

项目编号:20240402

暨南大学大学生创新创业训练计划 项目申报书

项目名称	<u>基于 GAT-TSDFormer 的 PM2.5 时空预测</u> <u>——基于轻量化与解释性视角</u>
项目类型	<u>创新训练项目</u>
项目类别	<u>一般项目</u>
项目负责人	<u>刘威腾</u>
学 号	<u>2021103160</u>
学院专业	<u>经济学院经济统计学</u>
联系电话	<u>15217271378</u>
是否为内外招联队	<u>否</u>
E-mail	<u>liuwt4408@126.com</u>
指导教师	<u>谢贤芬 王国长</u>
申报时间	<u>2024-03-05</u>

暨南大学教务处制

二〇二四年三月

填 写 须 知

一、项目类型分类说明：

1. 创新训练项目是本科生个人或团队，在导师指导下，自主完成创新性实验方法的设计、研究条件准备和项目实施、研究报告撰写、成果（学术）交流等工作。

2. 创业训练项目是本科生团队，在导师指导下，团队中每个学生在项目实施过程中承担一个或多个具体的角色，通过编制商业计划书、开展可行性研究、模拟企业运行、参加企业实践、撰写创业报告等工作。

3. 创业实践项目是学生团队，在学校导师和企业导师共同指导下，采用前期创新训练项目（或创新性实验）的成果，提出一项具有市场前景的创新性产品或者服务，以此为基础开展创业实践活动。

二、项目类别分为一般项目和重点支持领域项目两类。其中，重点支持领域项目旨在鼓励引导大学生根据国家经济社会发展和重大战略需求，结合创新创业教育发展趋势，在重点领域和关键环节取得突出创新创业成果。

三、申报书请按顺序逐项填写，填写内容必须实事求是，表达明确严谨。空缺项要填“无”。

四、填写时可以改变字体大小等，但要确保表格的样式没有被改变；填写完后用 A4 纸张双面打印，不得随意涂改。

项目名称		基于 GAT-TSDFormer 的 PM2.5 时空预测——基于轻量化与解释性视角							
项目类型		创新训练项目		项目类别		一般项目		项目研究周期	1 年
负责人情况	姓名	刘威腾	性别	男	平均绩点	4.12			
	学院	经济学院	专业	经济统计学	生源	<input checked="" type="checkbox"/> 内招 <input type="checkbox"/> 外招			
	联系电话	15217271378	QQ 号	3252303546	Email	liuwt4408@126.com			
项目组成员 (不含负责人)	姓名	性别	学号	学院	专业	内招/外招			
	王国杰	男	2022102173	信息科学技术学院	软件工程	内招			
	招铸峰	男	2022102815	经济学院	经济统计学	内招			
	杨添欣	女	2021102002	经济学院	金融学	内招			
	王萌	女	2021104222	经济学院	金融学	内招			
指导老师情况	姓名	谢贤芬	学科专业	统计学	所在单位	经济学院			
	职称/职务	高级实验师	学历学位	博士	联系电话	15915747152			
	姓名	王国长	学科专业	统计学	所在单位	暨南大学经济学院统计与数据科学系			
	职称/职务	教授、系主任	学历学位	博士	联系电话	13168829697			

一、项目简介

近年来，大气污染逐渐成为全球环境议题的焦点，引发了全球范围内人们对此问题的深刻关切。在中国，随着城市化和工业化的飞速发展，大量未经处理的污染物直接排放到生态环境中，引发了严重的雾霾天气，不仅对社会公众的生命健康构成严重威胁，还限制了我国未来高质量发展的前景。在雾霾治理中，PM2.5 由于其巨大危害和广泛影响，逐渐成为主要的预测、预报和预防对象。习近平总书记在北京考察时明确指出：“要加大大气污染治理力度，应对雾霾污染、改善空气质量的首要任务是控制 PM2.5”。

目前，以 PM2.5 为主要成分的雾霾天气对公众健康安全和社会经济发展构成了巨大威胁，急需精确的模型为社会公众、政府部门提供预测数据。目前主流的 PM2.5 预测模型虽精度较高，但存在着资源高消耗、CO2 高排放的问题，与绿色发展理念相悖，并且现有模型专注于预测准确性，忽略了模型的解释性问题，而目前污染物治理模式正逐渐从预警预报为重点转变为预警预报预防三位一体共同治理，亟需高可解释性的模型在 PM2.5 时间发展和空间传播方面提供科学指导。

在以上多种因素的共同作用下，如何构建高精度、轻量化、高可解释性的 PM2.5 时空预测模型成为中国环境治理、生态文明建设以及建设美丽中国的重要问题。本文针对此重大挑战，通过深化特征工程、综合目前各模型优点以及时间序列分解思想，设计了基于图神经网络（GNN）与注意力机制的 GAT 模型捕获空间依赖性、基于时间序列分解的 TSDFormer 模型捕获时间依赖性，创新性地构建了 GAT-TSDFormer 模型，实现了高精度、轻量化、解释性强的 PM2.5 时空预测。

具体来讲，本项目首先通过划分垂直扩散和水平传播因素的特征工程、基于 GAT 的空间依赖性建模、基于 TSDFormer 的时间依赖性建模建立了 GAT-TSDFormer 模型；

