关系: 一个简单基本的方法是, 对给定粒径 (例如 0.5um) 的标准粒子气体进行测量, 获得设备测得的信号分布图, 根据分布图取一个合适的脉冲大小作为该粒径粒子的标定脉冲. 由于光敏区光强分布不均、标准粒子本身大小不全都严格相同, 所以测定的脉冲信号也不是同样大小而表现为分布图的形式.
- 需要注意的是, 即便对于同等大小的粒子, 设备检测出来的信号大小也会不同, 这种现象我们称之为“单分散粒子光散射测量法的扩散效应”. 由于仪器的加工不能达到完全精确, 越是不精密的仪器, 其扩散效应也更严重. 而不同设备的光学、电路器件的不完全精密和不一致, 将导致不同设备的扩散效应的差异, 我们把这些差异统一等效为光敏区光强分布的差异. 由于标准粒子的测量对测量环境要求较高, 此外扩散效应严重时, 标准粒子标定法也将失效, 我们的目的是使用数学方法来避免扩散效应的影响, 完成对应粒子计数器设备的标定, 以及提高各设备之间的一致性.
- 为此我们提出一种拟合方法, 该方法基本原理是: 两台设备各自采样若干次, 再以数学方法对信号分布进行拟合, 使得两台设备对同样粒子分布样本的测量结果达到一致. 通过对扩散效应的分析可以知道, 该方法本质上是对光敏区光强分布函数的拟合. 目前我们正在采集更多的数据并进行更细致的分析.
- 污染源估计组
 - **背景:** 大气污染治理的根本在污染源的治理. 然而由于PM2.5来源的复杂性以及其随空气传播和流动的特点, 污染物的来源并不容易确定. 现有的通过生产活动强度来估计污染源的方法仍有待争议. 如何结合地面站点的空气质量的观测数据、气象信息、以及遥感图像来估计污染源具有重要的实际价值和科学意义.
 - **工作介绍:** 目前我们主要运用机器学习的方法, 以Bayesian网络为模型, 结合动态的气象信息, 从地面观测站点提供的历史数据中估计PM2.5污染源的位置和强度.
 - Bayesian网络以条件概率的方式来描述不同节点之间的相互作用, 广泛的应用于图像处理、语音识别和人工智能等诸多领域. 我们把不同地理区域看作是模型中的节点, 污染源在不同区域间的对流和扩散看作是不同的节点间的相互作用, 以此为模型的基本假设来学习污染源的分布和强度. 跟传统的Bayesian网络学习不同, 在我们的问题中, 由于大气流动的不稳定性, 导致网络的连结关系随时间变化. 如何改进传统的机器学习算法, 以适应当前随时间变化的网络重建是本课题中的关键.

- 下图展示了最初步的北京城区及其周边的污染源强度分布。算法和结果的正确性还在进一步测试中。

- **应用：**

- 绘制污染源地图
- 为PM2.5动力学模型提供输入
- 直接通过Bayesian网络进行PM2.5强度预测
- 通过资料同化等手段，结合遥感和雷达数据，提高估计和预测精度。



