利用基于卫星遥感的高精度 $PM_{2.5}$ 浓度预测分析济南市空气污染

行星数据（北京）

2019 年 6 月

I. 背景

济南市是山东省省会。由于人口集中和经济快速发展，近年来，济南市空气污染问题加剧，雾霾事件频发，带来了负面的健康效应并引起了社会的广泛关注。对于济南市 $PM_{2.5}$ 浓度的时空变化的精细化分析有助于评估已有排放控制措施的效果并为进一步制定合理的污染控制政策提供理论依据。然而，之前的相关研究依赖于地面空气质量监测站点提供的 $PM_{2.5}$ 浓度数据。由于我国在 2013 年之前并未将 $PM_{2.5}$ 浓度纳入常规监测，仅有少量以科研为目的对 $PM_{2.5}$ 零散观测，缺少时间上连续统一的 $PM_{2.5}$ 浓度数据；同时，由于地面监测站建设及维护成本较高，现有的地面监测网络倾向于在主要城市的中心区域设立站点，因而在空间上覆盖范围有限。地面监测数据的不足限制了对济南以及其他类似城市的空气污染进行长期时空特征的全面分析。

近年来，利用卫星遥感气溶胶光学厚度（AOD）反演地面 $PM_{2.5}$ 浓度的方法引起了研究者的广泛注意。AOD（Aerosol Optical Depth）是大气消光系数在垂直方向上的积分，描述了气溶胶对入射光的削减作用。卫星 AOD 数据具有高空间分辨率（可至 1 公里），覆盖范围广（全球覆盖）和历史长期连续观测（可追溯至 2000 年）等优势。之前的研究表明 AOD 与地面 $PM_{2.5}$ 浓度之间存在一定的相关性。因此，卫星遥感的 AOD 数据可以有效扩展地面监测网在时间和空间上的覆盖，进而提供历史长期无间断的、全覆盖的大气污染数据。已有研究通过大气化学传输模型、统计模型及机器学习算法等方法定量 AOD 和 $PM_{2.5}$ 浓度之间的关系，以利用遥感的 AOD 数据预测地面 $PM_{2.5}$ 浓度，定量描述在缺少 $PM_{2.5}$ 地面观测的地区或时间段的 $PM_{2.5}$ 时空分布。例如，在美国新英格兰地区和东南部地区，研究者利用包括混合效应模型和地理加权回归模型的两级空间统计学模型解析 AOD 和地面 $PM_{2.5}$ 浓度之间的非线性关系，预测没有观测数据地区的 $PM_{2.5}$ 浓度[*X. Hu et al.*, 2014; I. *Kloog et al.*, 2014]。该两级模型与地面观测值的相关系数可达 0.85 以上。类似的模型在中国同样准确预测了 $PM_{2.5}$ 浓度，与地面观测值的 R^2 达到 0.79 [Ma et al, 2016]。得益于遥感技术和统计模型及大数据的发展，我们在中国实现了 1 公里分辨率上对日平均的 $PM_{2.5}$ 浓度的准确预测[Xiao et al, 2017]以及在 6 公里分辨率上对小时平均的 $PM_{2.5}$ 浓度的准确预测[She et al, 2018]。

卫星 AOD 和 $PM_{2.5}$ 浓度间的关系存在复杂的时空变化，现有统计模型无法准确追算历史 $PM_{2.5}$ 浓度。由于缺少地面观测 $PM_{2.5}$ 浓度数据，现有模型依赖于 2013 年之后的数据来拟合模型，确定模型参数。然而之前的研究表明，当对训练数据时间区间之外的时间段的 $PM_{2.5}$ 浓度进行预测时，预测的准确性显著下降。例如，Ma et. al. (2016) 报告，当用 2013 年训练的数据预测 2014 年的 $PM_{2.5}$ 浓度时，预测值和观测值相比的 R^2 仅为 0.41。如何准确预测 2013 年之前的 $PM_{2.5}$ 浓度水平对于研究者是一个挑战。相比于传统的统计模型，机器学习算法可以更好的描述颗粒物浓度和卫星参数及气象参数间的复杂非线性关系。通过训练随机森林和 XGBoost（extreme gradient boosting）等机器学习模型，并融合不同机器学习算法得到的预测值，可以提高对 $PM_{2.5}$ 历史数据的预测准确度。我们之前的研究表明，在全国尺度上，包含机器学习算法的融合模型对 10 公里网格化的历史数据的预测比之前报道的统计模型的 R^2 提高了约 10%。融合模型预测得到的 $PM_{2.5}$ 浓度历史数据可以支持对 $PM_{2.5}$ 的长期时空趋势变化的研究[Xiao et al, 2018]。

本研究利用卫星遥感数据以及大气化学传输模型模拟数据、土地利用数据、气象数据等，建立基于机器学习算法的 $PM_{2.5}$ 预测模型，并进一步得到 2003-2016 年济南市的高分辨率（1 平方公里）网格化 $PM_{2.5}$ 浓度数据集。利用这一基于卫星数据的 $PM_{2.5}$ 浓度预测数据集，我们有机会对济南市长期局部尺度的空气污染特征进行分析，寻找影响 $PM_{2.5}$ 浓度的关键因素。

该项目建立的分析方法可以作为路线图促进对其他缺少 $PM_{2.5}$ 浓度监测的地区的空气污染特征进行分析，并为大气污染控制提供决策支持。

参考文献：

Hu X, Waller L, Lyapustin A, Wang Y, Liu Y. 2014. 10-year spatial and temporal trends of pm 2.5 concentrations in the southeastern us estimated using high-resolution satellite data. Atmospheric Chemistry and Physics 14:6301-6314.

Kloog I, Chudnovsky AA, Just AC, Nordio F, Koutrakis P, Coull BA, et al. 2014. A new hybrid spatio-temporal model for estimating daily multi-year pm 2.5 concentrations across northeastern USA using high resolution aerosol optical depth data. Atmospheric Environment 95:581-590.

Ma Z, Hu X, Sayer AM, Levy R, Zhang Q, Xue Y, et al. 2016. Satellite-based spatiotemporal trends in pm2. 5 concentrations: China, 2004-2013. Environmental Health Perspectives (Online) 124:184.

Xiao Q, Wang Y, Chang HH, Meng X, Geng G, Lyapustin A, et al. 2017. Full-coverage high-resolution daily pm2. 5 estimation using maiac aod in the yangtze river delta of china. Remote Sensing of Environment 199:437-446.

Xiao Q, Chang H H, Geng G, et al. An ensemble machine-learning model to predict historical PM2. 5 concentrations in China from satellite data. Environmental Science & Technology, 2018, 52(22): 13260-13269.

She Q, Xiao Q, Liu Y, et al. 2018. Satellite-based estimation of hourly PM2.5 levels during heavy winter pollution episodes in the Yangtze River Delta, China. AGU-JING, Xian, China. 2018.10

II. 方法

1. 数据收集与处理

为了得到足够的数据以支持模型训练，我们把研究区域扩展到覆盖整个山东省（图 1）。本研究收集和处理了研究区域内的各类数据，包括地面 $PM_{2.5}$ 观测数据、卫星气溶胶数据产品、大气化学模式模拟结果、气象数据、土地利用参数等，并对数据进行空间匹配及网格化处理，建立了约 27 万个 1 平方公里的网格，用于 $PM_{2.5}$ 预测模型的建立和提供网格化 $PM_{2.5}$ 浓度预测。

卫星 AOD 数据由搭载于 Terra（过境时间为当地时间 10:30）和 Aqua（过境时间为当地时间 13:30）卫星的中等分辨率成像分光计（Moderate Resolution Imaging Spectroradiometer，MODIS）传感器采集，并通过的最新 MAIAC（Multi-Angle Implementation of Atmospheric Correction）算法得到(Lyapustin et al. 2011a; Lyapustin et al. 2011b)。MAIAC 算法提供了 1 公里精度的网格化的 AOD 数据产品，使对局部尺度上气溶胶的异质性进行分析成为可能(Emili et al. 2011)。MAIAC 算法利用时间序列分析得到光谱的地面反射特征，进一步进行气溶胶反演。通过结合时间序列分析和空间分析有助于提高对云层和雪的识别质量。MAIAC 算法提供了质量评估变量来标示反演质量，包括云检测、地面及水面检测、雪检测等。我们移除了被标记为云或雪覆盖的 MAIAC 像素(Kloog et al. 2015)。为了使 AOD 数值更有效反应每天平均气溶胶柱浓度以及增加卫星数据的覆盖，我们每天建立了 Terra AOD

和 Aqua AOD 间的线性回归关系，并利用该关系预测仅有 Terra AOD 或仅有 Aqua AOD 时另一卫星反演的 AOD 数值。接下来，我们计算了 Aqua AOD 和 Terra AOD 的平均值以反映日平均的气溶胶厚度(Jinnagara Puttaswamy et al. 2014)并计算了月平均的 AOD 值作为之后建模的主要自变量。的为了进行网格化数据整合，我们利用泰森多边形算法建立了基于 MAIAC 像素经纬度的建模网格，并将其他数据集统一整合到该建模网格上。我们收集处理了 2013-2016 年间的数据用于构建并训练模型，并预测了 2003-2016 年的济南地区网格化 $PM_{2.5}$ 浓度以支持时空分析。

地面 $PM_{2.5}$ 观测数据收集自研究区域内的 99 个地面监测站点（图 1），数据覆盖 2013 至 2016 年。国家空气质量监测站点观测的每小时 $PM_{2.5}$ 浓度数据实时公布在中国国家环境监测中心网站上（http://www.cnemc.cn/）并通过该网站的镜像网站，PM25.in (http://pm25.in)下载。连续三小时重复的 $PM_{2.5}$ 观测值以及低于 1μg/m^3 的 $PM_{2.5}$ 观测值被认为是仪器异常产生的并被移除(Rohde and Muller 2015)。为了确保计算的月均 $PM_{2.5}$ 浓度能够正确反映每月的 $PM_{2.5}$ 水平，少于 18 个(60%)有效日均值的月份的月均 $PM_{2.5}$ 值被移除。

我们收集了 Terra 卫星搭载的 MODIS 的月平均归一化差分植被指数（Normalized Difference Vegetation Index, NDVI）1 公里数据产品 (MOD13A3)。为了填补少量缺失的 NDVI 数据，我们利用逆距离加权差值估算了缺失的 NDVI 数值，并将其空间整合到建模网格上。

我们从戈达德地球科学数据和信息服务中心（https://mirador.gsfc.nasa.gov/）下载了 OMI（Ozone Monitoring Instrument）传感器提供的二氧化氮柱浓度和气溶胶指数（aerosol index）。我们从 OMI NO2 二级数据集（OMNO2）中提取了参数 *ColumnAmountNO2Trop*，并从 OMI 气溶胶消光光学厚度和气溶胶种类二级数据集（OMAERO）中提取了参数 *AerosolIndexUV*。这两个参数的分辨率均为 13 × 24 平方公里。由于 OMI 传感器自 2007 年起存在扫描数据行异常，为了降低这一系统噪声，我们移除了所有标记为异常的像素并对 OMI 数据进行了过采样。具体来说，对于二氧化氮对流层柱浓度数据，我们在每一个网格中心点建立了 20 公里半径的采样区，并将每一季节落在采样区内的 OMI 像素取平均；对于气溶胶指数数据，我们移除了低于 0.5%百分率的异常值并在每一个网格中心点建立了 30 公里半径的采样区，并将每一季节落在采样区内的 OMI 像素取平均。

气象数据主要来自欧洲中期天气预报中心（European Centre for Medium-Range Weather Forecasts， ECMWF），分辨率为 0.25°。我们提取并处理了地面 2 米温度、相对湿度、地表气压、地面 10 米风速、地面 10 米风向、边界层高度、总降水等参数，并利用逆距离加权差值将气象场数据空间整合到建模网格上。海拔数据来自于高级星载热辐射反射探测仪（ASTER）的全球数字海拔模型（Global Digital Elevation Model，GDEM）版本 2，分辨率为 30 米。我们计算了 1 平方公里建模网格内的平均海拔。人口数据来自于网格化世界人口数据（Gridded Population of the World，GPW）全球人口数据，分辨率为 1 公里。我们利用空间整合将人口数据和建模网格相匹配。

我们提取了 MERRA-2（Modern-Era Retrospective Analysis for Research and Applications, Version 2）中的日均颗粒物浓度模拟结果(Randles et al. 2017)。MERRA-2 重分析数据提供了从 1980 年至今在 0.5° 纬度方向 × 0.625° 经度方向上全覆盖较准确的不同组分的颗粒物浓度模拟。这一重分析的大气化学模式模拟为我们提供了大尺度上 $PM_{2.5}$ 分布的信息。我们利用如下公式计算总 $PM_{2.5}$ 浓度(Buchard et al. 2016; Provençal et al. 2017):

$$\mathbf{PM_{2.5} = 1.375 \times 硫酸根离子 + 2.1 \times 有机碳 + 黑炭 + 灰尘_{2.5} + 海盐_{2.5}}$$

其中，灰尘 $_{2.5}$ 和海盐 $_{2.5}$ 是 MERRA-2 按颗粒物半径模拟结果中半径小于 2.5 微米的浓度求和。MERRA-2 按照五个半径区间模拟灰尘和海盐的浓度，我们计算了灰尘模拟结果中区间 1（半径 0.1~1.0 μm）、区间 2（半径 1~1.5 μm）和区间 3（半径 1.5~3.0 μm）的浓度和以及海盐模拟结果中区间 1（半径 0.03~0.1 μm）、区间 2（半径 0.1~0.5 μm）和区间 3（半径 0.5~1.5 μm）的浓度和。我们假设气溶胶中的硫酸盐主要以硫酸铵的形式存在，因此硫酸根质量浓度乘以 1.375 得到硫酸铵的质量浓度。有机碳和有机颗粒物质量比率，2.1，是通过比较中国 2013-2016 年的 $PM_{2.5}$ 观测值和 MERRA-2 有机碳模拟值估算得到的(Xiao et al., 2018)。之后，我们计算了月平均的 $PM_{2.5}$ 浓度模拟数据并利用逆距离加权差值将 MERRA-2 数据空间整合到 1 平方公里的建模网格上。

此外，我们还获取了 0.1 度网格化月平均排放清单数据（http://inventory.pku.edu.cn/home.html），并利用最近邻匹配将排放清单数据整合到模型网格。

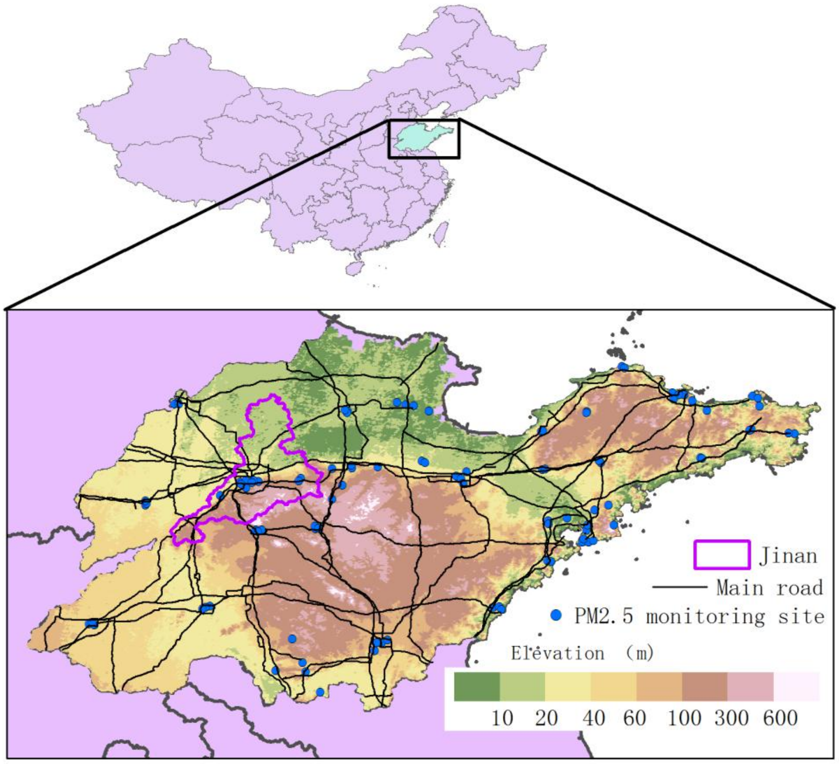


图 1. 研究区域（山东省）地形图。空气质量监测站点显示为蓝点，济南市以紫色高亮显示。

2. 模型训练和评估

本研究的技术路线如图 2 所示。首先，我们采取机器学习算法建立 $PM_{2.5}$ 浓度模型并利用十乘交叉检验和逐年交叉检验筛选出合适的模型参数及评价模型表现；之后，我们把表现最好的模型的预测值进行整合以提高历史追算的准确性和稳定性，并得到最终的 $PM_{2.5}$ 浓度预测。

在模型拟合阶段，为了筛选出合适的机器学习算法并和统计模型对比，我们建立了随机森林、XGBoost 和线性混合模型(LME) 三种模型预测月平均 $PM_{2.5}$ 浓度。线性混合模型被广泛应用

于基于卫星数据的区域及国家尺度的 $PM_{2.5}$ 浓度预测。随机森林和 XGBoost 都是基于决策树的算法，其优势包括：可以实现处理大量的特征而减少过度拟合；允许输入连续或分类变量；对于异常值的存在不敏感；可以提供模型中各参数的重要性排序，帮助筛选自变量。此外，随机森林算法还可以提供估计的泛化误差（out-of-bag error）。但是在数据量大或模型结构负责的情况下，随机森林算法所需的训练及预测时间都会显著增加。XGBoost 算法是 Gradient Boosting 算法的高效实现版本(Chen and Guestrin 2016)，其所需的训练和预测时间较短。在之前对山火事件造成的空气污染的研究中，Gradient Boosting 算法表现优于其他多种统计模型及机器学习模型(Reid et al. 2015)。这两种算法已被广泛应用于分类和回归建模。

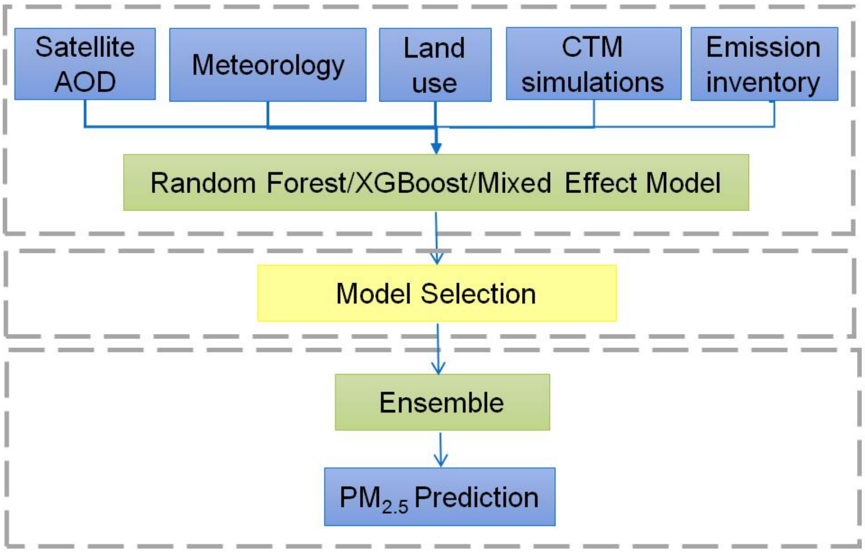


图 2. 技术路线图。

我们用 2013 至 2016 年的地面观测的 $PM_{2.5}$ 浓度数据训练模型。十乘交叉检验被用于优化模型超参数。在十乘交叉检验中，每次随机选择 10%的数据作为检验数据，利用剩下 90%的数据训练模型并预测检验数据。重复该过程十次使每一个数据都被选作检测数据一次。通过比较预测的 $PM_{2.5}$ 浓度和地面观测的 $PM_{2.5}$ 浓度评价模型表现。由于该模型致力于追算历史数据，我们倾向于选择结构较简单、预测表现稳定的超参数组合。十成交叉检验也被用于选择合适的预测参数。

为了评价模型表现，尤其是模型在追算历史数据时的表现，我们利用空间交叉检验和逐年交叉检验来选择合适的模型。之前的研究表明，$PM_{2.5}$ 预测模型在外推到训练数据时间段之外，即利用模型进行追算历史浓度时，预测 $PM_{2.5}$ 浓度时误差显著高于十乘交叉检验的结果。部分因为在十乘交叉检验中，随机选取的训练模型通常包含足够的观测值并预测出局域有效而无法外推的预测参数间的关系，导致时间和空间上的过拟合。因此，我们设计了逐年交叉检验的检验方法来评价模型在追算历史数据时的表现。在空间十乘交叉检验中，每次随机选择 10%的监测站的数据作为检验数据，利用剩下 90%的监测站的数据训练模型并预测检验数据。重复该过程十次使每一个监测站都被选作检测数据一次。逐年交叉检验依次选择一年的数据作为检测数据，并用其余年份的数据训练模型并预测检验数据。各个模型空间十乘交叉检验和逐年交叉检验结果可以更好地刻画模型在追算数据时的表现。

之后，我们融合了在交叉检验中表现较优的模型的预测值得到融合的预测值以平衡模型准确性和稳定性。我们利用十成交叉检验选择表现最好的融合模型并发现随机森林和 LME 模型