接着通过消融实验、敏感性实验对 GAT-TSDFormer 的性能进行全面分析，并通过基线模型对比实验与目前各个主流 PM2.5 预测模型进行全面对比；

然后通过模型中的 GAT 模块中的注意力机制组件挖掘 PM2.5 空间传播规律，通过 TSDFormer 中的序列分解结果以及各推理算子组件来分析 PM2.5 的时间发展规律，深入揭示了 PM2.5 的时空发展传播机制，为政府区域联防、城市联治提供科学指导，从而助力政府重点区域重点治理，重点时间重点出击；

最终，本文通过构建相关数据库，将模型接入实时数据网站，实现 PM2.5 实时预测，为社会公众、政府部门提供精准的 PM2.5 浓度预测数据。

本项目对 PM2.5 时空预测模型在高精度、轻量化、高可解释性研究方向提出了新的研究思路，填补当前这一方向的研究空缺，科学地丰富我国 PM2.5 时空预测领域的现有研究，为预测精准化、参数轻量化、高度可解释化的 PM2.5 时空预测模型做出进一步贡献，为公众防护、政府预警提供准确的预测支持，为政府预警预报预防三位一体共同治理、区域联防联控联治提供了科学的理论支撑。

二、立项依据（研究意义、国内外研究现状及发展动态分析等）

2.1 研究意义

2.1.1 研究背景

新中国成立以来，我国城市化率由建国初期的 7.3%跃升至 2022 年的 65.22%，且长期来看仍然处于增长的趋势，在这样的背景下，伴随着地方工业化进程的快速推进和全国机动车数量的急速增长，空气污染成了一道不可忽视的棘手难题。党的十八大以来，习近平总书记高度重视大气污染防治工作，在北京，他指出：“大气污染防治是北京发展面临的突出问题，要加大大气污染治理力度，应对雾霾污染。”在全国生态环境保护大会上，总书记要求：“要基本消除重污染天气，还老百姓蓝天白云、繁星闪烁”。国务院于 2024 年 1 月提出的《关于全面推进美丽中国建设的意见》中指出：到 2027 年，绿色发展深入推进，主要污染物排放总量持续减少，生态环境质量持续提升，美丽中国建设成效显著。而目前，我国生态文明建设仍处于压力叠加、负重前行的关键期，各类污染物总量仍居高位，美丽中国建设任务依然艰巨。

在新时代下，需要科学有效的污染物预测模型为美丽中国的建设保驾护航，而其中，PM2.5 污染物由于其危害之大、任务之重，成为目前污染物预测中的主要对象。2020 年 2 月，习近平总书记在北京考察时为大气治理指明了方向：“要加大大气污染治理力度，应对雾霾污染、改善空气质量的首要任务是控制 PM2.5”。

（1）公共健康与环境保护需求迫切

据美国宇航局(NASA)发布的全球细颗粒物(PM2.5)污染分布情况，我国长江中下游、华北地区已属于重度污染地区。根据生态环境部在 2020 年发布的环境质量简报显示，我国城市空气质量达标占比不到 60%，虽然近些年来我国的空气污染物浓度呈下降趋势，但仍然面临大量城市发展因空气污染受限的艰巨挑战。

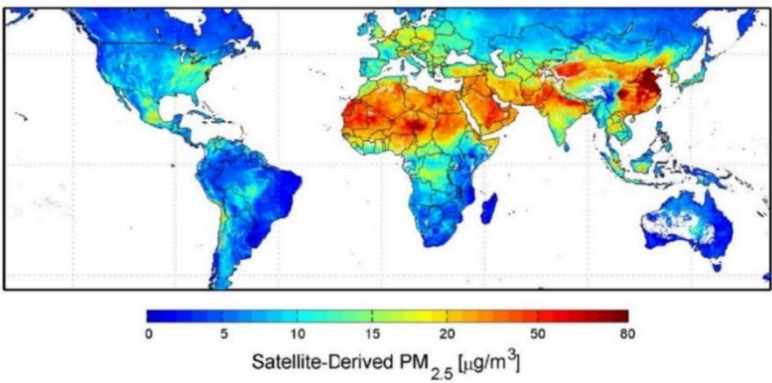


图 1 NASA 发布的 PM2.5 分布状况

在生命健康安全领域，研究表明大气悬浮微粒，特别是 PM2.5 级别的颗粒，能够穿过肺部的气体交换区域并进入血液循环系统，从而触发或加剧心脑血管疾病，并有可能促发恶性肿瘤，对人体威胁巨大；在道路交通领域，以 PM2.5 为主要成分的雾霾天气视野能见度极低，交通事故频繁发生，导致交通阻塞，航班起降滞后，严重影响人们的交通生活；在经济发展领域，雾霾天气不仅会影响农业、养殖业、旅游业等国内经济行业，同时也破坏了我国的国际形象，使得外界开始质疑中国的经济发展模式。