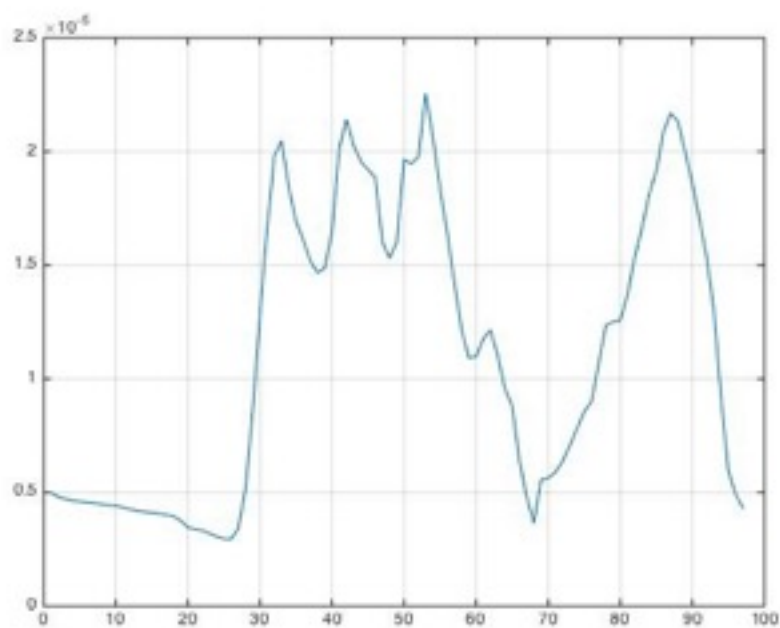
- **雾霾与天气耦合预测组**

- 研究内容：本小组主要负责大气污染与气象的耦合动力学预测研究。首先基于WRF/CHEM 建立了化学模式与气象模式的耦合预测平台，主要研究空气质量监测数据的处理方法、数据同化算法及理论、大气化学初始场的生成算法、污染源反演算法以及基于统计学习的预测结果校正方法。
- 输出结果数据格式说明：
 1. 数据集，可支持NC格式输出和MAT格式输出，目前输出文件为MAT文件。包含变量：

- (1) U10, V10 : 分别表示东西和南北方向离地面10米高度的风速, 为三维矩阵, 格式为 (LONG, LAT, Time) ;
- (2) T2: 离地面2米高度的温度, 三维矩阵, 格式为 (LONG, LAT, Time) ;
- (3) RAINC, RAINNC: 分别表示总的积云对流降水累积 和 总的格点降水累积, 为为三维矩阵, 格式为 (LONG, LAT, Time) ;
- (4) Q2:离地面2米高度的气压, 为为三维矩阵, 格式为 (LONG, LAT, Time) ;
- (5) QVAPOR: 水汽混合比, 为四维矩阵, 格式为 (LONG, LAT, Height, Time) ;
- (6) PM25: PM2.5 的即时浓度, 为四维矩阵, 格式为 (LONG, LAT, Height, Time) , 单位: "ug/kg-dryair" ;
- (7) PM10: PM10 的即时浓度, 为四维矩阵, 格式为 (LONG, LAT, Height, Time) , 单位: "ug/kg-dryair" ;
- (8) SO2: SO2 的即时浓度, 为四维矩阵, 格式为 (LONG, LAT, Height, Time) , 单位: ppmv;
- (9) NO: NO 的即时浓度, 为四维矩阵, 格式为 (LONG, LAT, Height, Time) , 单位: ppmv;
- (10)NO2: NO2 的即时浓度, 为四维矩阵, 格式为 (LONG, LAT, Height, Time) , 单位: ppmv;
- (11)CO: CO 的即时浓度, 为四维矩阵, 格式为 (LONG, LAT, Height, Time) , 单位: ppmv;
- (12)O3: O3的即时浓度, 为四维矩阵, 格式为 (LONG, LAT, Height, Time) , 单位: ppmv;

2. 图片:

- (1) 某一固定时刻, 某种物质浓度的空间分布, 如PM2.5:
- (2) 某一固定地点, 某种物质的浓度随时间变化图, 如城区一点的PM2.5变化图

SO₂浓度变化图：

3. 表格：

不同地点不同物质在某一时刻的浓度表，如：

Table 1

	AQI	PM2.5	PM10	CO	NO2	O3	SO2
万寿西宫	72	52	NaN	0.8	72	17	4
定陵	19	13	NaN	0.3	16	57	2
东四	63	45	NaN	0.6	58	23	3
天坛	40	25	NaN	0.9	80	14	6
农展馆	59	32	68	NaN	84	15	4
官园	63	29	76	1	80	12	5
海淀区万柳	54	20	57	1	86	8	4
顺义新城	48	18	48	NaN	4	2	2
怀柔镇	59	14	68	0.2	21	26	2

Table 1

	AQI	PM2.5	PM10	CO	NO2	O3	SO2
昌平镇	36	25	NaN	0.6	48	34	2
奥体中心	39	24	39	0.7	46	15	5
古城	86	47	121	1	83	14	4

- 数据库及可视化组
 - 搭建并维护数据库，提供数据库接口
 - 编写爬虫程序
 - 优化数据存储结构
 - 进行数据可视化
 - 搭建并维护项目专用 wiki，项目展示网站