的加权平均值表现最好。由于交叉检验中的斜率反映了 $PM_{2.5}$ 预测值的系统偏差，权重由交叉检验中斜率的倒数决定。

II. 结果

1. 模型表现

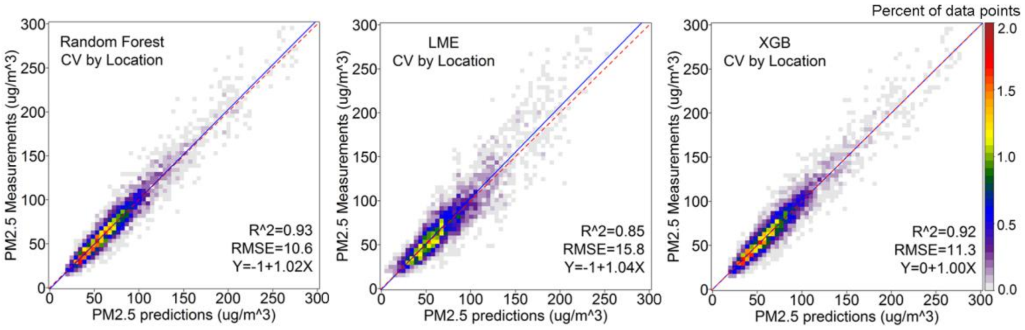


图 3 密度散点图显示随机森林（左）、线性混合模型（中）和 XGBoost（右）的空间十乘交叉检验结果。颜色代表落入格子的数据点占总数据点的百分比。

在空间十乘交叉检验中，三个模型的预测值和地面监测值都比较一致。拟合的线性关系非常靠近 45 度的标准线且斜率接近于 1，截距落在-1 和 0 的区间内（图 3）。其中，随机森林算法表现最好。与地面观测值相比，随机森林模型的预测值提供了最高的 R^2 （0.93）和最低的均方根误差（10.6 μg/m^3）。在三个模型中，线性混合模型的 R^2 最低（0.85）且均方根误差最高（15.8 μg/m^3），提示机器学习算法可以更好的描述 $PM_{2.5}$ 和其他预测变量间复杂非线性的关系。在训练模型时而我们注意到，虽然基于决策树的机器学习算法，随机森林和 XGBoost，提供了更准确的 $PM_{2.5}$ 浓度预测，但是这两种算法在处理空间变量时存在缺陷。在模型中加入不随时间变化的空间变量，例如经纬度，会导致预测地图出现空间上的不平滑。线性混合效应模型可以更好地处理空间变量。

图 4 展示了各个模型在逐年交叉检验中的表现。和之前的研究一致，我们发现预测模型在逐年交叉检验中的表现比其他交叉检验设计要差，提示未观测到的时间趋势对于 $PM_{2.5}$ 浓度预测有影响。经济发展、产业结构调整以及污染控制政策导致的年际间的 $PM_{2.5}$ 排放源变化可能影响 $PM_{2.5}$ 浓度和预测变量间的关系，但是现有的排放清单无法完全反映这些变化进一步损害了模型的追算预测表现。在逐年交叉检验中，线性混合效应模型表现最佳（R^2=0.76，均方根误差 21.2 μg/m^3），XGBoost 模型表现最差（R^2=0.71，均方根误差 23.5μg/m^3）。这一结果显示，在样本数量较小时，机器学习算法可能存在过拟合的风险。增加训练样本数量和调整模型超参数以降低模型复杂度可以一定程度上改善过拟合，提高模型稳定性。同时，我们注意到，所有模型的 $PM_{2.5}$ 浓度预测在高 $PM_{2.5}$ 观测值区（>150 微克/立方米）都有一定程度的低估，可能的原因包括对重霾污染过程的发生机制认识不够，缺少与重霾发生相关的关键变量。

为了得到在时间和空间上都稳定的预测值，我们融合了随机森林算法和线性混合效应模型预测的 $PM_{2.5}$ 浓度。随机森林算法和线性混合效应模型刻画了 $PM_{2.5}$ 浓度和预测变量间关系的不同方面，分别在空间交叉检验和逐年交叉检验中表现最佳。我们尝试了不同的融合方法，包括地理加权回归、扩展的加性模型、以及加权平均。基于交叉检验的结果，我们选择了加权平均得到融合的预测值。图 5 展示了融合的预测值在空间十乘交叉检验和逐年交叉检验中的表现。融合的预测值平衡了随机森林模型在空间上的稳定表现和线性混合效应模型在

时间上的稳定表现。在空间十乘交叉检验中，与地面观测 PM$_{2.5}$ 浓度相比，融合的预测值 R^2 达到 0.90，均方根误差为 12.6 μg/m^3，且其斜率（1.04）和截距（-2.7）非常接近 45° 标准线。说明和地面观测值相比，融合的预测值不存在系统性的偏差。我们也注意到，和单个模型的 PM$_{2.5}$ 浓度预测表现相似，融合的预测值在高 PM$_{2.5}$ 观测值区也略有低估。我们进一步对 PM$_{2.5}$ 浓度预测值和地面监测值的时间趋势进行分析比较（图 6）。 PM$_{2.5}$ 浓度的地面观测值和相应网格预测值的时间趋势非常一致（相关系数 0.96）。

为了进一步探究基于卫星数据的 PM$_{2.5}$ 浓度预测在 2013 年之前的表现，我们收集了 Donkelaar et al. (2016)预测的融合大气化学模式和卫星数据的 1 km 精度年平均 PM$_{2.5}$ 浓度预测并和本研究中得到的 PM$_{2.5}$ 浓度预测进行对比（图 7）。我们发现，和地面监测值相比，Donkelaar et al.的预测误差（R^2=0.68，均方根误差 17.7 μg/m^3）大于本研究中拟合模型的预测误差（R^2=0.95，均方根误差 6.9 μg/m^3）。比较 2003-2012 年间两种预测模型的 PM$_{2.5}$ 浓度预测，得到 R^2 为 0.71，均方根误为 11.4 μg/m^3。Donkelaar et al.的预测在高 PM$_{2.5}$ 值时对地面监测和本研究中的预测均有低估。

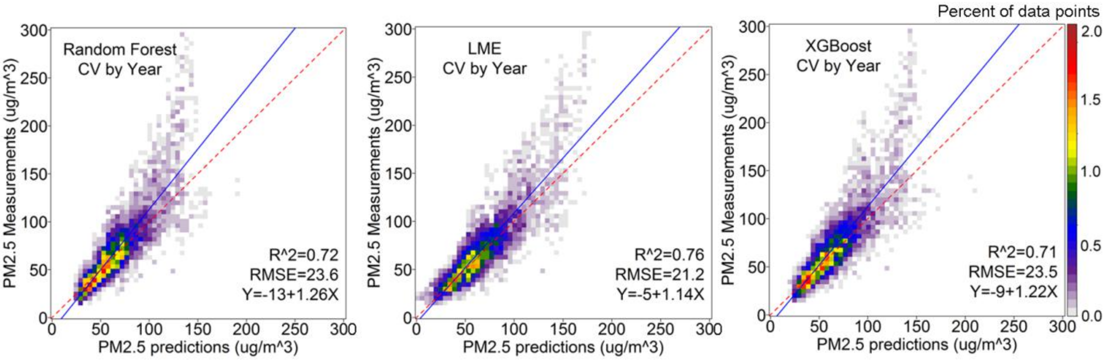


图 4 密度散点图显示随机森林（左）、线性混合模型（中）和 XGBoost（右）的逐年交叉检验结果。颜色代表落入格子的数据点占总数据点的百分比。

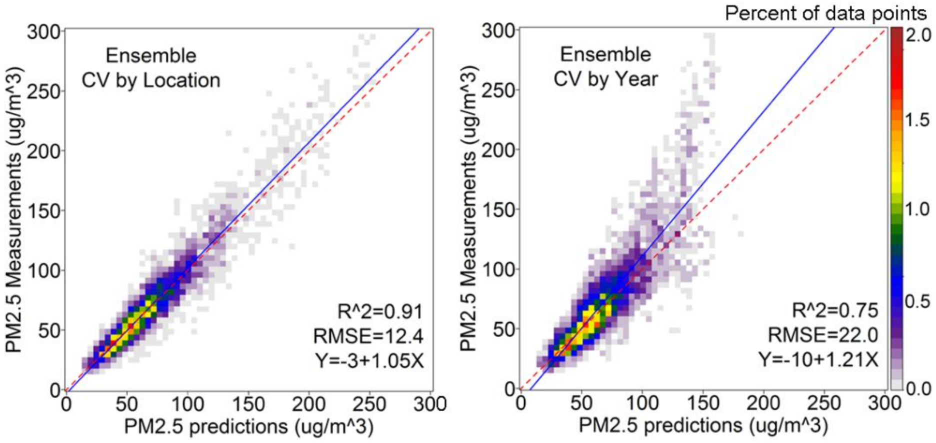


图 5 密度散点图显示融合的预测值在空间检查检验（左）和逐年交叉检验（右）中的结果。颜色代表落入格子的数据点占总数据点的百分比。

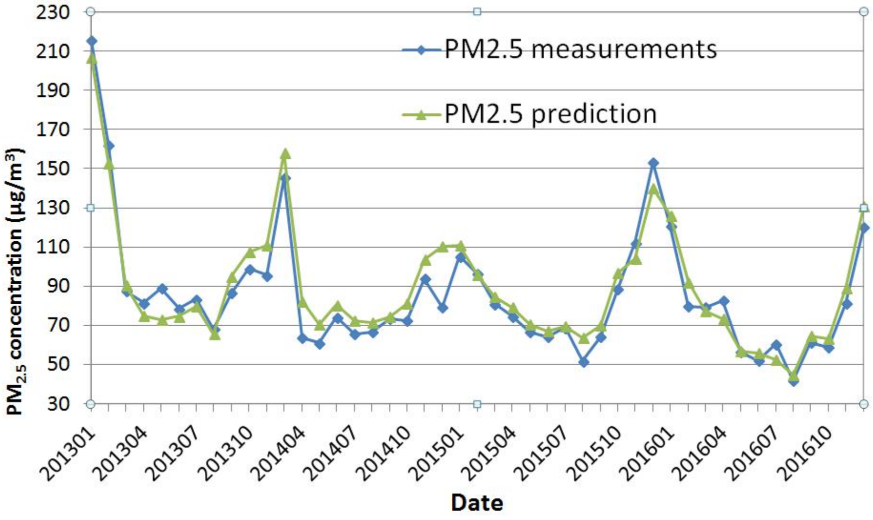


图 6 济南市 $PM_{2.5}$ 月平均的地面观测值和预测值的时间趋势对比。

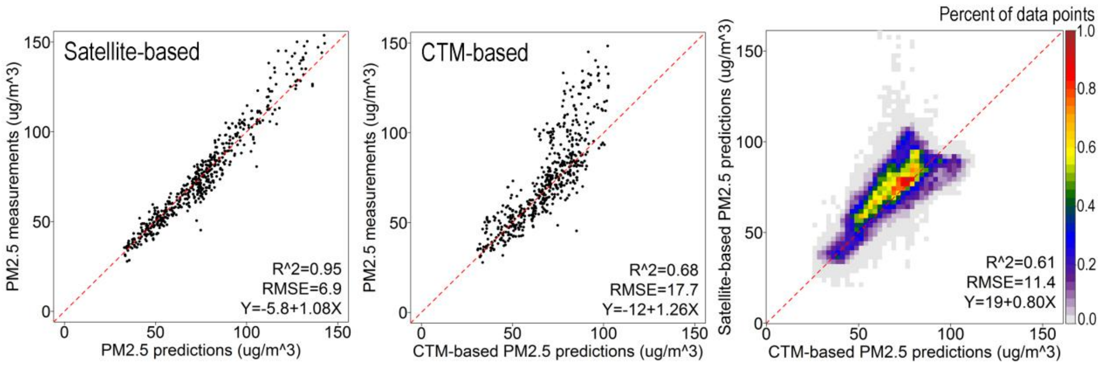


图 7 基于卫星和基于大气化学模式的 2013-2016 年年平均 $PM_{2.5}$ 浓度预测值与地面观测值对比。

2. $PM_{2.5}$ 空间分布

2016 年，济南市年平均 $PM_{2.5}$ 浓度为 74μg/m³，变化区间为 62 -85 μg/m³。如图 8 所示，2016 年 $PM_{2.5}$ 污染热点地区位于城市和乡镇中心。虽然多数 $PM_{2.5}$ 污染热点被地面空气质量检测网覆盖，但我们仍然识别出了一些没有被空气质量检测网覆盖的污染热点。例如，济南市东北区域的热点以及济南市西部聊城市的热点区域附近都没有地面监测站。济南市的北部和南部 $PM_{2.5}$ 污染水平相对较低。我们观察到，在济南市周边路网密集的区域，$PM_{2.5}$ 浓度沿主要道路升高，提示交通排放对于局部地区 $PM_{2.5}$ 污染有显著贡献。

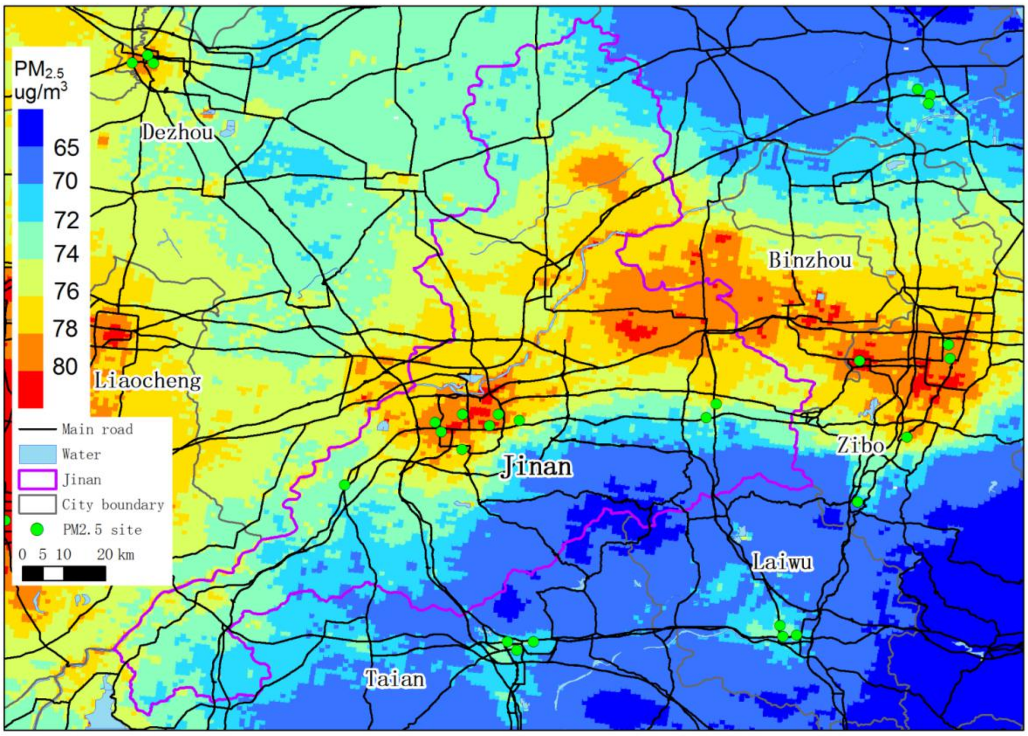


图 8. 济南市 2016 年年平均 $PM_{2.5}$ 分布。

3. $PM_{2.5}$ 时间趋势

图 9 显示了济南市 2003-2016 年间月平均 $PM_{2.5}$ 浓度的变化趋势。我们观察到 $PM_{2.5}$ 浓度呈现出明显的季节周期性：夏季（六月-八月）$PM_{2.5}$ 浓度最低，冬季（十二月、一月、二月）$PM_{2.5}$ 浓度最高。$PM_{2.5}$ 浓度的最高值从 2007 年到 2013 年上升显著，平均浓度从 116 μg/m^3 上升到 198 μg/m^3；而出现在夏季的 $PM_{2.5}$ 浓度最低值上升缓慢，平均浓度从 47 μg/m^3 上升到 71 μg/m^3。2013 年后，冬季和夏季的 $PM_{2.5}$ 浓度均开始显著下降。济南市年平均 $PM_{2.5}$ 浓度在 2005-2013 年间持续上升（图 9），并在 2013 年后迅速下降，平均降低速度达到每年 6 μg/m^3。2016 年汾渭平原的平均 $PM_{2.5}$ 浓度为 72.6 μg/m^3，低于 2003 年的平均 $PM_{2.5}$ 浓度 76.3 μg/m^3 和 2004 年的平均 $PM_{2.5}$ 浓度 75.4 μg/m^3，是 2003-2016 年间第二低的年平均 $PM_{2.5}$ 浓度。$PM_{2.5}$ 浓度的下降主要得益于环保部新的环境空气质量标准的发布和相关污染控制措施的实施。这一环境空气质量标准将 $PM_{2.5}$ 列为常规监测的环境空气污染物并制定了相应地浓度限值，促进了对 $PM_{2.5}$ 的监测和排放控制措施的制定实施。与季节平均 $PM_{2.5}$ 浓度对比，我们观察到不同的时间趋势：冬季和秋季（9 月-11 月）平均 $PM_{2.5}$ 浓度的时间趋势和年平均值非常一致，而春季（三月-五月）平均 $PM_{2.5}$ 浓度的时间趋势在 2008-2014 年间和年平均值变化相反。夏季平均平均 $PM_{2.5}$ 浓度一直保持稳定直到 2009 年出现一次上升，之后继续保持稳定直到 2014 年开始下降。值得注意的是，在 2015-2016 年，冬季的平均 $PM_{2.5}$ 浓度开始升高而其他季节的 $PM_{2.5}$ 浓度仍然保持下降趋势。

我们分析了各个季节的 $PM_{2.5}$ 水平变化对长期 $PM_{2.5}$ 浓度变化的贡献（图 10）。与预期一致，冬季 $PM_{2.5}$ 浓度变化主导了 $PM_{2.5}$ 浓度的年际变化。例如在 2005-2013 年间，冬季 $PM_{2.5}$ 浓度上升贡献了 58%的年平均 $PM_{2.5}$ 浓度上升；在 2013-2016 年间，冬季 $PM_{2.5}$ 浓度下降贡献了 41%的年平均 $PM_{2.5}$ 浓度下降。值得注意的是冬季是唯一一个季节其对 $PM_{2.5}$ 浓度升高的贡献大于

对 $PM_{2.5}$ 浓度下降的贡献。可能的原因包括人口流入导致的冬季供暖需求增加引起冬季污染排放增加。春季 $PM_{2.5}$ 浓度变化对年际间 $PM_{2.5}$ 浓度上升和下降的贡献最小，显示春季 $PM_{2.5}$ 浓度在研究时间段中相对稳定。综上所述，冬季 $PM_{2.5}$ 浓度的波动对年平均 $PM_{2.5}$ 浓度波动起主导作用。为了达到提高空气质量的目的仍然需要控制冬季 $PM_{2.5}$ 的排放。

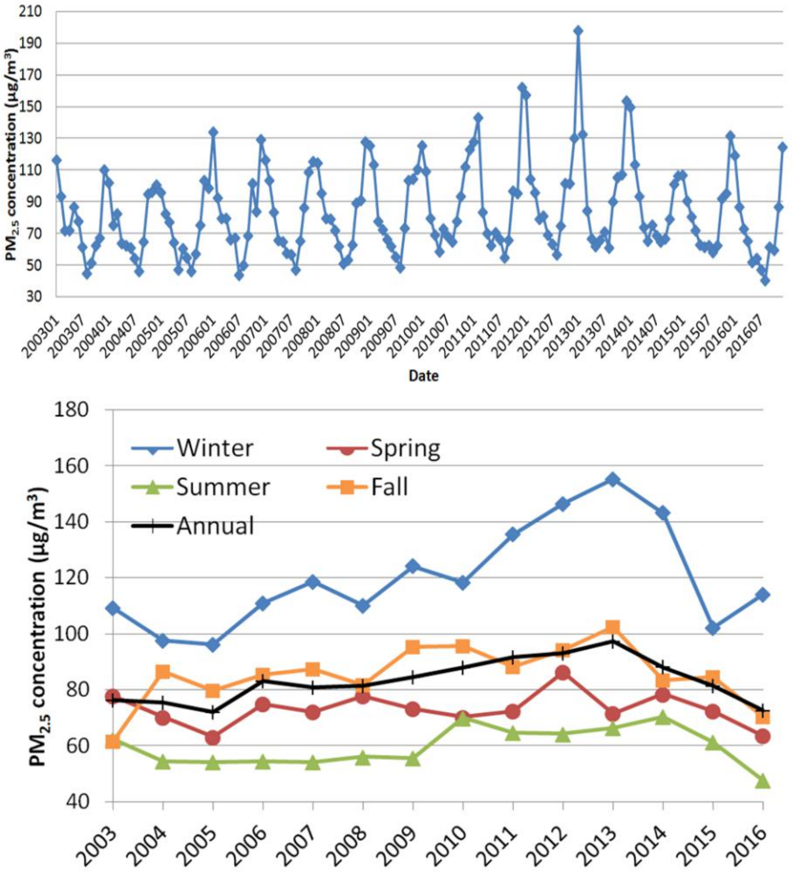


图 9. 济南市月平均（上）和季节平均（下） $PM_{2.5}$ 浓度。

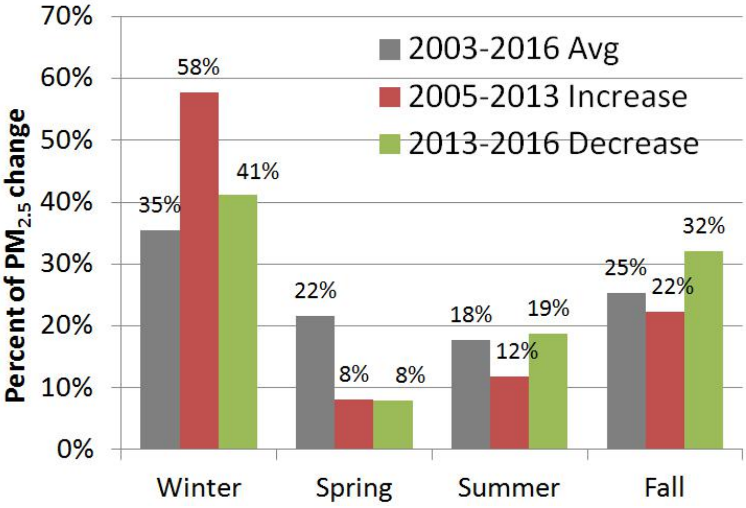


图 10. 济南市 2003-2016 年间季节 $PM_{2.5}$ 变化对年平均 $PM_{2.5}$ 变化的贡献。

4. PM$_{2.5}$ 季节性空间分布

为了分析济南市 PM$_{2.5}$ 热点在不同季节的分布，我们以百分比率为单位（图 11）和以浓度微克/立方米为单位（图 12）绘制了各个季节 PM$_{2.5}$ 浓度分布图。由于不同季节的平均 PM$_{2.5}$ 浓度差异极大，使用百分比率为单位并且将超出济南市 PM$_{2.5}$ 浓度 75%百分比率的区域定义为污染热点区域可以标准化 PM$_{2.5}$ 浓度平均水平的不同，有利于跨季节比较 PM$_{2.5}$ 污染热点的分布。我们注意到在春季，济南北部的 PM$_{2.5}$ 水平较低，但在秋冬季节，该地区为 PM$_{2.5}$ 污染热点。秋季济南市中心区域的污染热点有所收缩而相邻地区，尤其是济南市以西，存在高 PM$_{2.5}$ 污染的区域。临近地区的高污染可能传输到济南并影响济南市空气质量。与此相反，冬季时济南市，尤其是济南市中心和北部地区，成为这一区域的污染热点。污染物浓度高于相邻地区。