(2) 绿色发展理念下轻量化预测模型为大势所趋

现有的 PM2.5 等主要污染物的预测模型已经取得了一定阶段的胜利，但是这些模型的构建与部署都需要消耗大量自然资源，并产生巨量碳排放。有研究显示，目前主流的一套 PM2.5 浓度预测模型训练以此所产生的碳排放便高达 62 万 6155 磅的二氧化碳排放量，大约相当于五十辆汽车从出厂到报废的总二氧化碳排放规模，而这仅仅是训练过程，还不包含模型实时使用、模型存储、数据库维护等等各个方面。针对以深度学习为代表的大模型领域方面存在的此问题，国务院于 2024 年 1 月提出的《关于全面推进美丽中国建设的意见》中指出：以数据驱动的各种大模型以及派生出的诸如数据库维护等产业，应当加快探索低消耗、低排放的发展范式，探索符合绿色发展理念的新时代大模型。

(3) 新污染物治理体系亟需科学模型指导

由于现有模型对于 PM2.5 的预测的最长时限为 72h，因此对于政府部门而言，只能通过预警预报这一方面来减少 PM2.5 带来的人民生命健康以及社会发展的损失，难以采取科学合理的措施去提前遏制 PM2.5 的发展和传播。2020 年以来，国家环境保护局正式提出：“将污染物治理工作从预警预报为重点转变为预警预报预防三大体系联防联控”，但在实际执行中，由于缺乏科学的时空发展传播理论，常常面对着区域联防抓不到重要城市、城市联治找不对关键地区、政策实施落不到紧要窗口的问题，导致相关政策在所有城市基本相同，没有做好重点区域重点治理，重点时间重点出击，导致治理效率低下，治理效果不佳。

综上所述，以 PM2.5 为主要成分的雾霾天气给公众健康安全、社会经济发展构成巨大威胁；主流预测模型存在的资源高消耗、CO2 高排放问题与绿色发展相违背；新污染物治理体系亟需 PM2.5 时间发展和空间传播科学理论的指导。在以上多种因素的共同作用下，如何构建高精度、轻量化、高可解释性的 PM2.5 时空预测模型是我国环境治理、生态文明建设、建设美丽中国的新时代绿色中国道路上的重大挑战。

2.1.2 研究的理论价值和实践意义

(1) 理论价值

①高精度模型构建

面对传统预测模型构建依赖于气象学和环境科学领域的先验知识而导致的误差精度差，以及近几年现有研究中模型复杂度高、可解释性弱等问题，本研究设计相关组件，创新模型架构，构建对应模型，从而能够通过数据驱动的方式，自动学习和提取相关特征，构建高精度的 PM2.5 时空预测模型，为学界类似于 PM2.5 的相关污染物的时空预测模型提供新的研究思路。

②轻量化模型革新

本研究致力于解决目前 PM2.5 预测模型复杂度高，参数巨大的问题，通过相关模型组件去替代注意力机制和局部全连接，从而开发一种轻量化但不失预测精度的 PM2.5 时空预测模型，减少模型训练和实际使用上对资源和时间的消耗，为学界时空领域的轻量化转变提供新的研究方向和研究思路；

③解释性模型突破

本文所构建的 PM2.5 时空预测模型不仅注重模型的高精度、轻量化，还通过相关模型组件和架构设计深入揭示了 PM2.5 在时间发展和空间传播规律，实现了现有 PM2.5 时空预测模型在解释性上的突破。本文分别在空间依赖性建模和时间依赖性建模上引入相关重要组件，实现了在轻量高精度预测下仍然能够保留模型的高度可解释性，揭示各个相关因素之间在时间和空间上的复杂相互作用，为环境科学理论提供创新的思路。

(2) 实践意义

①精确预警预报，守护公众安全

本研究致力于构建一种精确的 PM2.5 浓度时空预测模型，旨在为政府部门、社会公众提供实时、准确的 PM2.5 浓度数据。对政府部门而言，准确的 PM2.5 浓度预测是科学有效发布空气质量预警预报的基础，只有预测模型精确预测，才能使得社会公众及时获取正确的气象信息，才能保证预警预报有作用、有意义；对社会公众而言，精确的预测使公众能够根据准确信息做出明智的日常生活决策，如调整户外活动计划、选择合适的防护装备，从而有效降低健康风险，提升整体生活质量；

②揭示时空规律，提供政策指引

本研究通过相关模型组件和架构，保留了预测模型的解释性，进而深入挖掘 PM2.5 的时间发展和空间传播规律，从而为城市群在污染物联防联控治理提供科学有效信息，帮助识别重点地区、重点时间窗口，提升政策效率；

首先，模型可揭示 **PM2.5 的时间发展规律**，从而为相关政府部门提供科学决策，实现**重点时间重点作为、重点窗口重点把握**，提升治理效率；

其次，模型可揭示 **PM2.5 的空间传播规律**，帮忙政府部门识别污染高风险区域、优化空气质量监测站点布局、有针对性地颁布相关限排限产政策，实现**重点区域重点治理、重点城市重点出击**。

2.2 国内外研究现状

2.2.1 PM2.5 生成扩散成因复杂，自然因素与人为因素共同作用

PM2.5 生成扩散的形成受多种自然因素的影响，国内学者以一些雾霾（PM2.5 为主要成分）天气高发地区（咸阳市、银川市、京津冀地区等）为例，通过收集气象数据探索影响 PM2.5 形成与维持的重要自然因素。目前研究认为主要有以下五个自然因素共同影响了 PM2.5 生成扩散：