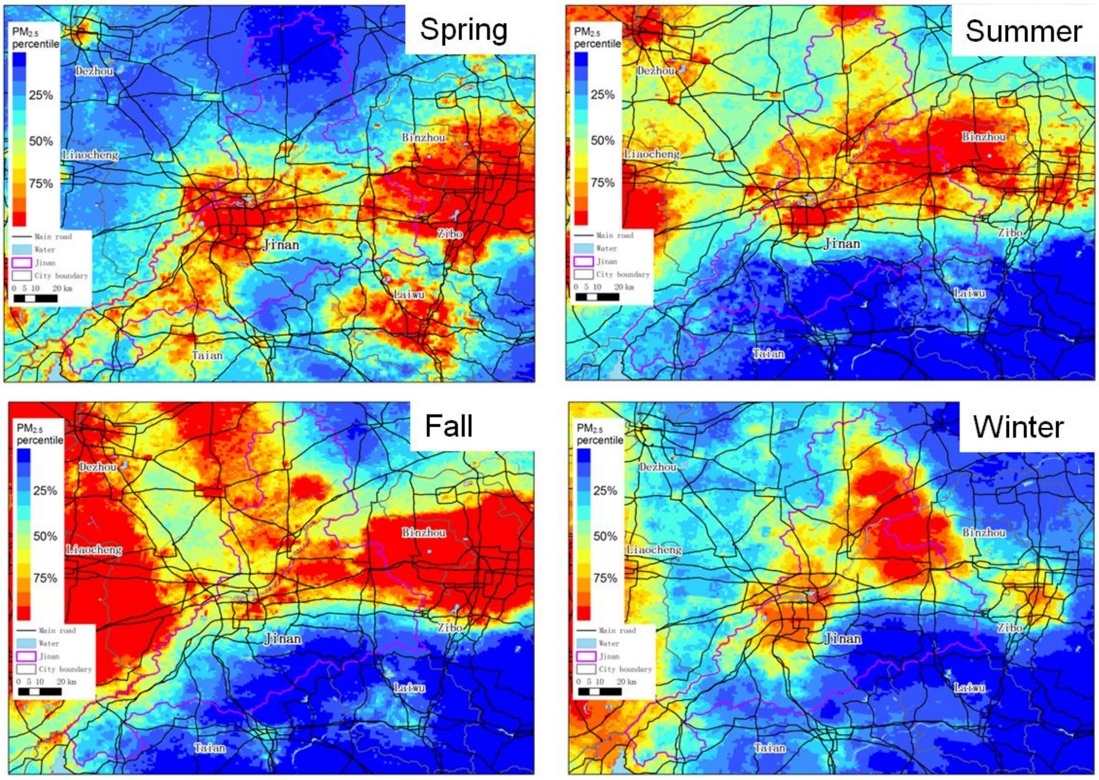


图 11. 以百分率为单位的济南市季节性 PM$_{2.5}$ 空间分布。

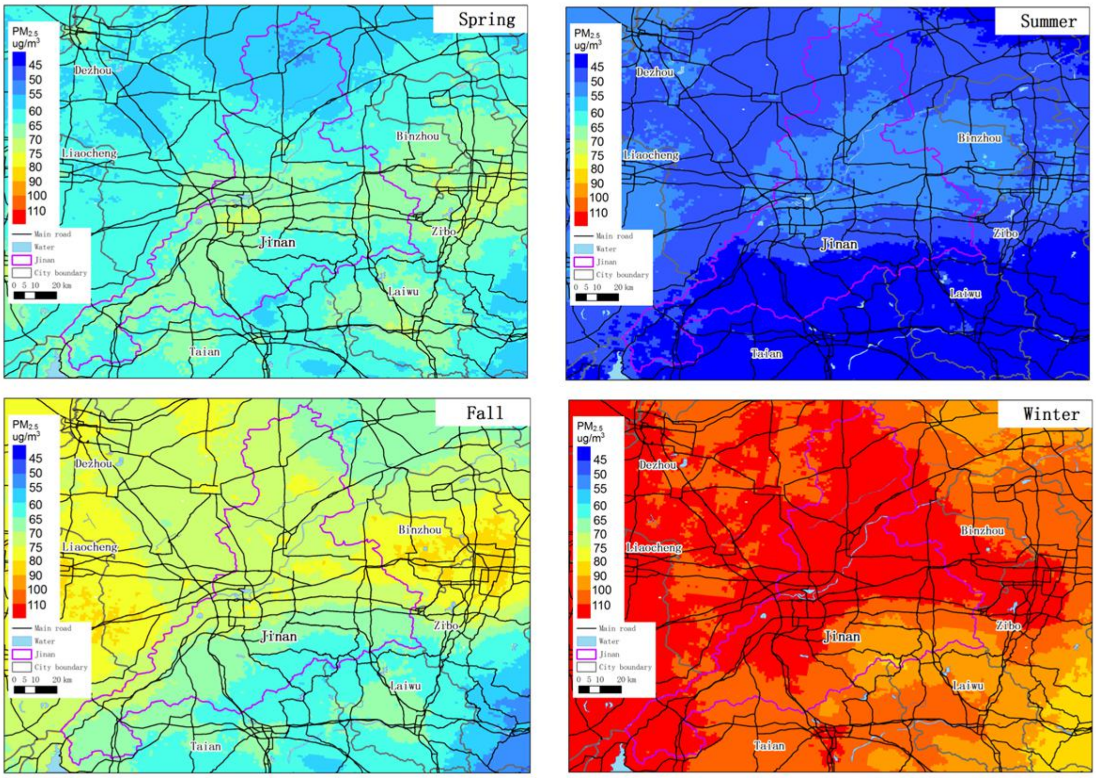
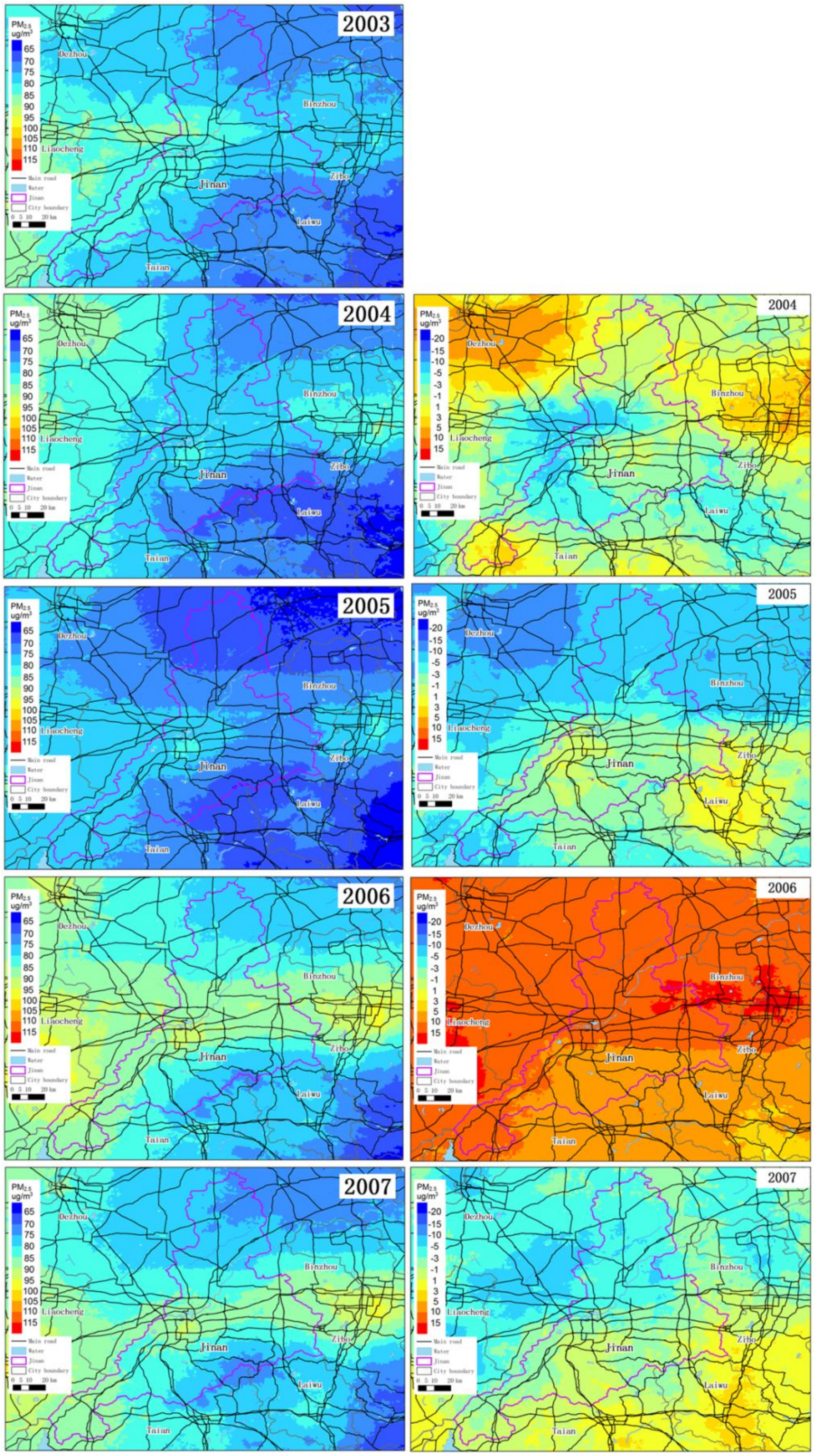
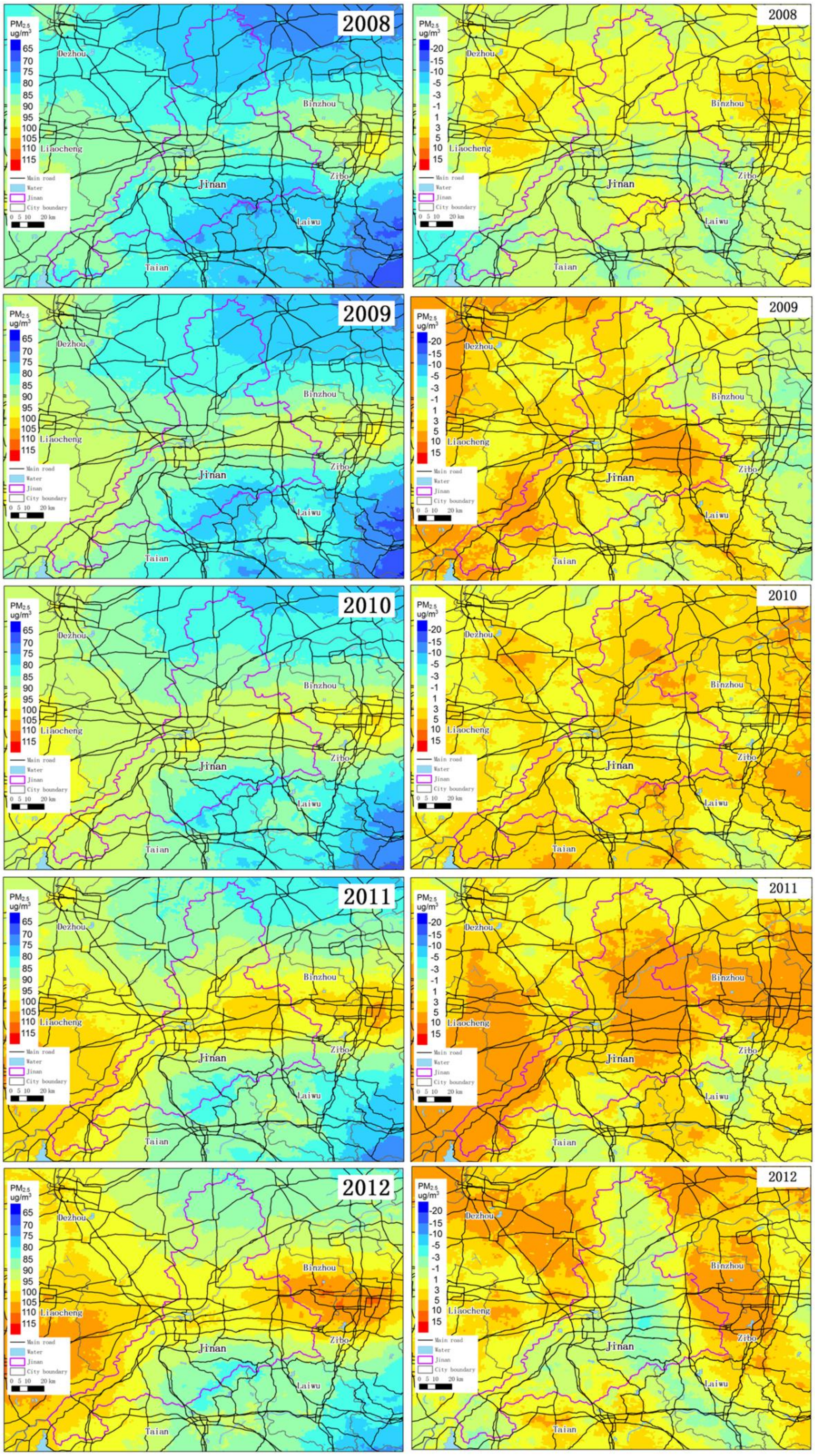


图 12. 以浓度（微克/立方米）为单位的济南市季节性 $PM_{2.5}$ 空间分布。

5. 济南市 $PM_{2.5}$ 历史浓度

图 13 展示了济南市 2003-2016 年间的历史 $PM_{2.5}$ 浓度分布及变化特征。我们观察到在 2005-2006 以及 2012-2013 出现了两次主要的 $PM_{2.5}$ 浓度升高，年平均 $PM_{2.5}$ 浓度分别升高了 11 $\mu g/m^3$ 和 4 $\mu g/m^3$。2005-2006 的 $PM_{2.5}$ 浓度升高主要发生在济南市东北部而 2012-2013 的 $PM_{2.5}$ 浓度升高主要发生在济南市市中心和北部地区。从 2007-2013，重污染区域有向济南市北部移动的趋势。2013 年之后，整个济南的污染物浓度不断下降。然而，2015-2016 年，济南市东北的污染热点出现污染物浓度的上升，同时济南市中心的 $PM_{2.5}$ 浓度仍然有显著的下降。由于济南市东北的污染热点未被地面监测网覆盖，该现象提示缺乏监管的污染热点可能在 2016 年出现污染排放的增加。此外，我们注意到相比于济南市的污染热点，周边相邻地区的污染热点的 $PM_{2.5}$ 浓度下降幅度更大，显示济南作为省会城市伴随大量人口流入、城市建设和工业发展，对空气污染的控制面临更大挑战。


2003
2004
2005
2006
2007
PM2.5 ug/m3
65
70
75
80
85
90
95
100
105
110
115
-20
-15
-10
-5
-3
-1
1
3
5
10
15
Main road
Water
Jinan
City boundary
0 5 10 20 km
Dezhou
Binzhou
Liaocheng
Jinan
Zibo
Laiwu
Taian


2008
2009
2010
2011
2012
PM2.5 ug/m3
65
70
75
80
85
90
95
100
105
110
115
-20
-15
-10
-5
-3
-1
1
3
5
10
15
Main road
Water
Jinan
City boundary
0 5 10 20 km
Dezhou
Binzhou
Liaocheng
Jinan
Zibo
Laiwu
Taian

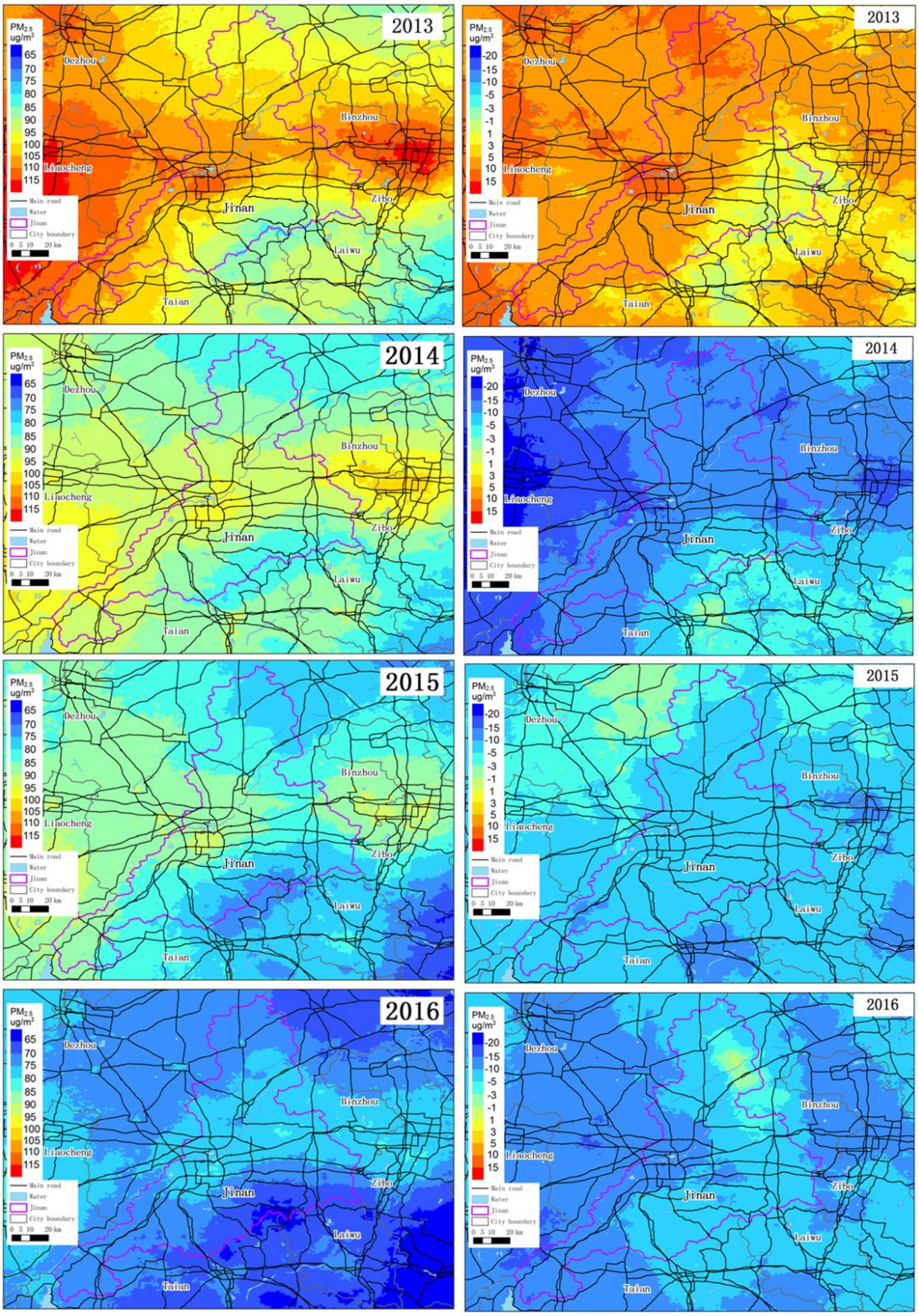


图 13. 济南 2003-2016 年平均 $PM_{2.5}$ 浓度（左）和与前一年相比 $PM_{2.5}$ 浓度的变化（右）。

6. 人口密度加权平均的 $PM_{2.5}$ 浓度

之前的研究表明，短期和长期的 $PM_{2.5}$ 暴露均会造成不良健康影响，包括呼吸系统疾病、心血管疾病及肺癌的风险提高等。为了直观表现济南市空气污染对公共健康造成的影响，

我们计算了人口密度加权平均的 $PM_{2.5}$ 浓度。权重为建模网格的人口密度/济南市平均人口密度。图 14 展示了济南市人口密度加权平均的 $PM_{2.5}$ 浓度历史变化趋势。人口密度加权平均的 $PM_{2.5}$ 浓度高于算术平均的 $PM_{2.5}$ 浓度（平均高 3.4 μg/m^3），显示济南市人口集中的区域和 $PM_{2.5}$ 污染水平高的区域比较一致。我们注意到，在 $PM_{2.5}$ 浓度的两次明显上升中，2005-2006 年的上升期间人口密度加权平均的 $PM_{2.5}$ 浓度和算术平均的 $PM_{2.5}$ 浓度升高幅度一致，分别升高了 12 μg/m^3 和 11 μg/m^3；而在 2012-2013 年的上升期间人口密度加权平均的 $PM_{2.5}$ 浓度升高幅度显著大于算术平均的 $PM_{2.5}$ 浓度升高幅度，分别升高了 7 μg/m^3 和 4 μg/m^3。这一结果说明，2005-2006 年的污染加重在整个区域发生，而 2012-2013 年的污染加重集中发生在人口聚集区域。

图 15 展示了 2016 年济南市人口密度加权平均的 $PM_{2.5}$ 浓度的空间分布。与算术平均的 $PM_{2.5}$ 浓度（图 8）相比，济南市东北部地区的污染热点由于人口稀疏，加权平均的 $PM_{2.5}$ 浓度较低。济南市南部沿主要道路人口聚集，显示较高的加权平均值。济南市市中心既是污染热点也是人口聚集区，因此有最高的加权平均污染物浓度。

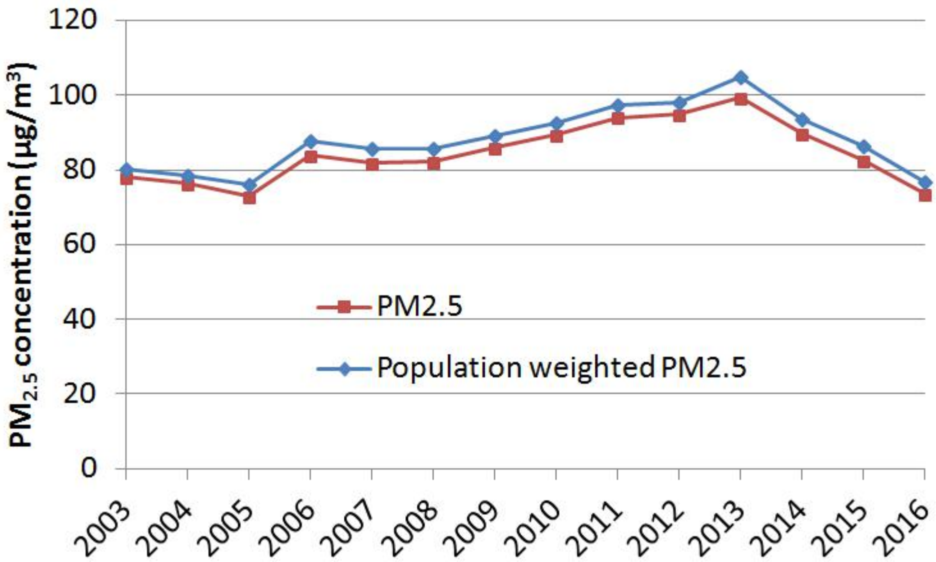


图 14 济南市年人口密度加权平均的 $PM_{2.5}$ 浓度和算术平均 $PM_{2.5}$ 浓度的时间趋势。

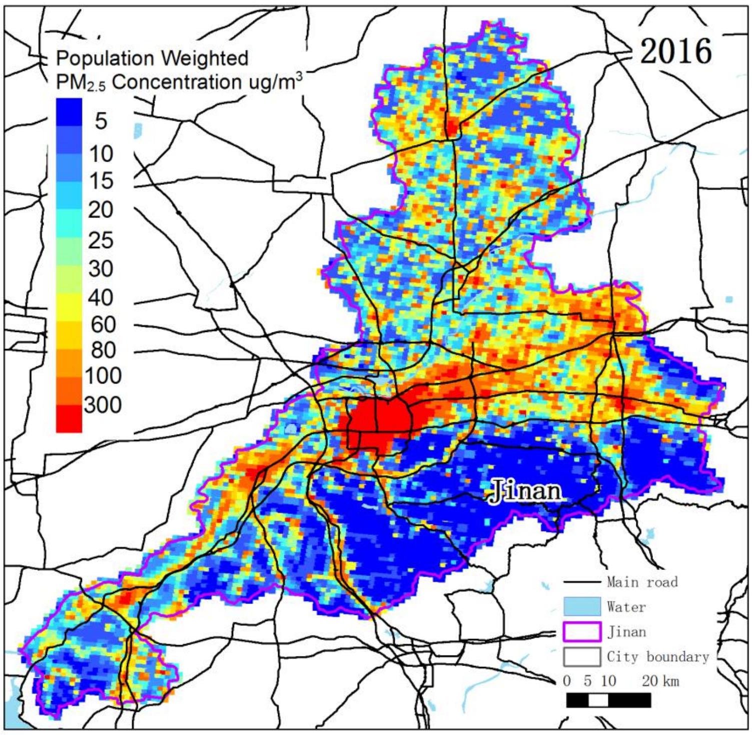


图 15 济南 2016 年人口密度加权平均的 $PM_{2.5}$ 浓度。

7. 可能影响 $PM_{2.5}$ 浓度的因素

7.1 气象因素

之前的研究报告了气象因素，例如风速、湿度、边界层高度等对短期 $PM_{2.5}$ 浓度波动有显著影响，并且参与触发重污染事件。我们也发现气象因素在月平均 $PM_{2.5}$ 预测模型中有显著贡献。然而，对长期 $PM_{2.5}$ 波动进行分析发现，年平均 $PM_{2.5}$ 浓度与气象因素无显著相关性（图 16）。年平均 $PM_{2.5}$ 浓度与温度、气压、边界层高度和风速的相关系数分别为-0.15、-0.21、0.31 和 0.20。虽然在 2015-2016 年 $PM_{2.5}$ 浓度的显著上升伴随着气象数据的较大变化，包括温度上升以及气压和风速的下降，但是类似的气象场变化在其他年份并没有伴随 $PM_{2.5}$ 浓度上升，因此气象场变化对 2015-2016 年的 $PM_{2.5}$ 浓度升高的作用并不明确。长期来看，$PM_{2.5}$ 浓度的长期波动，尤其是 2013 年之后的浓度下降，更可能归因于排放源的变化。

同时，我们也分析了气象因素，主要是风场，对济南市 $PM_{2.5}$ 传输的影响。我们对 2016 年日平均的 $PM_{2.5}$ 浓度进行建模，并结合风场数据对济南市 $PM_{2.5}$ 变化进行分析（图 17）。总的来看，济南市污染以本地为主，风速较大时污染物浓度较低，显示风的主要影响是增强污染物的扩散。秋季的西风风速上升伴随 $PM_{2.5}$ 浓度的升高，显示济南西部的污染热点区域的污染物传输对于济南 $PM_{2.5}$ 污染有一定的贡献。冬季的西南风在风速小于 8 米/秒时伴随 $PM_{2.5}$ 浓度的升高，但风速大于 8 米/秒时 $PM_{2.5}$ 浓度下降。由于济南市东南方向有山地阻碍污染物输入，与西南风相关的 $PM_{2.5}$ 浓度可能由于济南市内部的污染热点在区域内的污染传输。

之后，我们分析了风速较大时济南市 $PM_{2.5}$ 的可能传输通道。由于济南市位于平原地区，周围地势平缓，只在东南行政边界上有海拔低于 800 米的山脉（图 1），因此山脉对污

染物扩散的限制较弱，污染物在各个方向均可以传输，传输方向主要依赖于风向（图 18）。2016 年第 319 天，济南地区盛行东北风。济南市东部的污染热点的污染物沿风向进入济南市东部造成局部的高污染区。2016 年第 335 天，济南盛行西风。济南市西部的污染物传输进入济南市。从图中可以看到，济南市东南的山地对污染物传输有一定的阻碍作用。2016 年第 3 天和第 4 天出现了由东北风驱动的污染扩散过程，污染物浓度大幅降低。这一过程中济南市全境都受到污染物扩散的影响。综上所诉，济南周围地形平坦，不存在特殊的污染物输送通道。东南部山地对污染物扩散略有阻碍。由于气象场数据在空间精度较低，因此局部地区小气象场对污染物传输的影响仍需要进一步研究。

同时，我们汇总了 2012-2016 年五年间月平均的 PM$_{2.5}$ 浓度数据和风场数据（图 19）。长期分析的结果和 2016 年类似，秋季的西风和升高的 PM$_{2.5}$ 浓度相关，其他季节没有明显的污染物传输。

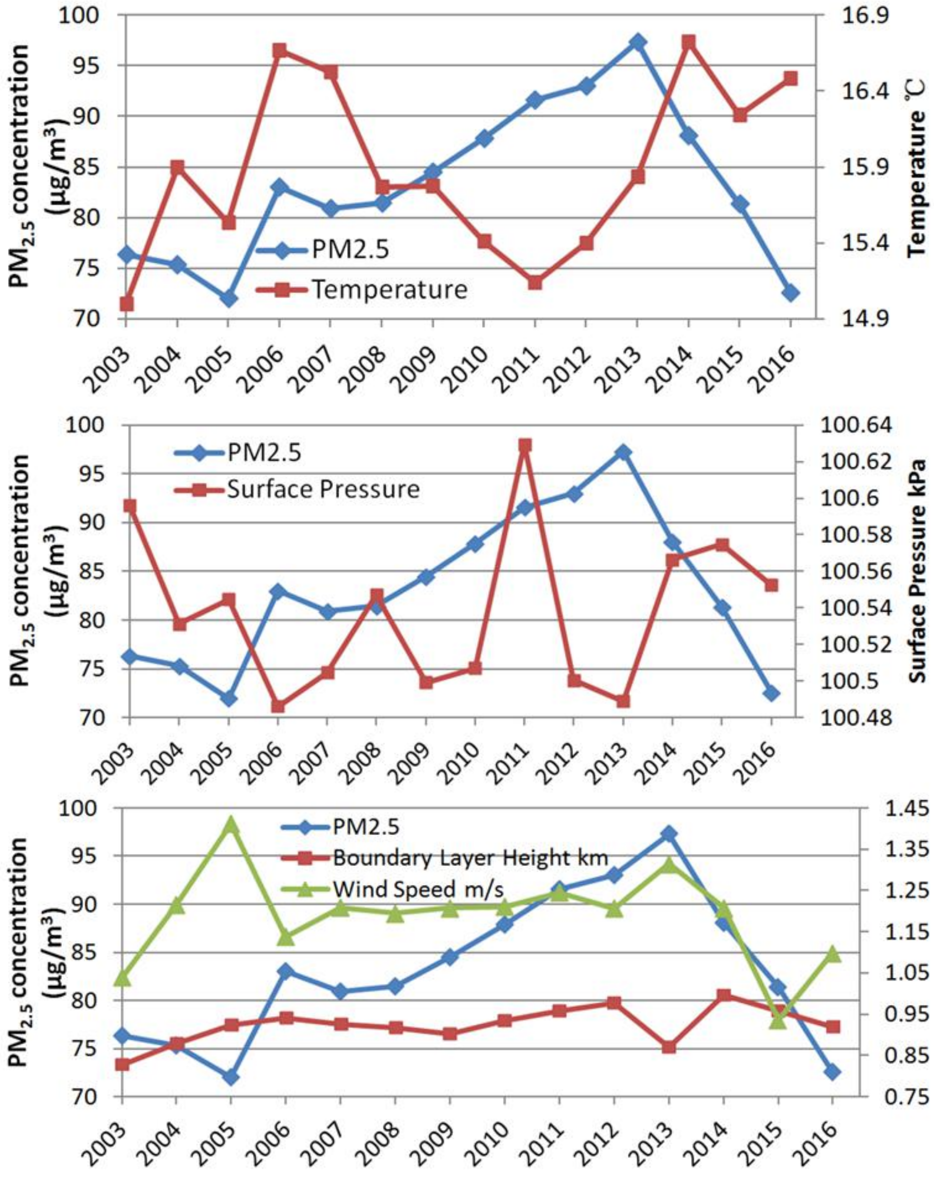


图 16 济南市年平均 PM$_{2.5}$ 浓度和气象因素的时间趋势。

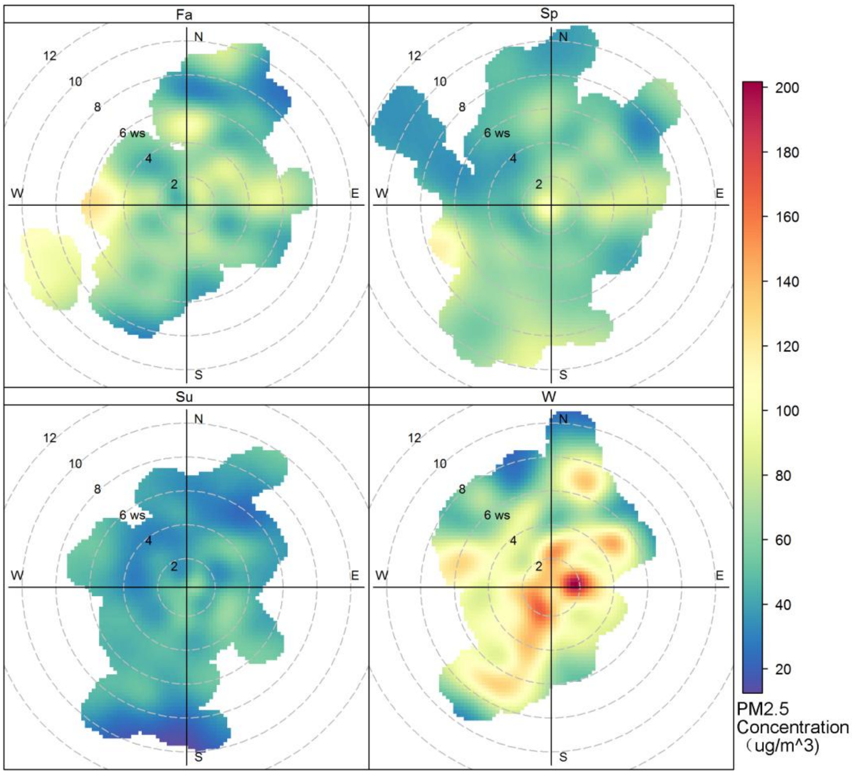


图 17 济南市 2016 年分季度日平均 PM2.5 浓度与风场的热力图。

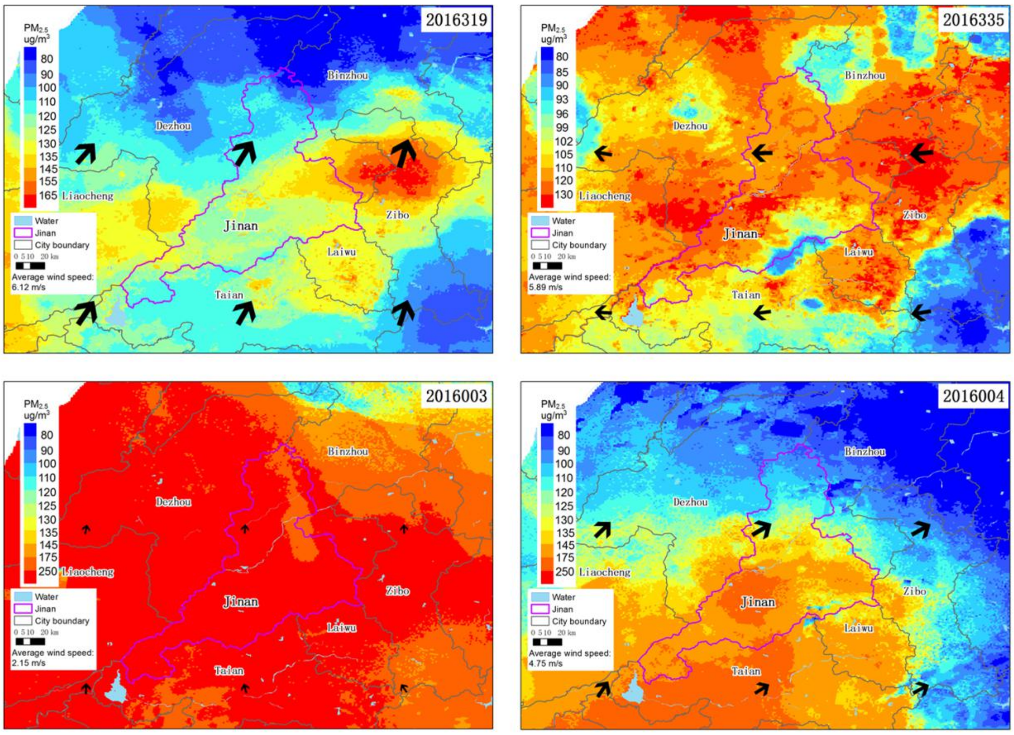


图 18 济南市 2016 年部分天日平均 PM2.5 浓度分布和风场分布。

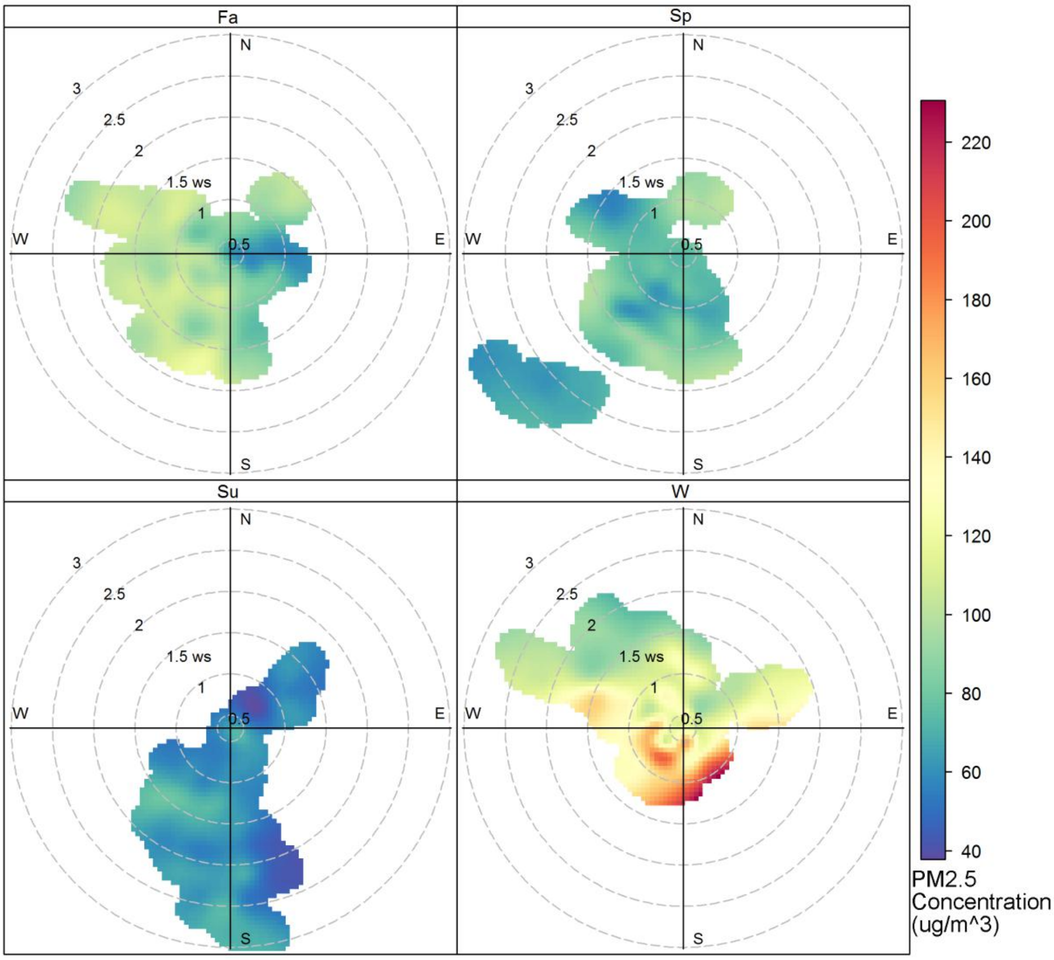


图 19 济南市 2012-2016 年分季度月平均 $PM_{2.5}$ 浓度与风场的热力图。

7.2 其他因素

我们比较了济南市年平均 $PM_{2.5}$ 浓度和济南市统计年鉴中的多种参数，包括总用电量、工业用电量、居民用电量、集中供暖面积、以及工业颗粒物排放。工业用电量和年平均 $PM_{2.5}$ 浓度正相关。在 2003-2016 年间相关系数为 0.68（图 20），特别是在 2005 年之后，相关系数达到 0.84。$PM_{2.5}$ 浓度的变化相比于工业用电量有大约一年的时延。在 2012 年，总工业用电量出现下降，之后保持稳定；$PM_{2.5}$ 浓度在 2013 年后持续下降。这一变化模式提示 $PM_{2.5}$ 浓度的下降得益于工业生产的下降（工业生产外迁）以及对工业废气的处理后排放。

对比济南市 $PM_{2.5}$ 浓度的时间趋势和我国其他重污染地区的 $PM_{2.5}$ 浓度趋势（图 21），可以看到华北地区的济南、京津冀和汾渭平原 $PM_{2.5}$ 浓度在 2005-2006 年均出现显著上升，长三角和四川盆地则没有这一趋势。在 2012-2013 年，济南、京津冀和长三角的 $PM_{2.5}$ 浓度均出现显著上升，汾渭平原和四川盆地则没有这一趋势。$PM_{2.5}$ 浓度的变化可能与这些地区的经济发展、产业结构调整、以及气象场变化相关。