（1）**静小风和逆温现象**：静小风条件下，大气的扩散能力较弱，有利于近地层污染物的积累和雾霾天气的形成，而逆温现象即大气温度随高度增加而升高，会阻碍污染物向上扩散，导致污染物在近地面层积聚（李艳春等，2017）；

（2）**混合层高度**：混合层高度与 PM2.5 浓度呈显著的负相关。持续较低的混合层高度和较弱的垂直上升运动抑制了近地面层大气垂直风向的湍流交换，不利于污染物在垂直方向上的扩散，从而有利于雾霾天气的发展和维持（侯梦玲等，2017）；

（3）**相对湿度**：雾和霾天气的相对湿度在垂直分布上存在显著差异。雾上面的干层使得地面向上的长波净辐射通量较大，有利于辐射雾的形成和发展，而霾天较厚的湿层使得地面向上的长波净辐射通量较小，冷却率也较小，不利于地面降温，阻碍辐射雾的形成（卢一凡等，2021）；

（4）**静稳天气条件**：稳定的高空天气形势与地面弱高压均压场相互配置，为雾霾天气过程提供了有利的环流背景（高萌等，2021）

除了自然环境因素，PM2.5 生成扩散还受到人为因素的影响，目前很多学者已经注意到雾霾污染存在较为显著的**空间溢出效应**。邵帅等（2016）指出，在大气环流等自然因素以及产业转移、交通运输等社会经济因素的作用下，PM2.5 等主要污染物可以在多个区域间传播。邓光耀等（2024）使用社会网络分析对黄河流域雾霾污染空间关联网络的结构特征进行探索，发现黄河流域雾霾污染表现出复杂的**空间关联关系**，一些能源型城市是地区雾霾污染空间关联网络的中心，对城市间雾霾污染空间关联起到关键作用。

2.2.2 PM2.5 时空发展传播机制复杂，时间发展规律与空间传播规律共同作用

现有研究显示，PM2.5 的发展传播在时间与空间上呈现出了一定的特征规律：

在 PM2.5 的时间发展规律上，学者们考虑在 PM2.5 浓度的时间特征时，往往最重视的是季节因素。方莎莎等人（2022）认为武汉市的 PM2.5 浓度存在较为明显的日，月，季差异，总体呈现出冬春高，秋夏低的趋势。Miao Y 等人（2018）对东北、华北平原、华东和四川盆地 4 个人口密集、污染严重的城市区域进行了研究和比较，研究发现 PM2.5 浓度具有明显的季节性差异，其中以冬季最高，秋季和春季次之，夏季最小。Jedruszkiewicz J 等人（2017）却得出冬季 PM2.5 浓度会低于夏季的结果，出现这样的结果可能是由于全球不同地区各个季节的气象因素存在差异而导致的；

在 PM2.5 的空间传播规律上，由于雾霾污染是一种大气污染现象，故会随着空气流动向周围传播，这会加重雾霾污染的空间相关性，同时，不同地区由于地理环境、城市的空间形态、城市执行政策以及产业结构等存在差异，同样会对 PM2.5 浓度产生影响。王静等（2014）对上海市 2013 年 11 月 30 日至 2013 年 12 月 9 日一次严重的雾霾持续性污染过程进行分析，结果表明，此次污染的污染物是由西北气流远距离传送而来，而本地的静稳天气导致污染物不能及时扩散，因此造成了持续十天的雾霾污染。另外，他们提供的卫星监测图也显示雾霾污染在长三角地区也呈现出空间聚集性的特点。秦伟良、许影运用地统计学的方法对江苏省 13 个市雾霾污染进行分析，结果表明，在晴天江苏省的雾霾污染呈现出高度的空间相关性；而在雨雪天气时，江苏省的雾霾污染空间相关性则不是那么明显，天气变化对雾霾的传播也有显著的影响。周侗等人（2022）的对长江流域三大城市群的 PM2.5 浓度研究，显示不同于成渝城市群，长江中游城市群和长江三角洲城市群的 PM2.5 浓度存在明显的空间自相关。吴舒祺等人（2023）认为长三角城市群 PM2.5 浓度的变化呈现出西北高东北低的情况。李梦航等人（2021）从局部角度分析，认为各气象因子在影响 PM2.5 浓度的时候存在空间异质性。

2.2.3 知识驱动预测受阻，数据驱动预测表现优异

传统的大气污染物浓度预测研究主要是基于观测和模拟的方法，CMAQ（Byun，1999）、GEOS Chem（Bey 等，2001）和 WRF/Chem（Grell 等，2005）应用当前对大气物理和化学过程的理解构建微分方程来预测不同污染物的质量浓度，然而，缺乏准确的输入数据（如排放源和排放量）等因素可能会导致这些模型的预测结果不甚理想。

随着环境监测技术的逐渐发展，相关的气象数据和污染物数据逐渐丰富，数据驱动的 PM2.5 预测模型逐渐成为主流，主要可以分为单时间序列预测、多元时间序列预测以及时空预测。

2.2.3.1 单时间序列预测方法

在单时间序列预测中，主要是以 PM2.5 往期数据作为输入，并不引入其它特征，例如 Zhou J 等人（2018）利用 ARIMA 模型对福州市的 PM2.5 浓度进行分析与预测，取得了较好的预测结果。然而这种简单的单变量数据处理方式对时序数据有较多的假设要求，往往很难满足，因此仅适用于某些特殊区域，不具备普适性。

2.2.3.2 多元时间序列预测方法

在多元时间序列预测中，主要以传统线性模型、机器学习模型和深度学习模型为代表。

在**传统线性模型**方面，基于 PM2.5 与相关解释变量之间线性关系的假设，学者们提出了例如地理加权回归（Bruna J.等，2014）、地理和时间加权回归（GTWR）（He and Huang，2018）和土地利用回归（LUR）（Shamsoddini 等，2017）等线性方法。但 PM2.5 浓度预测中各变量之间存在大量的非线性关系，且互相影响，传统线性模型难以胜任该预测任务；

在**机器学习模型**方面，主要模型有支持向量机回归（SVR）、决策树模型（DT）、K 近邻算法（KNN）以及基于集成学习思想的随机森林（RF）、XGBOOST 等。近几年来，康俊锋（2020）等人分别采用 K 临近模型、BP 神经网络模型、支持向量机回归模型、高斯过程回归模型、XGBOOST 模型和随机森林模型对 PM2.5 浓度进行预测。这六种模型进行对比实验发现，针对 PM2.5 浓度的预测问题 XGBOOST 有着更好的表现，并且对冬季的 PM2.5 浓度预测准确考虑更高；刘侠和何春旺（2021）建立了决策树，RF，GBDT 和 XGBOOST 在内的四种机器学习模型用于预测并对北京市 PM2.5 质量等级进行分类，结果显示 XGBOOST 的模型精度要更高一些；祭海洋（2023）建立了 XGBOOST，RF，GBDT 和 ADABOOST 四种机器学习模型，结果显示 GBDT 效果最好；赵滨等（2023）使用岭回归等模型作为元学习器的 Stacking 融合模型与另外四种单一机器学习模型做对比，结果显示前者预测效果优于后者，认为 Stacking 融合模型更适用于大范围的空气质量预测。

在**深度学习模型**方面，主要模型有 MLP、LSTM、GRU 以及其改进模型。黄明等（2012）、Dedovic 等（2015）、Agarwal 等人（2020）各自在其对 PM2.5 等各种污染物的预测中使用了传统的 MLP 模型，后续部分学者基于 MLP 基本架构，针对 PM2.5 预测任务的特点，进行了相关改进，傅明磊等（2015）开发了一个具有滚动机制和累积生成操作的灰色模型的 MLP 网络来预测中国三个城市的 PM2.5 和 PM10 浓度，江继宇等（2016）提出了一种 MLP 模型，利用参与式传感的地理传感器数据来预测中国三个城市的 AQI，Thanavanich 等（2020）引入了一种采用粒群优化的改进 MLP 模型，用于预测中国的 AQI 数据，三者均实现了一定的性能提升；作为 RNN 的变体，LSTM、GRU 模型由于其在时间序列上的有效性能，也被用于空气质量预测模型。王俊山等（2018）在 LSTM 模型基础上引入了三个额外组件来预测北京的 AQI，实现了对 PM2.5 时间发展特征的有效学习。宋希娟等（2019）提出了一种 LSTM-Kalman 模型来预测实验数据集上的三种污染物，结果所提出的模型比经典的 LSTM 模型具有更高的准确性。兰玉霄等（2020）提出了一种考虑时间窗口大小的时间优化方法的 GRU 模型来预测杭州的 AQI。此外张克飞等（2020）、Sonawani 等（2021）、刘波等（2021）提出了增强的 GRU 模型，分别用于预测淮海经济区的 AQI、印度浦那的 NO2 和北京的 PM2.5。

2.2.3.3 时空预测方法

关于空气污染数据序列中复杂的时间依赖性，先前的学者已经开发了各种模型用于不同时间跨度的空气污染预测并取得了不错的效果，但没有考虑监测站之间的空间依赖性，为提升模型预测精度，部分学者集中于解决空间依赖性建模问题来建立时空预测模型，以进行更加准确的预测，目前主要是通过 CNN、图论和 Transformer 变体来捕获空间依赖性。

(1) 运用 CNN 对空间依赖性进行建模存在局限

在地理统计学领域，非线性的空间自相关性构成了关键的考量因素之一，对局部异质性的精细化识别有助于提升预测模型的精确度，有研究通过将监测站点数据人工重构为二维阵列，以适配传统的欧氏空间卷积神经网络 (CNN) 架构，来提取空气质量监测点数据的空间自相关性 (Wu C.等, 2018); 但是这种处理方式破坏了数据的原有空间结构，并未能充分考虑时间序列的相关性; Huang 与 C.J. (2018) 的研究则尝试通过 ConvLSTM 模型解决这一问题，该模型融合了 CNN 与长短期记忆网络 (LSTM) 的框架，以捕获监测站数据的时空动态，但该方法依旧基于 CNN，理论上只适用于欧几里得空间下的空间关系建模。