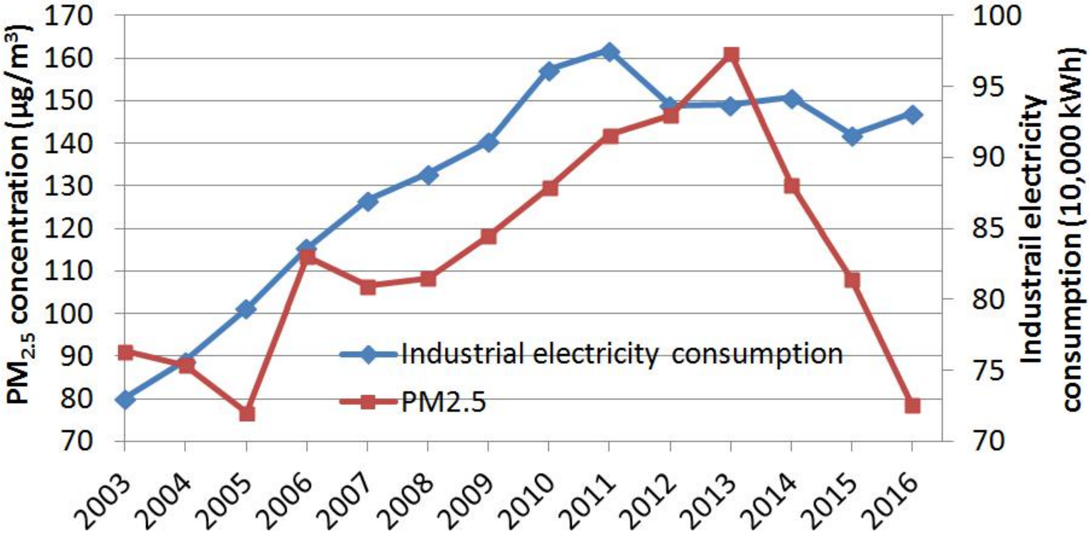


图 20 济南市年平均 $PM_{2.5}$ 浓度和工业用电量的时间趋势。

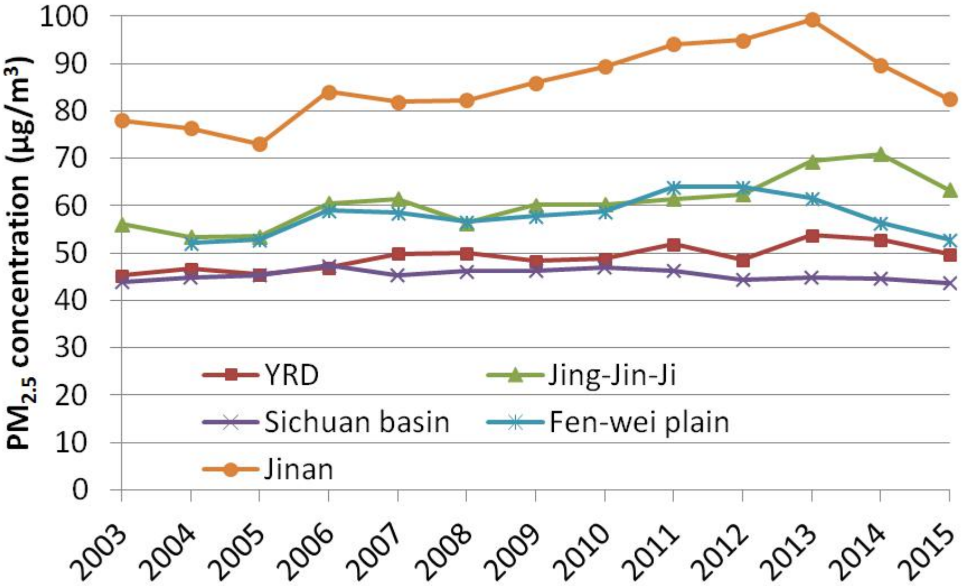


图 21 济南市和其他污染重点地区（长三角（YRD）、京津冀（Jing-Jin-Ji）、四川盆地、汾渭平原）的污染水平变化历史趋势对比。

8. 讨论

8.1 利用卫星遥感数据监测 $PM_{2.5}$ 的方法优势和局限性

卫星遥感数据在近几年被越来越多的应用于大气污染物监测。相比如地面监测站，卫星数据具有空间精度高、覆盖范围广、历史数据连续性好以及运行成本低等优势。本研究使用的所有数据均可以公开获取，方法普适性强，可以扩展到其他省市和地区。卫星数据提供的污染物浓度预测可以支持评估人群长期暴露水平、分析污染物时空变化规律、以及发现局部排放源等工作。同时，引入卫星遥感数据有助于我国数字化生态环境监测网络的建立。

卫星数据反演 $PM_{2.5}$ 浓度的方法也存在一定的局限。，卫星数据反演 $PM_{2.5}$ 浓度的准确性比地面测量低，时间精度也达不到地面连续监测的要求。下一步我们的研究希望整合极轨卫星和定轨卫星的数据，提高反演的时间精度，从而更好的对重污染事件的发展变化和污染物传输进行监控。此外，卫星数据提供的是对现实情况的观测，无法对可能的污染情景进行模拟，因此在评估气象条件对短期重污染事件的影响及污染预报预警等方面存在不足。通过结合卫星数据及大气化学模式数据，可以对大气化学模式模拟结果进行校准，筛选输入数据，提高模拟质量

8.2 政策建议

我们发现，虽然现有监测站点覆盖了城市中心人口密集区域，但是未能完全覆盖 $PM_{2.5}$ 热点地区，尤其对市郊的工业排放源监测不足。这种监测的不平衡可以部分解释市中心的污染热点和市郊的污染热点在近几年的不同变化趋势。因此，增加监测站点或者调整现有监测站点的位置以加强对污染热点地区的大气污染监测十分必要，有助于进一步加强对大气污染物排放的监管。

由于济南市地处平原，没有明显的污染物传输通道；并且济南市污染水平较高，是区域 $PM_{2.5}$ 热点，因此济南市的 $PM_{2.5}$ 污染以本地排放为主。在高污染事件时，济南本地污染物浓度和邻近地区污染物浓度均较高，存在地区混合效应。因此区域联防联控对于控制高污染事件的发生有积极意义。

同时我们观察到，随着对工业排放的控制，主要道路附近的 $PM_{2.5}$ 浓度相对较高，尤其在城市郊区这一现象更明显。我们认为进一步分析交通流动源对城市 $PM_{2.5}$ 污染的贡献以及加强对交通流动源排放的监管十分必要。

参考文献

Van, Donkelaar A, et al. "Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. " *Environmental Science & Technology*50.7(2016):3762.

Understanding the regional air pollution in Jinan based on satellite-derived high-resolution $PM_{2.5}$ concentration data

Planet Data Inc.

June, 2019

I. Background

Jinan, a major city in China, suffers from severe air pollution that leads to adverse health effects. Detailed analyses of temporal and spatial variations in $PM_{2.5}$ concentrations in Jinan could support policy making and benefit air pollution control. However, previous studies in China rely on $PM_{2.5}$ measurements from the national air quality monitoring network that is unavailable before 2012 and rarely covers rural regions. Thus, it is hard to conduct a comprehensive analysis of the long-term spatiotemporal characteristics of air pollution in Jinan as well as other cities in China. To estimate historical $PM_{2.5}$ levels in Jinan, we developed a $PM_{2.5}$ prediction model including satellite remote sensing retrievals, chemical transport model (CTM) simulations, meteorological information and land use variables. We constructed the prediction model with machine learning algorithms that better describe the non- linear complex relationship between $PM_{2.5}$, satellite data, and other predictors. Our model provide satellite-based $PM_{2.5}$ predictions with high spatial resolution (1 km), long time period (2003-2016), and complete coverage. Taking advantage of the satellite-base $PM_{2.5}$ predictions, we have the opportunity to analyze the long-term trend and local-scale variations in $PM_{2.5}$ levels in Jinan. The methodology developed in this project can serve as a roadmap to analyze air quality characteristics in other regions lacking air quality monitoring. Accurate assessment of historical air pollution levels would support policy evaluation and air quality management.

II. Methods

3. Data collection and processing

To obtain sufficient data for model training, we extended the study domain to cover Shandong Province (Figure 1). The latest AOD data retrieved by the Multi-Angle Implementation of Atmospheric Correction (MAIAC) algorithm from measurements of the Aqua (crossover at 1:30 pm local time) and Terra (crossover at 10:30 am local time) Moderate Resolution Imaging Spectroradiometer (MODIS) instruments during 2003 to 2016 were obtained in this study (Lyapustin et al. 2011a; Lyapustin et al. 2011b). The fine spatial resolution (1 km) and high accuracy of MAIAC AOD makes it possible to characterize local-scale aerosol heterogeneity (Emili et al. 2011). MAIAC algorithm uses time series analysis to characterize spectral surface reflectance which is required for aerosol retrievals. The combination of the time series and spatial analysis helps improve quality of cloud and snow detection.

MAIAC provides quality assurance (QA) flags indicating the retrieval quality, including cloud mask, land/water/snow mask and adjacency mask (i.e., proximity to cloud or snow). MAIAC pixels that were cloud contaminated or covered with snow were excluded (Kloog et al. 2015). To improve the coverage of MAIAC retrievals, a linear regression between daily Aqua and Terra MAIAC AOD was fitted and the regression coefficients were used to estimate missing Aqua/Terra AOD when only one of them is present. Then the observed and predicted AOD values were averaged to reflect daily aerosol loadings (Jinnagara Puttaswamy et al. 2014) and monthly average AOD values were calculated to serve as a major predictor in the $PM_{2.5}$ prediction model. We constructed a 1 km modeling grid from MAIAC AOD retrieval pixels by the Thiessen polygon algorithm for data integration. Data during 2013-2016 were collected for model training and we predicted historical $PM_{2.5}$ levels from 2003-2012 for the trend analyses.

There are 99 air quality monitoring stations in the study region (Figure 1). Hourly $PM_{2.5}$ measurements from these stations are published in real time by the China National Environmental Monitoring Center (CNEMC, http://www.cnemc.cn/). Measurements were downloaded from PM25.in (http://pm25.in/), a direct mirror of data from CNEMC. Repeated identical measurements for at least three continuous hours were removed since these measurements are likely caused by instrument malfunction (Rohde and Muller 2015). Hourly measurements less than $1\mu g/m^3$ were also removed because it is below the instruments' limit of detection. Monthly average $PM_{2.5}$ concentrations, calculated from daily average measurements, were used as output of our prediction model. Months with less than 18 (60%) valid daily measurements were excluded. $PM_{2.5}$ measurements were matched to the modeling grid by location.

Normalized Difference Vegetation Index (NDVI) data were obtained from Terra MODIS monthly global NDVI dataset at 1 km resolution (MOD13A3). The NDVI value of each grid cell was assigned as the average of NDVI pixels falling within the corresponding grid cell. Missing data in NDVI were interpolated by inverse distance weighting.

The tropospheric vertical column NO_2 density and absorbing aerosol index (AAI) data in UV light from Ozone Monitoring Instrument (OMI) was downloaded from the Goddard Earth Sciences Data and Information Services Center (https://mirador.gsfc.nasa.gov/). We extracted and processed the parameters *ColumnAmountNO2Trop* from the OMI NO2 level 2 data (OMNO2), and *AerosolIndexUV* from the OMI Aerosol Extinction Optical Depth and Aerosol Types level 2 data (OMAERO). These retrievals are at 13×24 km^2 resolution at nadir. Due to row anomaly started from 2007, retrievals with the cross track anomaly flag as nonzero were removed and oversampling was conducted to smooth the systematic noise. Regarding the NO_2 column density, the value of each 1 km grid cell was assigned as the average of samples from a 20 km-radius buffer centered on this grid cell during each season. Regarding the AAI parameters, retrievals with lower than 0.5% percentile were removed and the value of each 1 km grid cell were assigned as the average of samples from a 30 km-radius buffer centered on this grid cell during each season.

Meteorological data were obtained from European Centre for Medium-Range Weather Forecasts (ECMWF) at $0.125°\times0.125°$resolution. We extracted and processed surface pressure, temperature at 2 meter, wind speed at 10 meter, u and v component of wind speed at 10 meter, boundary layer height, relative humidity, total precipitation, surface solar radiation, and surface thermal radiation. The meteorological data were downscaled to 1 km grid cell through a daily smooth surface estimated by inverse distance weighting.

The elevation data were obtained from the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model (GDEM) version 2 at 30 m resolution. Population density data were obtained from Gridded Population of the World (GPW) at 1 km resolution.

We obtained daily $PM_{2.5}$ simulations from the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) (Randles et al. 2017). The MERRA-2 $PM_{2.5}$ simulations have complete coverage and relatively high accuracy at 0.5° latitude × 0.625° longitude resolution. Evaluation studies in the U.S. showed that MERRA-2 $PM_{2.5}$ simulations agreed well with ground measurements (Buchard et al. 2017). MERRA-2 data provided additional information on $PM_{2.5}$ distribution at broad scale. The total concentration of $PM_{2.5}$ was calculated using the following equation (Buchard et al. 2016; Provençal et al. 2017):

$$\mathrm{PM}_{2.5} = 1.375 \times \mathrm{SO}_4 + 2.1 \times \mathrm{OC} + \mathrm{BC} + \mathrm{Dust}_{2.5} + \mathrm{Sea\ salt}_{2.5}$$

where SO_4, OC, BC represent the MERRA-2 concentration of sulfate ion, organic carbon, and black carbon, respectively. $\mathrm{Dust}_{2.5}$ and $\mathrm{Sea\ salt}_{2.5}$ are the concentration of dust and sea salt with a radius less than 2.5 µm. Since MERRA-2 simulates dust and sea salt by five size bins, we summed dust concentrations of Bin 1 (radius 0.1~1.0 µm), Bin 2 (radius 1~1.5 µm), and Bin 3 (radius 1.5~3.0 µm), and sea salt concentrations of Bin 1 (radius 0.03~0.1 µm), Bin 2 (radius 0.1~0.5 µm), and Bin 3 (radius 0.5~1.5 µm). We multiplied SO_4 by 1.375 to get the concentration of sulfate aerosol, assuming that sulfate is primarily presented as ammonium sulfate. The ratio between organic carbon and organic matter, 2.1, was estimated from $PM_{2.5}$ observations and MERRA-2 organic carbon simulations in China during 2013-2016 (Xiao et al., 2018). The MERRA-2 $PM_{2.5}$ simulations at 50 km resolution was interpolated by inverse distance weighting to the 1 km modeling grid.

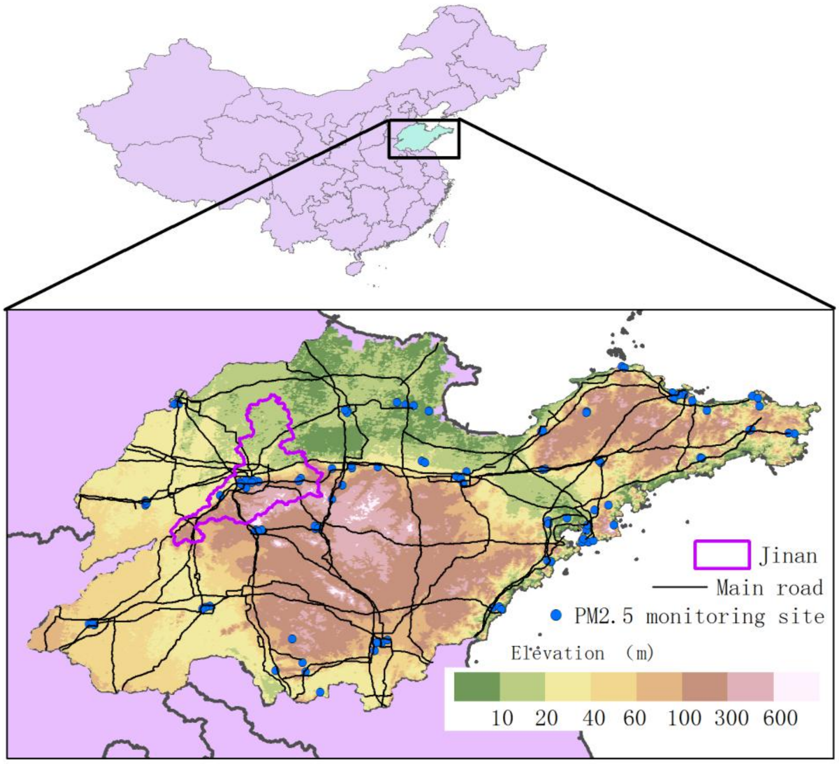


Figure 1. Map of the study domain with elevation as background. Air quality monitors are shown as blue dots and the Jinan City is highlighted.

4. Model training and evaluation

A diagram of our modeling method is shown in Figure 2. First, we trained three models, including random forest, extreme gradient boosting (XGBoost), and linear mixed effect (LME) model, separately. The random forest algorithm is a bagged classifier based on decision tree and has been widely used for classification and regression. This algorithm offers several advantages over other machine learning algorithms: it can handle a large number of features without overfitting; it allows both continuous and categorical input variables; it is robust to outliers; and it provides variable importance as well as out of bag error for model evaluation. However, the random forest model training and prediction time can increase significantly with the increase in complexity of the model. The XGBoost algorithm is developed from gradient boosting (Chen and Guestrin 2016). Gradient boosting model has been shown to outperform various statistical and machine learning models in predicting $PM_{2.5}$ levels during a wildfire event (Reid et al. 2015). XGBoost requires less training and predicting time than random forest and has been widely used in data mining competitions (Anand and Renov 2015; Mangal and Kumar 2016). The LME model has been widely used to predict $PM_{2.5}$ concentrations from satellite data in previous studies.

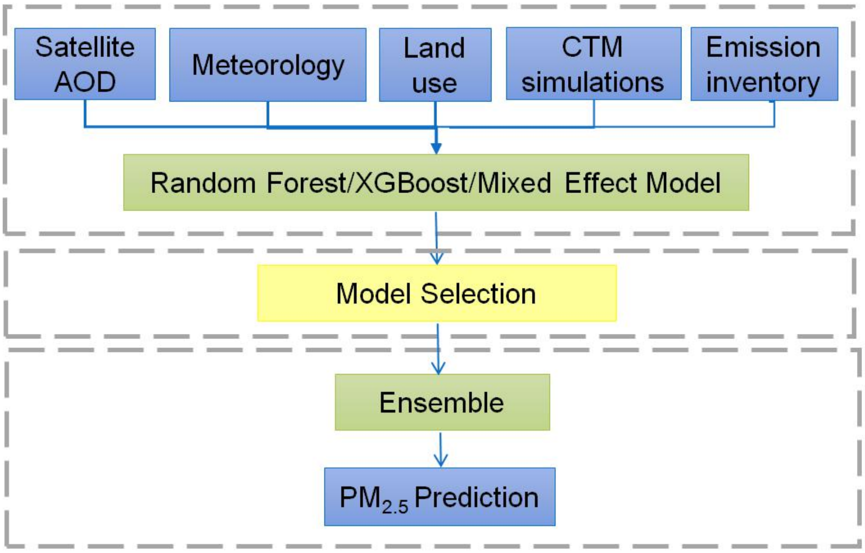


Figure 2. Modeling process.

We optimized hyperparameters of machine learning algorithms through 10-fold cross validation (CV) that we randomly selected 90% of data to train the model, and then we used the remaining 10% of data to examine the model performance. This process was repeated 10 times so that each data record was left for testing once. Since this study aims to train a model with accurate hindcast predictions, we favor low variance than low bias in the bias-variance trade off. We also selected appropriate predictors by 10-fold CV and the decision tree based algorithms, random forest and XGboost, also provided estimated importance of predictors that guided parameter selection (Hu et al. 2017; Reid et al. 2015).

Then we evaluate model performance spatially and temporally to select models with best performance. Because in the standard 10-fold cross validation, the randomly selected training dataset usually contains enough observations to estimate local spatial and temporal trends that may not hold constant outside the model fitting domain and period, we also conducted spatial 10-fold CV and by year CV to detect potential spatial and temporal over fitting. For the spatial CV, we used data from randomly selected 90% of monitors to train the models and used data from the remaining 10% of monitors to test the model. Similarly, for the by year CV, we used data from one year during the modeling period to train the model and used data from the remaining years to test the model. The spatial CV and by year CV result could better characterize the prediction error of our model.

Finally, we combined predictions from the selected individual models as the ensemble prediction to balance model accuracy and robustness. We took a weighted average of predictions from random forest and LME model. The weight was calculated from the inverse of the slope of the evaluation line in CV since the slope of the evaluation line represent the system bias of $PM_{2.5}$ predictions.

II. Results

8. Model performance

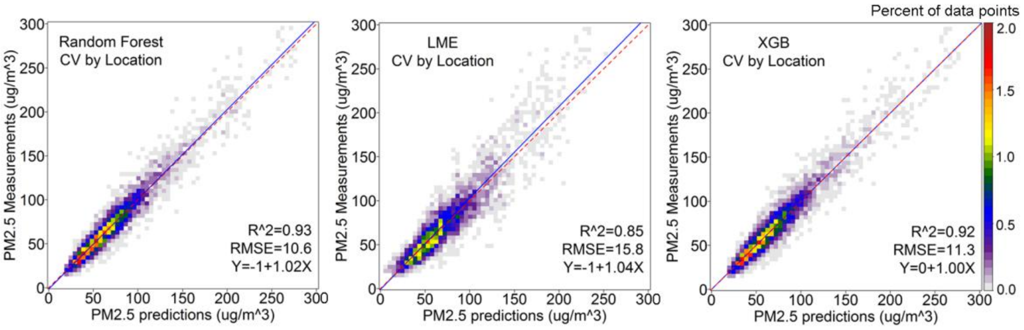


Figure 3 Density plots present model spatial cross validation results using random forest algorithm (left), linear mixed effect model (middle), and XGBoost algorithm (right). The color scale shows percent of points falling in each grid.

Figure 3 shows the spatial CV performance of individual model. The random forest model performed the best with the highest R^2 (0.93) and the lowest root mean square error (10.6 $\mu g/m^3$). The LME model had the lowest R^2 (0.85) and the highest root mean square error (15.8 $\mu g/m^3$) among these three models. All the three models provided predictions matched well with in-situ measurements and the evaluation lines were close to the standard 45°line. The slopes were close to 1 and the intercepts ranged between -1 and 0. We noticed that while generating accurate $PM_{2.5}$ estimates, decision tree based machine learning algorithms, e.g. random forest and XGBoost, have difficulties handling spatial predictors and including time-fixed spatial parameters led to unsmooth prediction maps.

Figure 4 shows the by year CV performance of each model. Consistent with previous studies, prediction models in the by year CV had lower accuracy than in other cross validations, indicating that unobserved temporal trends contributed to the prediction of $PM_{2.5}$ (Figure 4). The intra-annual changes in $PM_{2.5}$ emission sources due to economic development and pollution control policies might affect the relationships between $PM_{2.5}$ and its predictors, but such changes were not well characterized in emission profiles and negatively affected the model performance. The LME model performed the best in the by year CV, with the highest R^2 (0.76) and lowest root mean square error (21.2 μg/m^3). The XGBoost algorithm performed the worst, partly due to the relative small training dataset in this project.