(2) 图论为非欧式空间依赖性建模提供了全新思路

为了处理非欧式空间数据，F. Scarselli 等 (2008) 提出了 GNN 模型，它可以处理包括无环图、有环图、有向图和无向图在内的多种图类型。近年来，图神经网络 (GNNs) 因其在处理空间信息方面的优异性能，在社会网络分析、生物技术、制药、知识图谱和网络流量预测等领域引起了极大的关注 (Scarselli 等 2009、Wu Z.等 2020、Zhou J.等, 2018); 迄今为止，只有少数研究将 GNN 应用于 PM_{2.5} 空气污染预测，并取得了良好的效果。Lin Y.等 (2018) 建立了一个基于 GNN 的模型来获得站点之间的相关性进行空间依赖性建模; Yanlin Qi 等 (2019) 基于图卷积模型 (GCN) 来提取不同站点之间的空间相关性，并使用 LSTM 来捕捉不同时间观测之间的时间相关性，发现该混合模型相较于其他模型能实现更高的预测精度。除了构建无向图之外，部分学者尝试将部分气象因素添加到图的边缘以此来构建有向图，Z. Wang 等 (2020) 还将气象特征添加到图的边缘，构建了有向图，利用知识增强的图神经网络 GNN-PM_{2.5} 来学习 PM_{2.5} 在城市之间的空间传输，并使用递归神经网络来捕捉时间扩散过程，准确预测未来 72 小时多个城市群的 PM_{2.5} 浓度。尽管如此，上述方法只能预测监测站附近区域的 PM_{2.5} 浓度，而现实中监测站通常分布稀疏，因此对于明显远离监测站的站点，这些方法的预测误差将增大，为此 D. Liang 等人 (2021) 通过建立城市群之间的注意力通道来捕获这种距离较远的城市之间的空间相关性，Z. Liu 等人 (2022) 通过构建多层图卷积层扩大感受野的方法来捕获远距离城市的空间依赖性。L. Su 等 (2023) 在空间注意力机制的模块上，设计了一个名为具有时间注意 GRU 的图卷积网络 (简称 GCNTAG) 的 PM_{2.5} 浓度 48 小时预测模型，通过结合图神经网络 (GNN) 和门控循环单元 (GRUs)，有效地捕捉空气质量时间序列中的时空依赖性特征。

(3) Transformer 提供全新建模思路，改变时空预测领域研究范式

Transformer (A. Vaswani, 2017) 在自然语言处理 (NLP) 和计算机视觉 (CV) 领域取得了巨大成功，由于 Transformer 模型一开始是应用于机器翻译领域，在时空预测领域上不适配，因此部分学者开始对其进行改进，开创性的工作如 LogTrans (S. Li 等,2019) 通过稀疏注意力机制解决了长序列的计算挑战，随后的研究，如 Informer (H. Zhou 等,2021)、Autoformer (H. Wu 等,2021) 和 Fedformer (T. Zhou 等,2022) 引入了创新方法来通过时间特征处理、基于自相关的序列相似性和频域转换来提高预测精度；近两年中，如 PatchTST (Y. Nie 等,2023)、TimesNet (H. Wu 等,2023)、Crossformer (Y. Zhang 等,2023) 和 iTransformer (Y. Liu 等,2023) 结合了 Transformers 的长程依赖能力和 MLP 的局部全连接建模优势，获得了更好的预测性能。

但目前也有学者对这一研究范式进行了反思，Deng J 等 (2024) 利用条件相关性和条件自相关性作为调查工具，揭示了目前基于注意力机制和局部全连接模块在时空预测上的显著冗余（特别是在空间依赖性建模上），提出使用信道保持和时间保持这一研究范式，并以此建立了 HDformer 模型，使得模型更有效地利用数据中的复杂空间和时间规律，保证了预测精度的情况下，模型规模较小。

2.2.4 研究总体现状

综上所述，目前国内外对于 PM2.5 的生成扩散因素较为全面，在 PM2.5 时空预测模型的研究上已经有了较为深刻的见解，取得了不错的预测性能，但本项目认为目前的预测模型在特征工程深化工作、时空预测模型轻量化工作以及模型在 PM2.5 时空发展传播机制的可解释性工作上仍存在空白，故在 PM2.5 时空预测领域还存在以下研究要点和问题：

(1) 结合气象科学领域知识，深化预测模型的特征工程：现有研究大多数直接将 PM2.5 浓度及其影响因素直接输入模型，并没有针对各个影响因素在 PM2.5 在垂直扩散和水平传播上的差异性进行有效区别，没有在特征工程构建上考虑该特点，因此如何结合 PM2.5 在垂直扩散和水平传播上的差异性来构建特征工程，是一个值得研究问题；

(2) 结合各模型优劣势与时间序列分解性思想，构建轻量化的预测模型：目前的 PM2.5 时空预测模型基本通过注意力机制和局部全连接来捕获相关序列的空间、时间相关性，导致模型参数量庞大，复杂度高，因此如何去结合目前各个 PM2.5 的时空预测模型的优劣性，依据时间序列可分解性构建高精度、轻量化的 PM2.5 时空预测模型很值得研究；

(3) 结合模型相关组件，分析 PM2.5 时间发展和空间传播规律：当前的 PM2.5 时空预测模型主要关注预测结果的准确性，却忽略了对于污染物时间发展和空间传播规律的解释，因此如何构建模型相关组件，从数据驱动角度分析 PM2.5 时间发展和空间传播规律，也是很值得研究的一个问题。