To obtained spatially and temporally robust predictions, we combined predictions from random forest model and LME model that performed the best in spatial CV and by year CV respectively. These two models characterized different aspects of the complex relationships between $PM_{2.5}$ and predictors. Various ensemble methods were employed and evaluated, including geographically weighted regression, generalized additive model, and different average methods. We selected weighted average as the ensemble method regarding the CV results. Figure 5 shows the performance of ensemble predictions in spatial CV and by year CV. The ensemble prediction balanced the spatial robustness of random forest model and the temporal robustness of LME model. The ensemble prediction provided a R^2 of 0.90, RMSE of 12.6 μg/m^3, slope of 1.04 and intercept of -2.7 in 10-fold CV. Figure 6 shows the temporal comparison between $PM_{2.5}$ measurements and $PM_{2.5}$ predictions. The ensemble prediction agreed well with measurements, with correlation coefficient of 0.96. We did not observe systemic bias of the satellite-based $PM_{2.5}$ prediction.

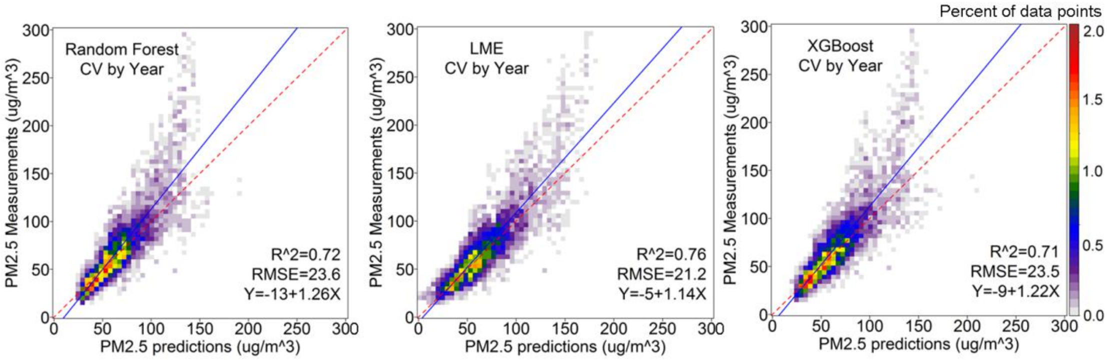


Figure 4 Density plots present the model by year cross validation results using random forest algorithm (left), linear mixed effect model (middle), and XGBoost algorithm (right). The color scale shows percent of points falling in each grid.

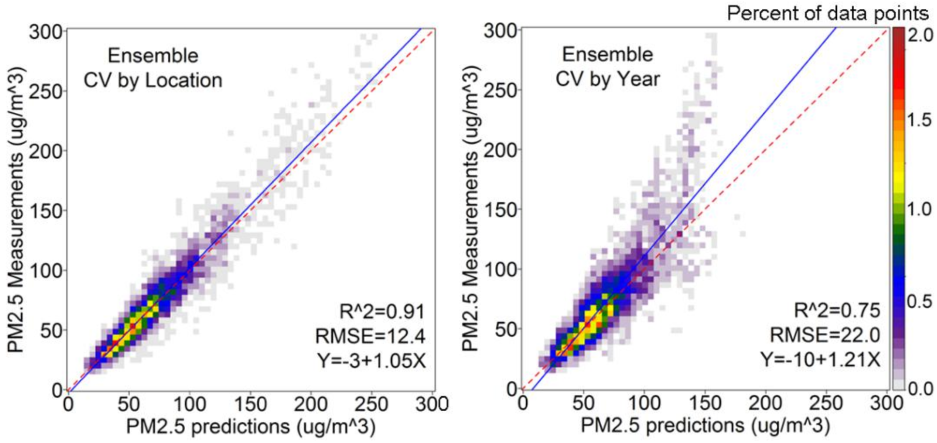


Figure 5 Density plots present the ensemble prediction performance in the spatial CV (left) and by year CV (right). The color scale shows percent of points falling in each grid.

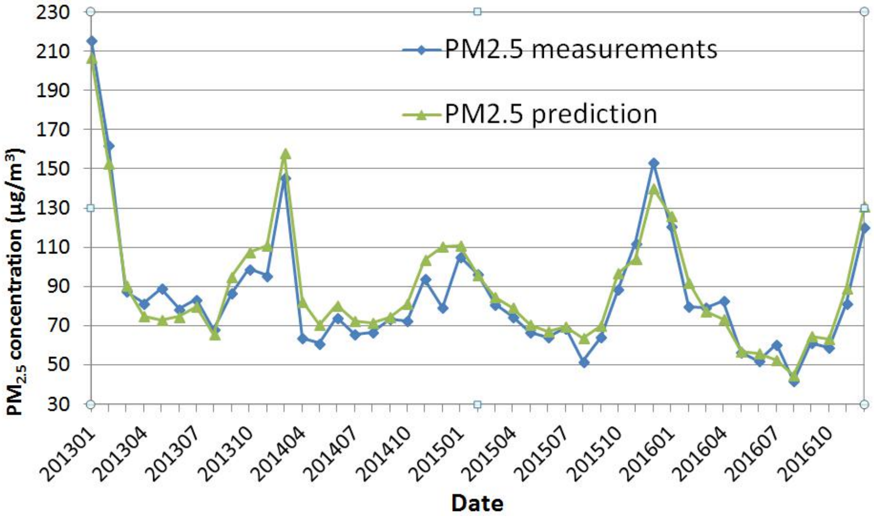


Figure 6 Temporal trend of ground $PM_{2.5}$ concentration measurements and satellite-based $PM_{2.5}$ concentration predictions.

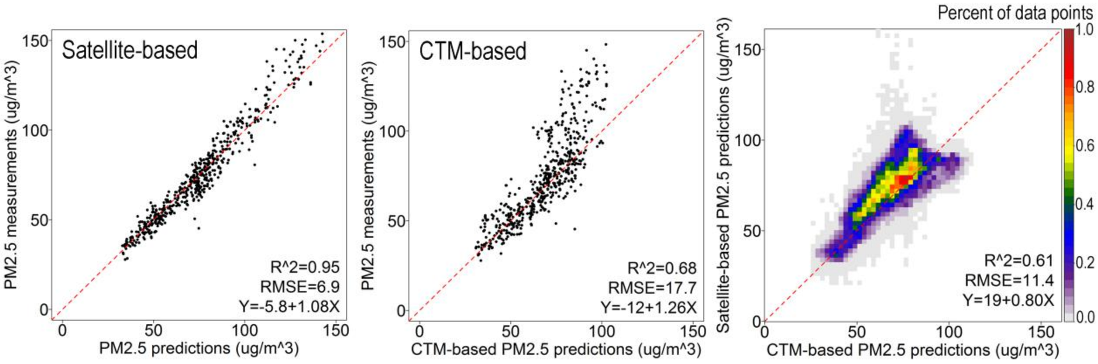


Figure 7 Density plots present the comparison results between ground measurements, satellite-based $PM_{2.5}$ predictions, and CTM-based $PM_{2.5}$ predictions. The color scale shows percent of points falling in each grid.

To obtain some information on the performance of our satellite-based $PM_{2.5}$ predictions before 2013 when the ground measurements are unavailable, we collected the 1-km annual average $PM_{2.5}$ predictions from Donkelaar et al. (2016) (website) and compared them with our model predictions (Figure 7). Donkelaar et al. (2016) provided the $PM_{2.5}$ concentration predictions that combined information from chemical transport model (CTM), ground measurements, as well as satellite data. We observed that compared to ground annual measurements during 2013-2016, the prediction error of Donkelaar et al. (R^2=0.68， RMSE 17.7 μg/m^3) was larger than the predicition error of our model (R^2=0.95，RMSE 6.9 μg/m^3). When comparing predictions from our model and from Donkelaar et al., we got the R^2 of 0.71 and RMSE of 11.4 μg/m^3. Predictions from Donkelaar et al. underestimate ground measurements as well as predictions from our model when the $PM_{2.5}$ level is high.

9. $PM_{2.5}$ spatial distribution

In 2016, the annual average $PM_{2.5}$ concentration in Jinan city was 74μg/m^3, ranging between 62 and 85 µg/m^3. As shown in Figure 8, the $PM_{2.5}$ pollution hotspots are located at city/county centers in 2016. Residents and industry were clustered around city/county centers with relatively high air pollutant emission. Most $PM_{2.5}$ pollution hotspots were covered by the national air quality monitoring networks; however, we identified some hotspots outside the monitoring network. For example, hotspots in the northeast of Jinan and to the west of Jinan (in Liaocheng city) had no ground monitoring station. The north and south regions of Jinan had relatively low $PM_{2.5}$ levels. We observed that the $PM_{2.5}$ level increased along the main roads in regions outside Jinan city, indicating that emission from vehicles significantly contributed to regional $PM_{2.5}$ levels in regions with high road density.

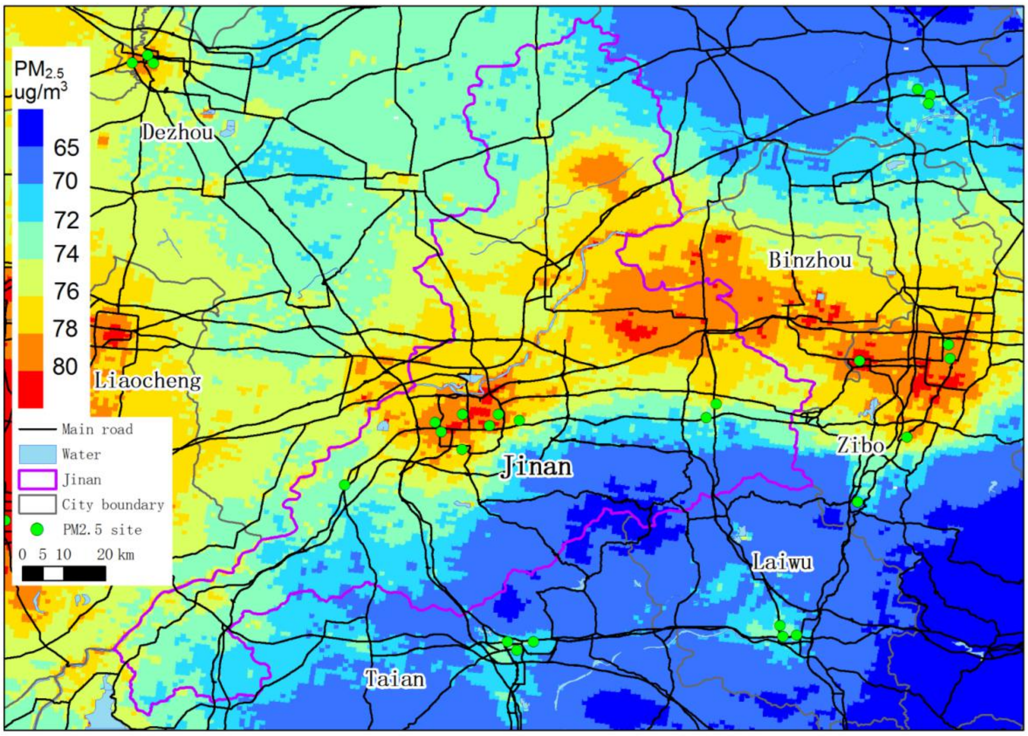


Figure 8. Annual average $PM_{2.5}$ distribution over Jinan in 2016.

10. $PM_{2.5}$ temporal trends

Figure 9 shows the temporal trends of $PM_{2.5}$ in Jinan during 2003-2016. We observed significant seasonal variations in $PM_{2.5}$ concentrations that the highest $PM_{2.5}$ levels occurred in winter and the lowest $PM_{2.5}$ levels occurred in summer. From 2007 to 2013, the winter highest $PM_{2.5}$ concentrations increased significantly from 116 to 198 µg/m^3, and the summer lowest $PM_{2.5}$ concentrations increased slowly from 47 to 71 µg/m^3. After 2013, both the winter and summer $PM_{2.5}$ concentrations decreased sharply. The annual average $PM_{2.5}$ concentrations showed increase from 2005 to 2013 (Figure 9), and declined rapidly after 2013. Benefited from various pollution control policies, the annual average $PM_{2.5}$ concentration in 2016 (73 µg/m^3) has been lower than that in 2003 (76 µg/m^3). The seasonal average $PM_{2.5}$ concentrations showed different trends: the winter and fall average $PM_{2.5}$ concentrations followed the annual average trend well, but the spring average $PM_{2.5}$

concentrations changed oppositely during 2008-2014. The summer average $PM_{2.5}$ concentrations kept constant until a increase occurred in 2009, then it kept constant and started decreasing in 2014. It is notable that different from $PM_{2.5}$ concentrations in other seasons, the winter average $PM_{2.5}$ started increasing during 2015-2016.

We analyzed the contribution of each season to the long-term variations in $PM_{2.5}$ concentrations (Figure 10). As expected, changes in winter average $PM_{2.5}$ dominated the intra-annual changes in $PM_{2.5}$. Winter $PM_{2.5}$ contributed 58% of annual $PM_{2.5}$ increase during 2005-2013, and 41% of annual $PM_{2.5}$ decrease during 2013-2016. Winter is also the only season that contributed more in $PM_{2.5}$ increase than in $PM_{2.5}$ decrease. One reason might be the increase in winter heating resulted from raised population density leading to increase in air pollutant emission in winter. Spring contributed the least in both $PM_{2.5}$ increase period and $PM_{2.5}$ decrease period, indicating that the $PM_{2.5}$ levels in spring were relatively stable during the study period.

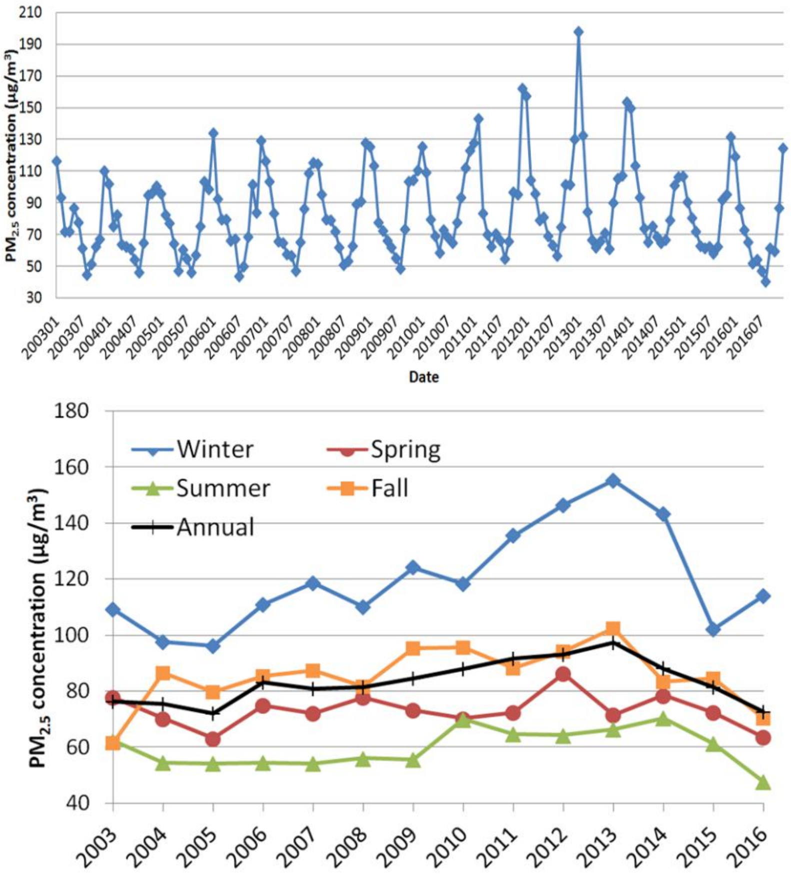


Figure 9. Monthly (above) and seasonal (below) average $PM_{2.5}$ concentrations in Jinan.

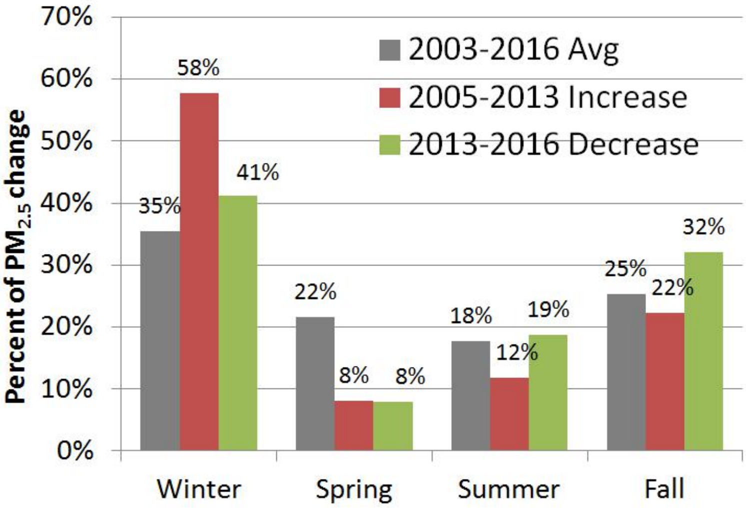


Figure 10. Seasonal contribution to $PM_{2.5}$ changes during 2003-2016.

11. $PM_{2.5}$ seasonal distribution

In order to analyze the distribution of $PM_{2.5}$ hotspots in different seasons, we plotted $PM_{2.5}$ concentration in percentile (Figure 11) and in ug/m^3 (Figure 12) in each season, because the average $PM_{2.5}$ level in each season varied considerable. Using percentile as plotting unit and defining regions with $PM_{2.5}$ concentrations above 75% percentile as hotspots could normalize the difference in average $PM_{2.5}$ levels thus to benefit comparisons across seasons. We noticed that in spring, the north region in Jinan showed low $PM_{2.5}$ levels, but in fall and winter, this region was a $PM_{2.5}$ hotspot. In fall, the hotspot covering Jinan city center shrunk and the surrounding region, especially to the northwest of Jinan, had relatively high $PM_{2.5}$ levels. Thus, $PM_{2.5}$ might transport from surround regions to Jinan and affect air quality in Jinan. Different from other seasons, in winter Jinan became the regional $PM_{2.5}$ hotspot with $PM_{2.5}$ concentrations higher than surrounding regions.

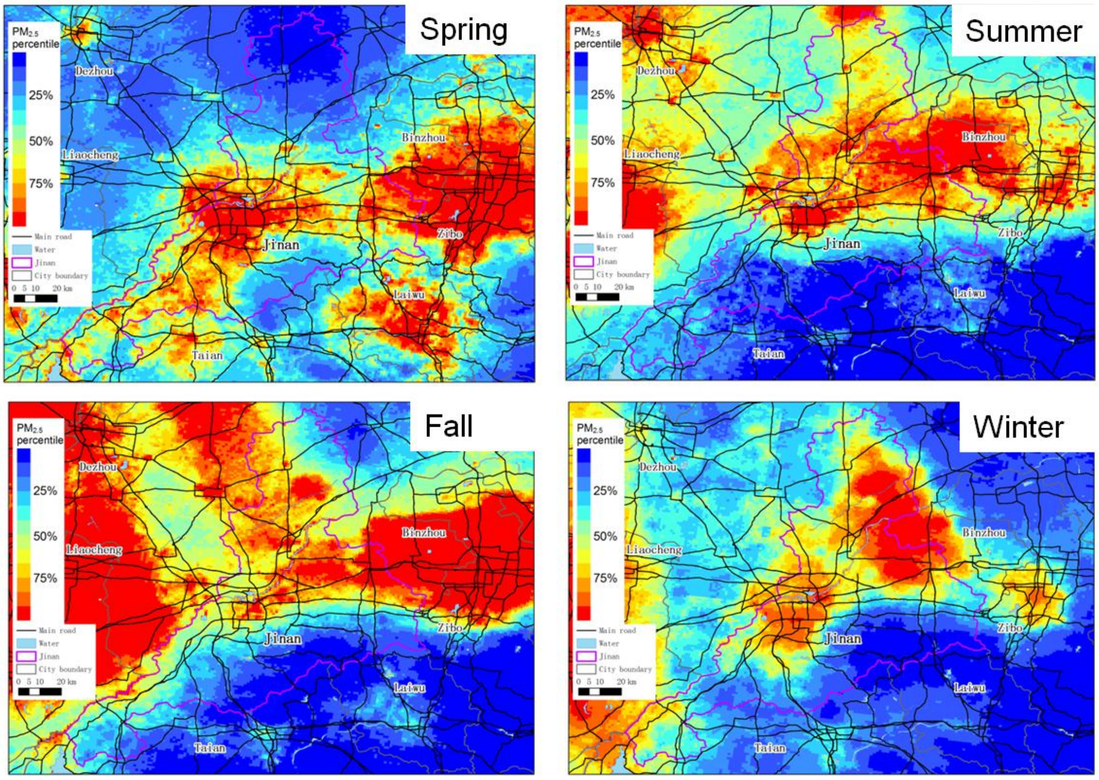


Figure 11. Seasonal $PM_{2.5}$ distribution in Jinan shown in percentile.