主要参考文献:

- [1] Fan,J.,Li,Q.,Hou,J.,Feng,X.,Karimian,H., Lin,S.,2017. A spatiotemporal prediction framework for air pollution based on deep RNN. ISPRS Ann. Photogramm. Remote. Sens. Spat. Inf. Sci. IV-4/W2, 15–22.
- [2] Li, X., Peng, L., Yao, X., Cui, S., Hu, Y., You, C., Chi, T., 2017. Long short-term memory neural network for air pollutant concentration predictions: method development and evaluation. Environ. Pollut. 231, 997 – 1004.
- [3] Wu, C., Li, Q., Hou, J., Karimian, H., Chen, G., 2018. PM2.5 concentration prediction using convolutional neural networks. Sci. Surv. Mapp. 43, 68 – 75.
- [4] Huang, C.J., Kuo, P.H., 2018. A deep CNN-LSTM model for particulate matter (PM2.5) forecasting in smart cities. Sensors 18.
- [5] Bruna J.,Zaremba, W.,Szlam,A., Lecun,Y.,2014. Spectral networks and locally connected networks on graphs. Proceedings of the International Conference on Learning Representations.
- [6] Defferrard, M., Bresson, X., Vandergheynst, P., 2016. Convolutional neural networks on graphs with fast localized spectral filtering. Proceedings of the Advances in Neural Information Processing Systems. pp. 3844–3852.
- [7] Kipf, T.N., Welling, M., 2017. Semi-supervised classification with graph convolutional networks. Proceedings of the International Conference on Learning Representations.
- [8] Qi Y, Li Q, Karimian H, 等 A hybrid model for spatiotemporal forecasting of PM2. 5 based on graph convolutional neural network and long short-term memory[J]. Science of the Total Environment, 2019, 664: 1-10.
- [9] F. Scarselli, M. Gori, A. C. Tsoi, M. Hagenbuchner and G. Monfardini, "The Graph Neural Network Model," in IEEE Transactions on Neural Networks, vol. 20, no. 1, pp. 61-80, Jan. 2009.
- [10] Wang S, Li Y, Zhang J, 等 PM2. 5-GNN: A domain knowledge enhanced graph neural network for PM2. 5 forecasting[C]//Proceedings of the 28th international conference on advances in geographic information systems. 2020: 163-166.
- [11] Scarselli, F., Gori, M., Tsoi, A. C., Hagenbuchner, M., & Monfardini, G. (2009) . The graph neural network model. IEEE Transactions on Neural Networks, 20 (1) , 61 – 80.
- [12] Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., & Yu, P. S. (2020) . A comprehensive survey on graph neural networks. IEEE Transactions on Neural Networks and Learning Systems, 1 – 21.
- [13] Zhou J., Cui G., Zhang Z., Yang, C., Liu, Z., Wang, L., Li, C., & Sun, M. (2018) . Graph neural networks: A review of methods and applications. URL <http://arxiv.org/abs/1812.08434>.
- [14] Lin, Y., Mago, N., Gao, Y., Li, Y., Chiang, Y. Y., Shahabi, C., & Ambite, J. L. (2018) . Exploiting spatiotemporal patterns for accurate air quality forecasting using deep learning. In GIS: Proceedings of the ACM international symposium on advances in geographic information systems (pp. 359 – 368) . New York, NY, USA: Association for Computing Machinery.

- [15] Su I F, Chung Y C, Lee C, 等 Effective PM2. 5 concentration forecasting based on multiple spatial – temporal GNN for areas without monitoring stations[J]. Expert Systems with Applications, 2023, 234: 121074.
- [16] Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, “Attention is all you need,” Advances in neural information processing systems, vol. 30, 2017.
- [17] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, “Bert: Pre-training of deep bidirectional transformers for language understanding,” arXiv preprint arXiv:1810.04805, 2018.
- [18] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, “Transformers in vision: A survey,” ACM computing surveys (CSUR) , vol. 54, no. 10s, pp. 1 – 41, 2022.
- [19] S. Li, X. Jin, Y. Xuan, X. Zhou, W. Chen, Y.-X. Wang, and X. Yan, “Enhancing the locality and breaking the memory bottleneck of transformer on time series forecasting,” Advances in Neural Information Processing Systems, vol. 32, pp. 5243 – 5253, 2019.
- [20] H. Zhou, S. Zhang, J. Peng, S. Zhang, J. Li, H. Xiong, and W. Zhang, “Informer: Beyond efficient transformer for long sequence time-series forecasting,” in Proceedings of AAAI, 2021.
- [21] H. Wu, J. Xu, J. Wang, and M. Long, “Autoformer: Decomposition transformers with auto-correlation for long-term series forecasting,” in Thirty-Fifth Conference on Neural Information Processing Systems, 2021.
- [22] T. Zhou, Z. Ma, Q. Wen, X. Wang, L. Sun, and R. Jin, “Fedformer: Frequency enhanced decomposed transformer for long-term series forecasting,” in International Conference on Machine Learning. PMLR, 2022, pp. 27 268 – 27 286.
- [23] Y. Liu, Q. Liu, J.-W. Zhang, H. Feng, Z. Wang, Z. Zhou, and W. Chen, “Multivariate time-series forecasting with temporal polynomial graph neural networks,” in Advances in Neural Information Processing Systems, A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho, Eds., 2022. [Online]. Available: <https://openreview.net/forum?id=pMumil2EJh>
- [24] R.-G. Cirstea, C. Guo, B. Yang, T. Kieu, X. Dong, and S. Pan, “Triformer: Triangular, variable-specific attentions for long sequence multivariate time series forecasting – full version,” arXiv preprint arXiv:2204.13767, 2022.
- [25] Y. Liu, Q. Liu, J.-W. Zhang, H. Feng, Z. Wang, Z. Zhou, and W. Chen, “Multivariate time-series forecasting with temporal polynomial graph neural networks,” in Advances in Neural Information Processing Systems, A. H. Oh, A. Agarwal, D. Belgrave, and K. Cho, Eds., 2022. [Online]. Available: <https://openreview.net/forum?id=pMumil2EJh> [39] R.-G. Cirstea, C. Guo, B. Yang, T. Kieu, X. Dong, and S. Pan, “Triformer: Triangular, variable-specific attentions for long sequence multivariate time series forecasting – full version,” arXiv preprint arXiv:2204.13767, 2022.
- [26] H. Wu, T. Hu, Y. Liu, H. Zhou, J. Wang, and M. Long, “Timesnet: Temporal 2d-variation modeling for general time series analysis,” in The Eleventh International Conference on Learning Representations, 2023.