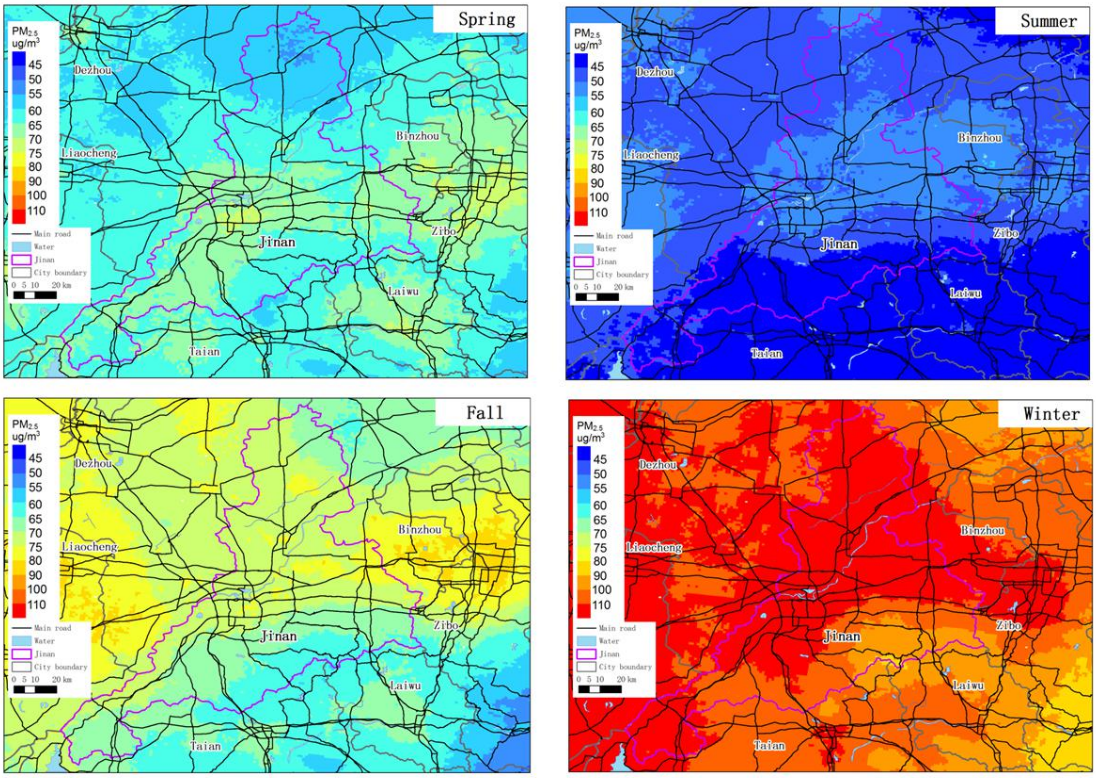
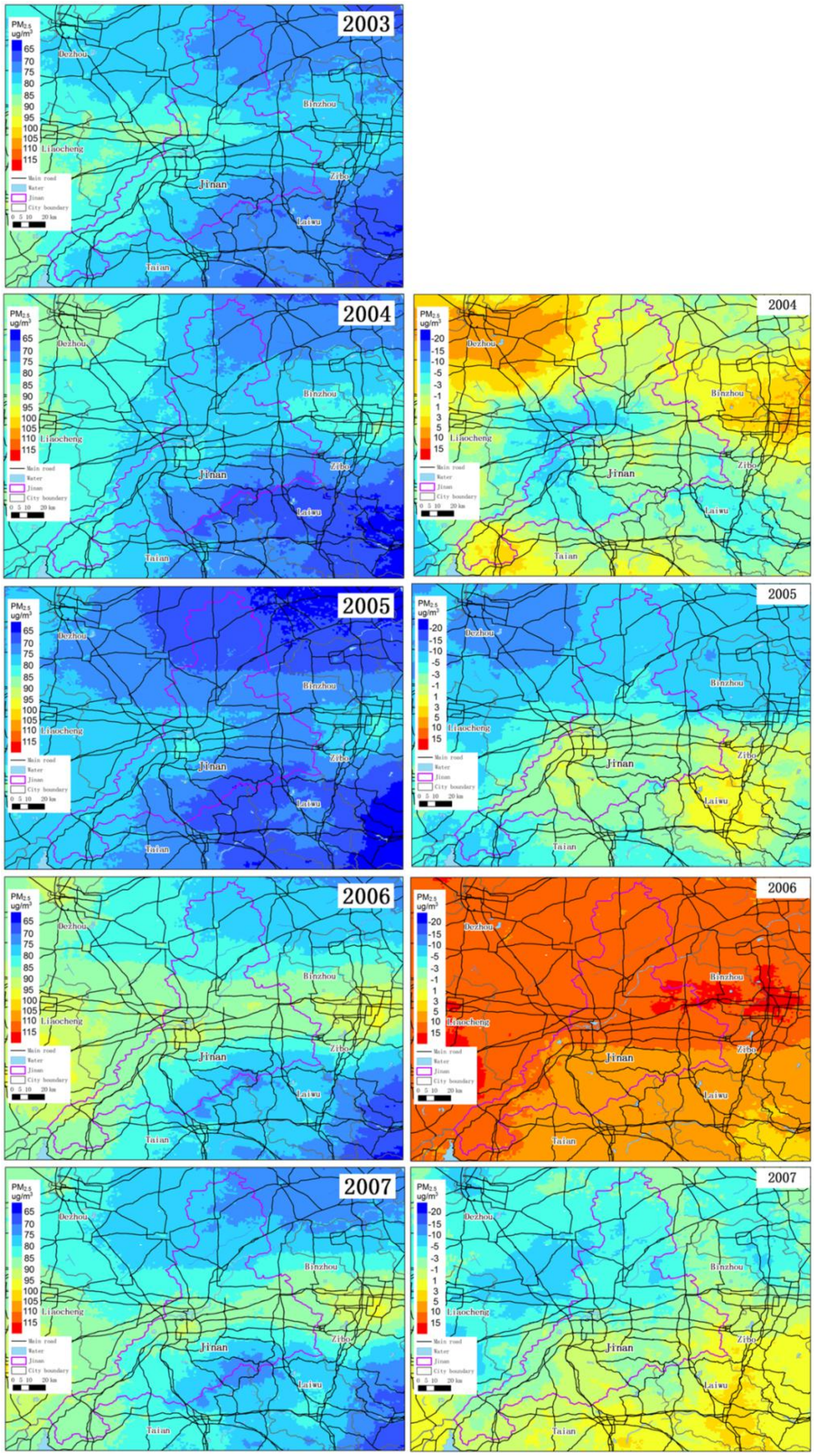
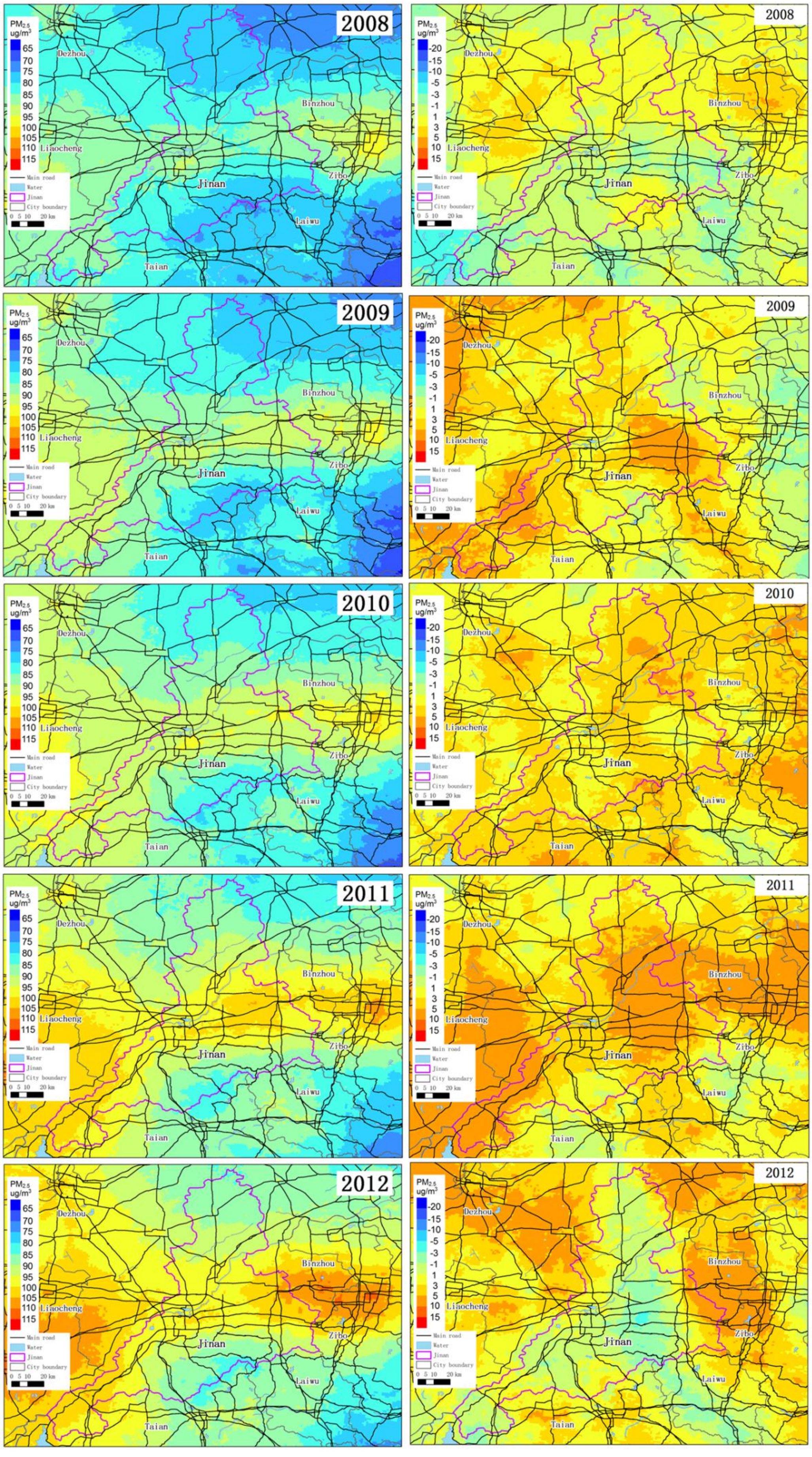


Figure 12. Seasonal $PM_{2.5}$ distribution in Jinan shown in ug/m^3.

12. Historical $PM_{2.5}$ levels and changes

Figure 13 summarized the historical $PM_{2.5}$ levels and changes during 2003-2016 over Jinan. We observed two major $PM_{2.5}$ increases occurred during 2005-2006 and 2012-2013, with annual average $PM_{2.5}$ concentrations increased by 11 and 4 $\mu g/m^3$, respectively. The increase during 2005-2006 occurred mainly in northeast Jinan and the increase during 2012-2013 occurred mainly in city center and the north region. From 2007 to 2013, the highly polluted region moved toward the north region of Jinan. After 2013, the whole study domain showed decrease in $PM_{2.5}$ concentrations. However, during 2015-2016, the hotspot in northeast Jinan was observed with increasing $PM_{2.5}$ levels while the city center was observed with a larger decrease in $PM_{2.5}$ concentrations compared to other regions. These different change patterns indicated that emissions in hotspot regions lacking monitoring need to be controlled. Additionally, we noticed that the hotspot regions outside Jinan showed larger decrease in $PM_{2.5}$ concentrations. Jinan, as the capital of Shandong province, may face more challenges in pollution control relative to other cities.


2003
2004
2005
2006
2007
PM2.5 ug/m3
65
70
75
80
85
90
95
100
105
110
115
-20
-15
-10
-5
-3
-1
1
3
5
10
15
Main road
Water
Jinan
City boundary
0 5 10 20 km
Dezhou
Binzhou
Liaocheng
Jinan
Zibo
Laiwu
Taian


2008
2009
2010
2011
2012
PM2.5 ug/m3
65
70
75
80
85
90
95
100
105
110
115
-20
-15
-10
-5
-3
-1
1
3
5
10
15
Main road
Water
Jinan
City boundary
0 5 10 20 km
Dezhou
Binzhou
Liaocheng
Jinan
Zibo
Laiwu
Taian

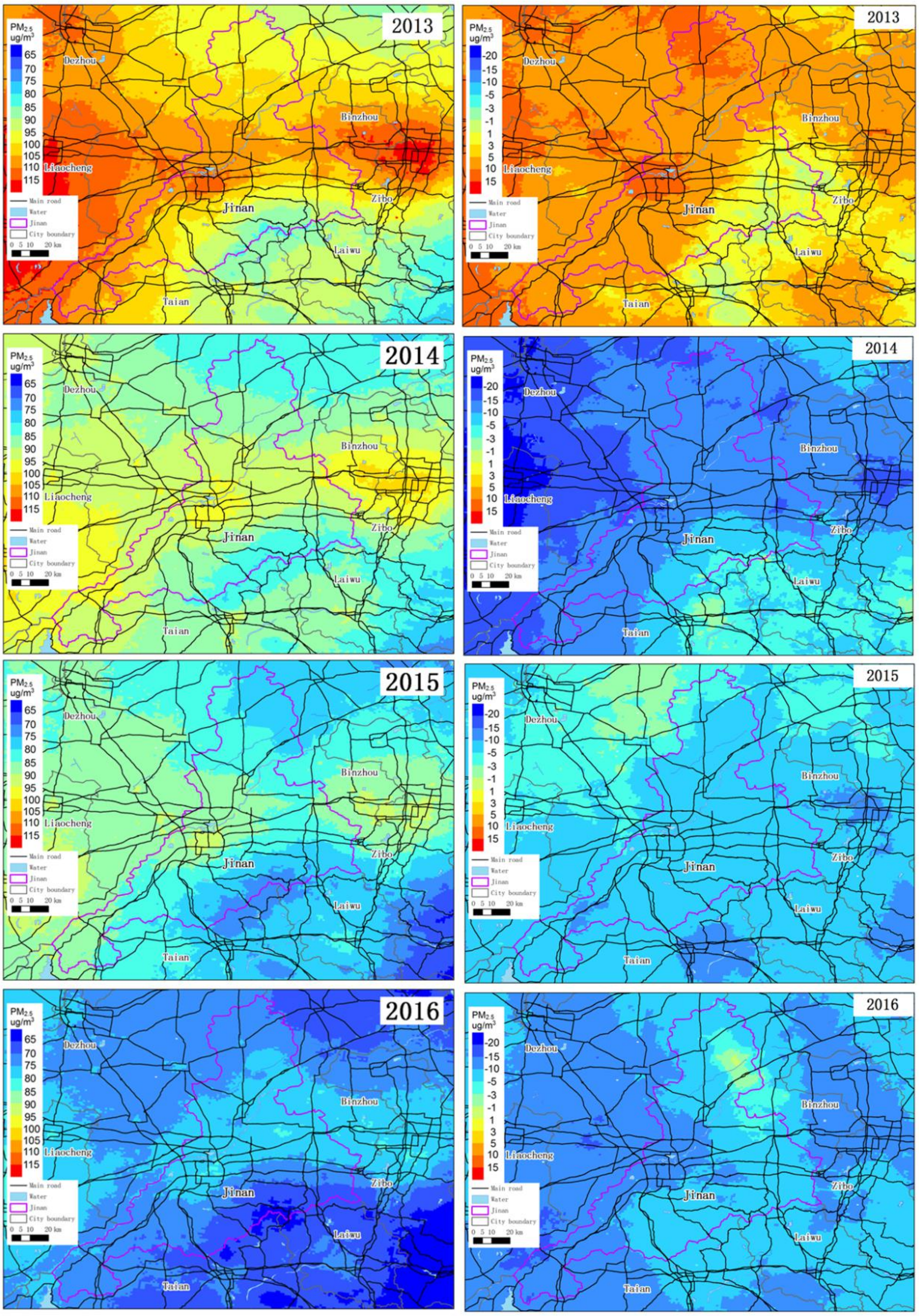


Figure 13. Annual $PM_{2.5}$ distribution (left) and $PM_{2.5}$ changes comparing to the previous year (right) during 2003-2016.

13. Population density weighted average $PM_{2.5}$ concentration

Previous studies reported that both long-term and short-term exposure to $PM_{2.5}$ was associated with adverse health effects, including increased risk of respiratory diseases, cardiovascular diseases, and lung cancer. To directly show the health burden attributable to $PM_{2.5}$ in Jinan, we calculated the population density weighted average $PM_{2.5}$ concentration. The weight is the population density of each grid divided by the average population density in Jinan. Figure 14 shows the temporal trend of the population density weighted average $PM_{2.5}$ concentration in Jinan. The population weighted $PM_{2.5}$ concentration is higher than the arithmetic average $PM_{2.5}$ concentration by 3.4 $\mu g/m^3$ on average, indicating that the population density region consistent with high $PM_{2.5}$ level region. We noticed that, regarding the two $PM_{2.5}$ increase events during 2005-2006 and 2012-2013, the increase in population weighted average $PM_{2.5}$ (12 $\mu g/m^3$) is the same as the increase in arithmetic average $PM_{2.5}$ (11 $\mu g/m^3$) during 2005-2006. However, during 2012-2013, the increase in population weighted average $PM_{2.5}$ (7 $\mu g/m^3$) is much higher than the increase in arithmetic average $PM_{2.5}$ (4 $\mu g/m^3$) during 2005-2006. This result indicates that during 2005-2006, the $PM_{2.5}$ pollution increased over the whole Jinan domain, while in 2012-2013, the $PM_{2.5}$ pollution increase mainly occurred in population density regions.

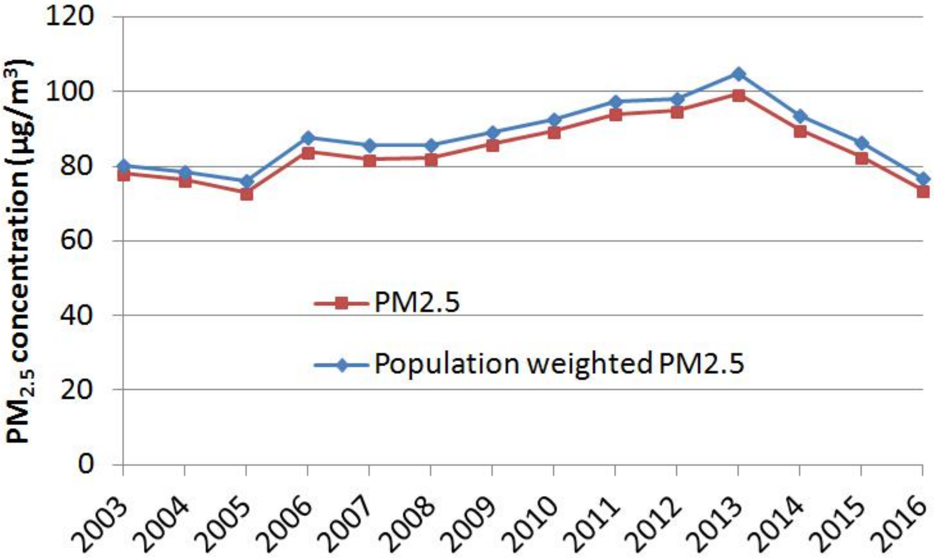


Figure 14 The temporal trends of population density weighted average $PM_{2.5}$ concentration and arithmetic average $PM_{2.5}$ concentration in Jinan during 2003-2016.

Figure 15 shows the spatial distribution of population density weighted average $PM_{2.5}$ concentration in Jinan in 2016. Compared with the spatial distribution of arithmetic average $PM_{2.5}$ concentration (Figure 8), the $PM_{2.5}$ hotspot in the northeast of Jinan had relatively low population-weighted value due to the low population density. Population clustered along the main roads in the south of Jinan, leading to a high population weighted $PM_{2.5}$ level. The city center of Jinan is both pollution hotspot as well as population center, thus showing the highest population-weighted $PM_{2.5}$ concentration.

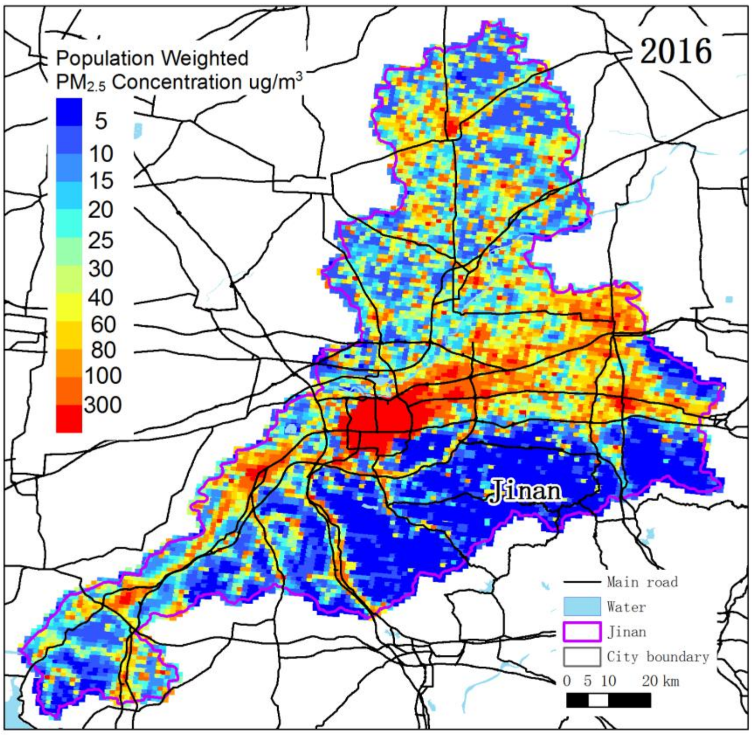


Figure 15 The distribution of population density weighted average $PM_{2.5}$ concentration in Jinan in 2016.

14. Factors that may affect $PM_{2.5}$ levels

14.1 Meteorological factors

Previous studies reported that meteorological factors affected short-term $PM_{2.5}$ concentrations and triggered haze events. We also observed that meteorological factors contributed significantly in our monthly $PM_{2.5}$ prediction model. However, regarding long-term $PM_{2.5}$ trends, we noticed that annual average $PM_{2.5}$ concentrations were significantly correlated with no meteorological factor (Figure 16). The correlation coefficients between annual average $PM_{2.5}$ concentrations and temperature, surface pressure, boundary layer height, and wind speed were -0.15, -0.21, 0.31, and 0.20, respectively. Thus, the long-term change in $PM_{2.5}$ concentrations were more likely due to changes in emission sources.

Pollutants transportation has raised researcher's attention recently. Emissions from one place may significantly affect air quality in surrounding regions. Thus, we analyzed the effect of meteorological factors, especially wind, on the $PM_{2.5}$ transportation over Jinan. We developed an additional model to predict daily average $PM_{2.5}$ concentrations in 2016 in Jinan. The model structure is the same as the monthly prediction model. We plotted the predicted daily $PM_{2.5}$ concentrations in polar coordinates by wind speed and direction in each season to show the effect of wind on $PM_{2.5}$ concentration (Figure 17). Basically, high $PM_{2.5}$ concentrations were associated with low wind speed, indicating that the $PM_{2.5}$ pollution in Jinan is majorly due to local emissions. In fall, the southeast wind was associated with elevated $PM_{2.5}$ concentration, showing that the pollution hotspot in the east of Jinan may contribute to pollution in Jinan. In winter, the weak southeast wind was associated with increased $PM_{2.5}$ concentration. Since mountains in the south of Jinan block the transportation of pollutants, this increase in $PM_{2.5}$ concentration may be due to the transportation of pollutant from the south hotspot in Jinan to the northern region of Jinan.

Then we selected days with high satellite data coverage (low cloud cover) and relatively high wind speed to detect potential $PM_{2.5}$ transport paths. Since Jinan located in Shandong Peninsula with flat terrain (Figure 1), pollutants are able to transport from almost all directions. The mountains located on the east south boundary of Jinan may slightly limit transport of pollutants over the mountain. Thus, wind direction determined the transport direction (Figure 18). On the 319th day of 2016, Jinan region was covered by northeast wind. The $PM_{2.5}$ pollution transported from the hotspot of the east of Jinan to Jinan, leading to regional high pollution levels in the east of Jinan. On the 335th day in 2016, Jinan was covered by west wind and $PM_{2.5}$ pollution transported from the west to enter Jinan. Figure 16 shows that the mountain region in the southeast of Jinan limited pollutant transportation leading to a relatively low $PM_{2.5}$ level. On the third and fourth day in 2016, we observed a strong transportation event driven by east north wind, leading to a significant decrease in $PM_{2.5}$ concentrations. During this event, the whole Jinan domain was affect by the strong wind. To summarize, the terrain of Jinan is flat without significant elevation change, thus there exists no special pollutant transportation path. The mountain region in the east south of Jinan slightly limited pollution transportation. Since the spatial resolution of the meteorological data is low, further studies are needed to analyze the effects of local meteorological fields on pollution transportation.

We also plotted the monthly $PM_{2.5}$ concentrations in polar coordinates by wind speed and wind direction in Jinan during 2012-2016 to analyze the long-term effect of wind field on $PM_{2.5}$ concentration (Figure 19). The results are similar as the daily results in 2016. The east wind was associated with increase in $PM_{2.5}$ concentration in fall. No significant transportation effect was observed in other seasons.

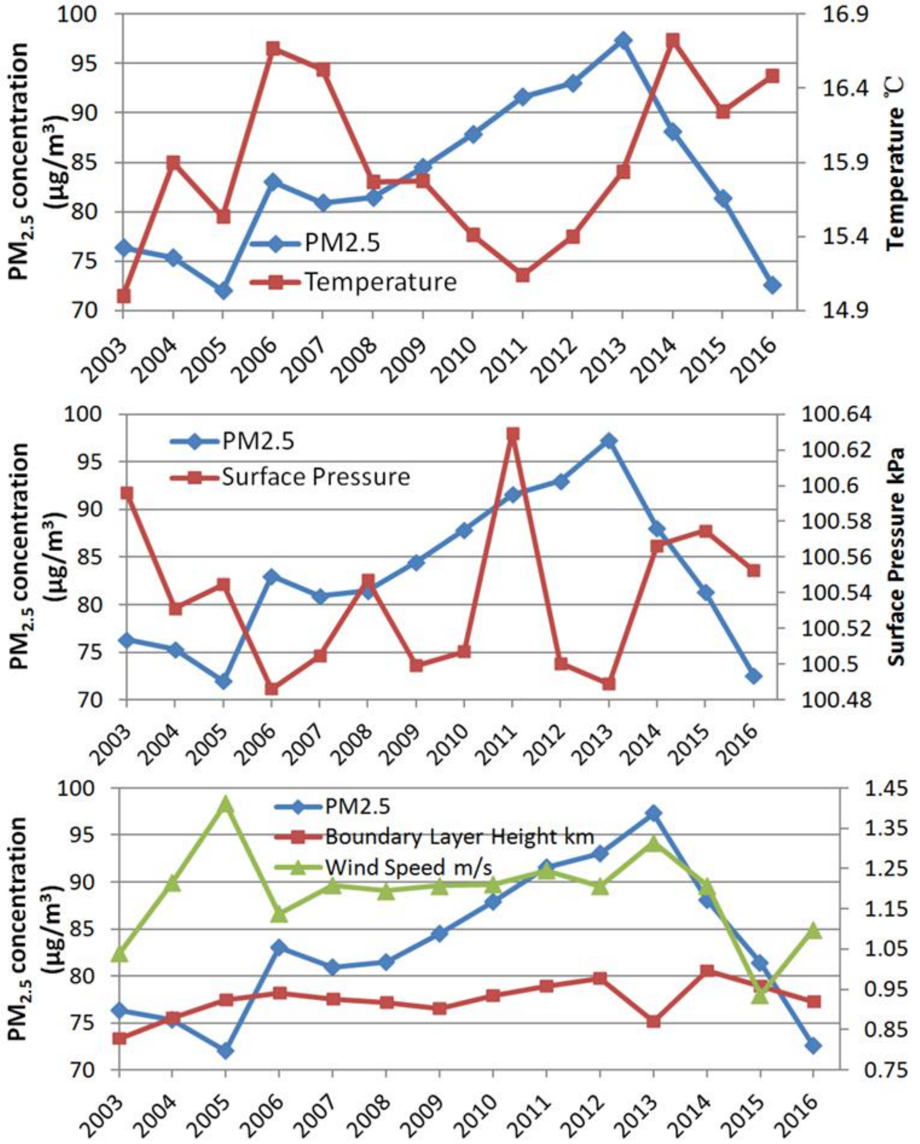


Figure 16 Temporal trends of annual average $PM_{2.5}$ concentrations and meteorological factors.

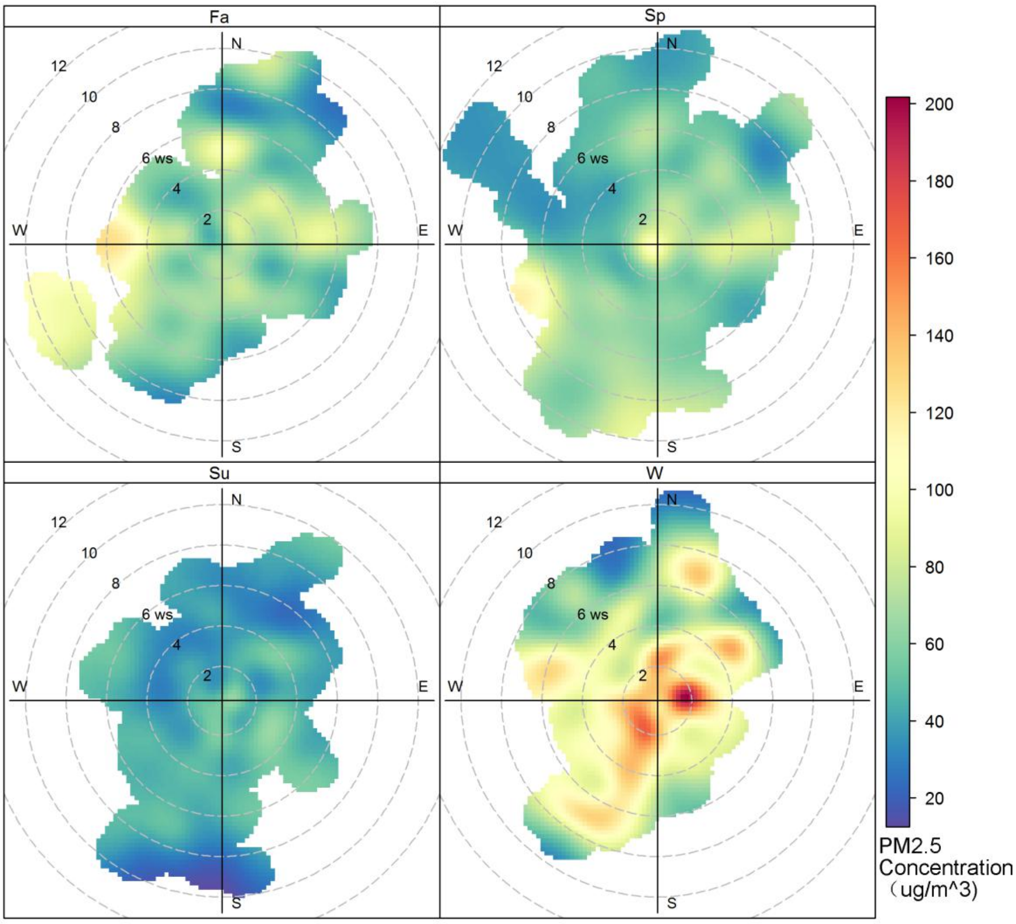


Figure 17 Seasonal analyses by plotting daily $PM_{2.5}$ concentration in polar coordinates by wind speed and direction in Jinan in 2016.

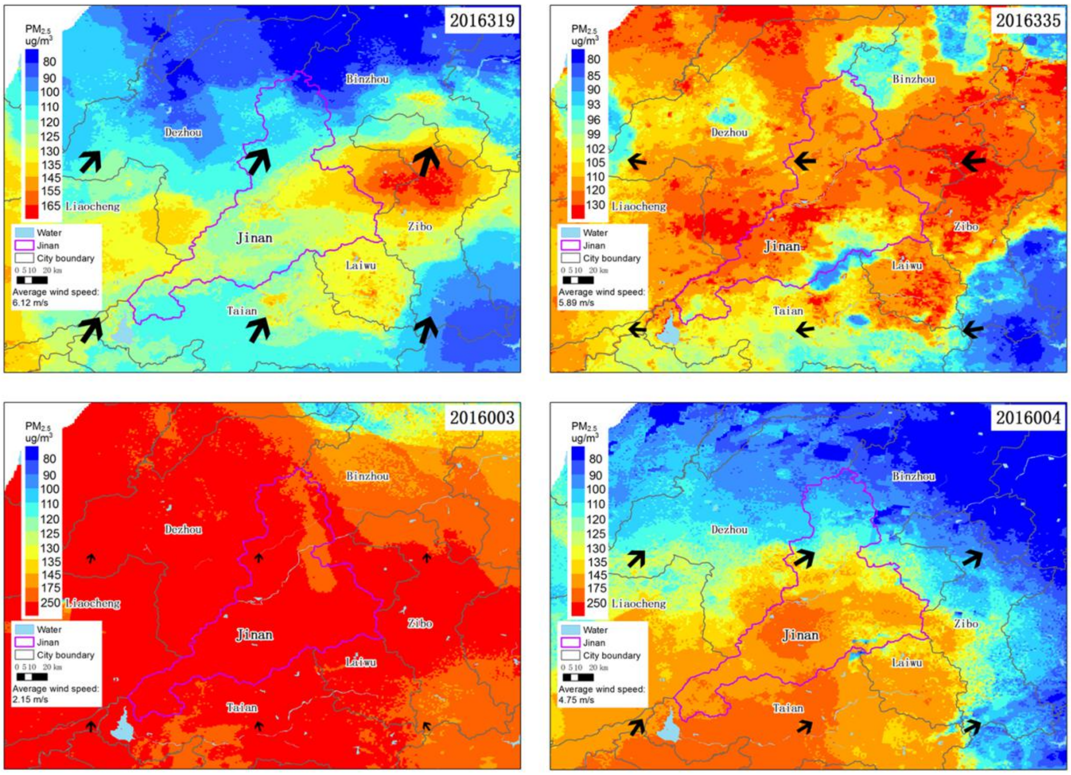


Figure 18 Four cases show the distribution of daily average $PM_{2.5}$ concentration and wind field in Jinan.

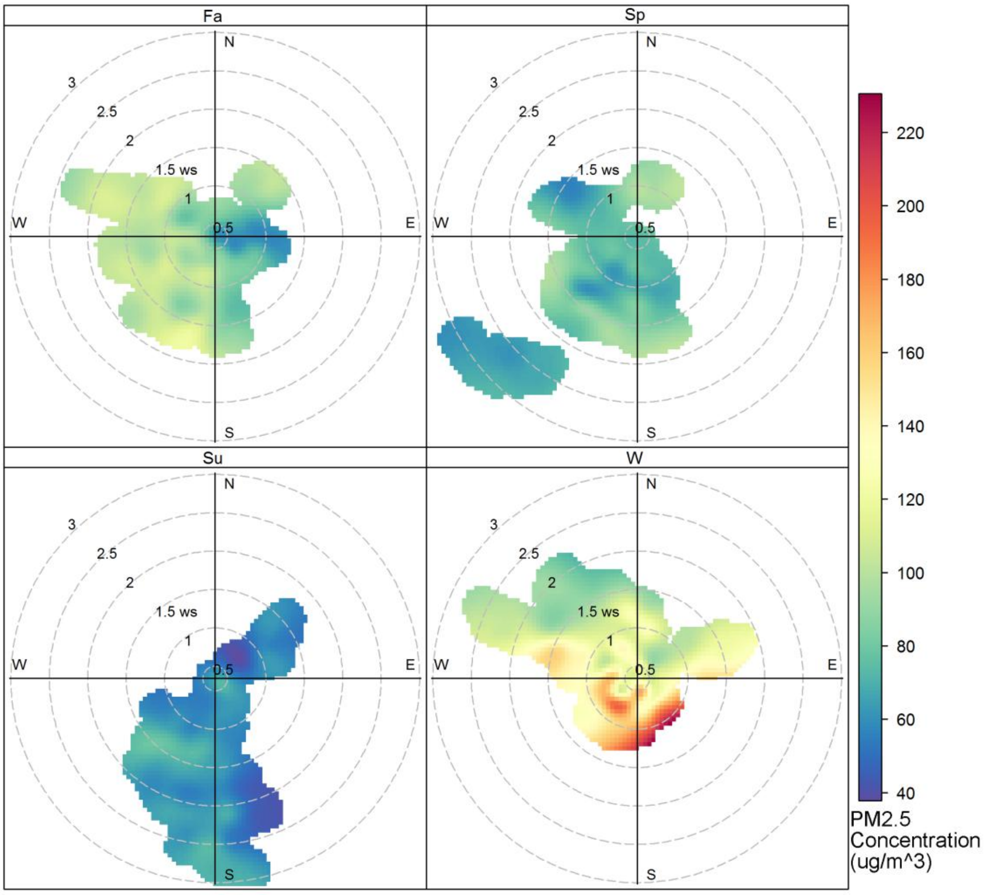


Figure 19 Seasonal analyses by plotting monthly $PM_{2.5}$ concentration in polar coordinates by wind speed and direction in Jinan during 2012-2016.

14.2 Other factors

We compared annual $PM_{2.5}$ concentrations with various factors in the Jinan statistics yearbook, including total electricity consumption, industrial electricity consumption, residential electricity consumption, central heating area, and industrial particulate matter emission. The industrial electricity consumption was positively correlated with annual $PM_{2.5}$ concentration (r=0.68 during 2003-2016) (Figure 20), especially after 2005 (r=0.84 during 2005-2016). The change trend of $PM_{2.5}$ concentrations had a one-year lag with the change trend of industrial electricity consumption. In 2012, the industrial electricity consumption decreased slightly and then kept constant, while the $PM_{2.5}$ concentrations decreased significantly after 2013. This pattern indicated that the decrease in $PM_{2.5}$ concentrations benefited from both the decrease in industrial production as well as better industrial emission process methods.

We compared the temporal trends of $PM_{2.5}$ concentrations in Jinan and in other highly polluted regions in China (Figure 21). We observed that the $PM_{2.5}$ concentrations over Jinan, Jing-Jin-Ji, and Fen-wei Plain showed a significant increase during 2005-2006; however, the

$PM_{2.5}$ concentrations over Yangtze River Delta (YRD) and Sichuan basin did not show this increase. During 2012-2013, the $PM_{2.5}$ concentrations over Jinan, Jing-Jin-Ji and YRD showed a significant increase, while the $PM_{2.5}$ concentrations over Fen-wei plain and Sichuan basin did not show this trend. The different trends of $PM_{2.5}$ concentrations may be associated with regional economic development.

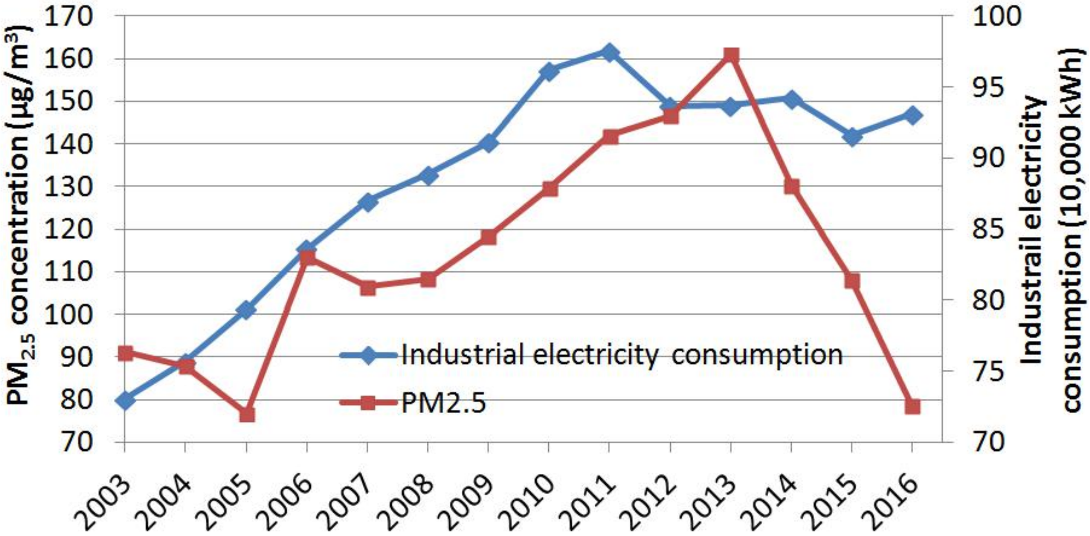


Figure 20 Temporal trends of annual average $PM_{2.5}$ concentrations and industrial electricity consumption.

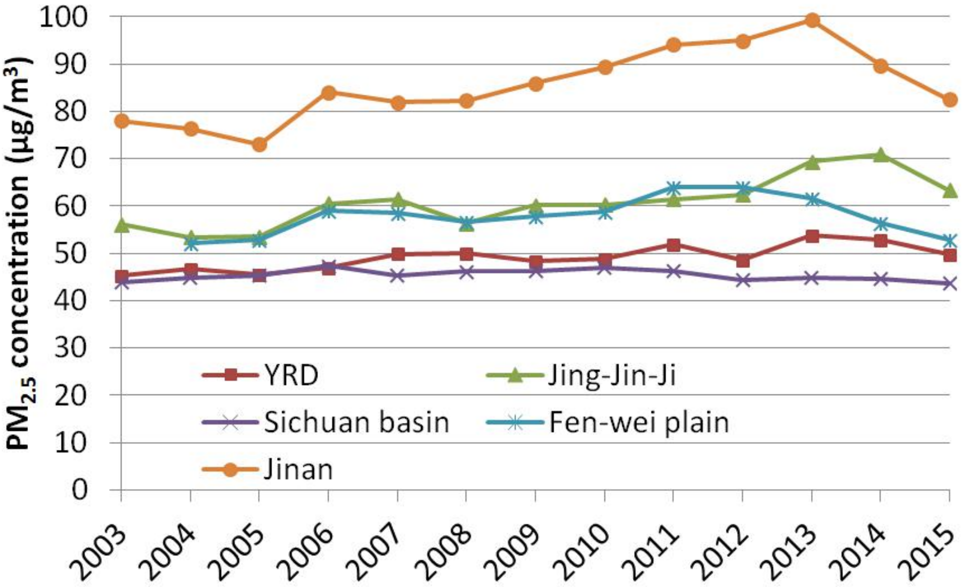


Figure 21 The comparison of historical $PM_{2.5}$ pollution levels in Jinan and other highly polluted regions (YRD, Jing-Jin-Ji, Sichuan basin, and Fen-wei plain).

8. Discussion

8.1 The advantages and limitations of monitoring air quality by satellite remote sensing

In the past decade, satellite remote sensing data have been increasingly used in air pollution monitoring. Compared with ground monitoring stations, satellite data have several advantages. For example, satellite data with high spatial resolution and global/ regional coverage provide continuous pollution surface that benefits the characterization of pollution spatial distribution and detection of local emission sources. Moreover, the long data record and low operating costs of satellite remote sensing could extend the temporal coverage of monitoring and support the analyze of long-term trend of pollution. Satellite data also have some limitations. The accuracy of satellite-based pollutant concentration estimates is lower than measurements from central stations. The ground measurements have a higher temporal frequency than satellite data. Our future research will aim to combine data from polar orbit satellite and geostationary orbit satellite to improve the temporal resolution of satellite retrievals. Thus, we could better characterize haze events as well as monitor pollution transportation.

Satellite remote sensing provides observations of the air quality and it cannot simulate pollution levels under different what-if scenarios. Thus, satellite remote sensing data can hardly contribute to the alarm of high-pollution events or assess the contribution of different factors on haze events. However, employing satellite data together with chemical transport models (CTM) could optimize input data of CTM, thus improve simulation quality.

8.2 Policy advices

We notice that although the current air quality monitoring network covers city centers with high population density, some air pollution hotspots, especially industrial emission sources in rural areas, were not covered by the monitoring net work. This lack of monitoring may lead to the different trends in $PM_{2.5}$ concentrations at city center and at rural hotspots during recent years. Thus, additionally monitoring stations or adjustment of locations of current stations are necessary for monitoring of air pollutant emissions.

Since Jinan located in Shandong Peninsula with flat terrain, we did not observe significant $PM_{2.5}$ transport paths. Additionally, since air pollution level in Jinan was higher than surrounding regions, we observed that the pollution in Jinan is majorly due to local emissions. We noticed that the $PM_{2.5}$ concentrations alone major roads were relatively high, especially in rural regions, indicating that emissions from transportation vehicles is critical for air pollution is Jinan. Further studies on the contribution of transportation emission on air quality is needed.

References:

Van, Donkelaar A, et al. "Global Estimates of Fine Particulate Matter using a Combined Geophysical-Statistical Method with Information from Satellites, Models, and Monitors. " *Environmental Science & Technology*50.7(2016):3762.