- [27] Y. Zhang and J. Yan, “Crossformer: Transformer utilizing crossdimension dependency for multivariate time series forecasting,” in The Eleventh International Conference on Learning Representations, 2023.
- [28] Y. Liu, T. Hu, H. Zhang, H. Wu, S. Wang, L. Ma, and M. Long, “itransformer: Inverted transformers are effective for time series forecasting,” arXiv preprint arXiv:2310.06625, 2023.
- [29] Deng J, Song X, Tsang I W, 等 The Bigger the Better? Rethinking the Effective Model Scale in Long-term Time Series Forecasting[J]. arXiv preprint arXiv:2401.11929, 2024.
- [30] 张可可, 胡冬梅, 闫雨龙, 等. 基于小波变换的山西省 PM_{2.5} 污染特征及影响因素 [J]. 环境科学, 2022, 43 (3) : 1226-1234.
- [31] 周侗, 张帅倩, 闫金伟, 等. 长江经济带三大城市群 PM_{2.5} 时空分布特征及影响因素研究 [J]. 长江流域资源与环境, 2022, 31 (4) : 878-889.
- [32] 吴舒祺, 么嘉祺, 杨冉, 等. 长三角城市群 PM_{2.5} 时空变化和影响因素分析 [J/OL]. 环境科学, 2023, 1-14[2023-03-25] .
- [33] 李梦航, 阙翔, 刘金福, 等. 中国城市 PM_{2.5} 影响因素的空间异质性分析 [J]. 赤峰学院学报 (自然科学版), 2021, 37 (1) : 37-41.
- [34] 方莎莎, 廖可文, 刘火胜, 等. 武汉 2018~2019 年 PM_{2.5} 浓度变化特征及气象因素影响分析 [J]. 绿色科技, 2022, 24 (24) : 211-216.
- [35] 戴春雨. 哈尔滨城市街区 PM_{2.5} 浓度预测及街区形态优化策略研究[D].东北林业大学,2023.
- [36] 张宏波. 机器学习在 PM_{2.5} 浓度预测中的应用研究[D].西南大学,2023.
- [37] 王静, 施润和, 李龙, 等. 上海市一次重雾霾过程的天气特征及成因分析[J]. 环境科学学报, 2014, 11: 25-32.
- [38] 秦伟良, 许影. 基于地统计学的江苏 PM_{2.5}空间变异性分析[J]. 环境工程, 2014, 01: 73-76+91.
- [39] Guo Y A, Zhang X M. Analysis of the pollution characteristics and influence factors of PM_{2.5} in Chinese main capital city [C]. Advanced Materials Research. Trans Tech Publications Ltd, 2014, 1023: 247-251.
- [40] Jędruszkiewicz J, Czernecki B, Marosz M. The variability of PM₁₀ and PM_{2.5} concentrations in selected Polish agglomerations: the role of meteorological conditions, 2006 – 2016 [J]. International Journal of Environmental Health Research, 2017, 27 (6) : 441-462.
- [41] Miao Y, Liu S, Guo J, 等 Unraveling the relationships between boundary layer height and PM_{2.5} pollution in China based on four-year radiosonde measurements[J]. Environmental pollution, 2018, 243: 1186-1195.
- [42] Silcox G D, Kelly K E, Crosman E T, 等 Wintertime PM_{2.5} concentrations during persistent, multi-day cold-air pools in a mountain valley[J]. Atmospheric Environment, 2012, 46 (1) : 17-24.
- [43] Shi Y, Xie X, Fung C, 等 Identifying critical building morphological design factors of street-level air pollution dispersion in high-density built environment using mobile monitoring[J]. Building and Environment, 2018, 128: 248-259.

三、项目方案（具体研究内容、研究目标、计划、技术路线、人员分工等）

3.1 研究思路

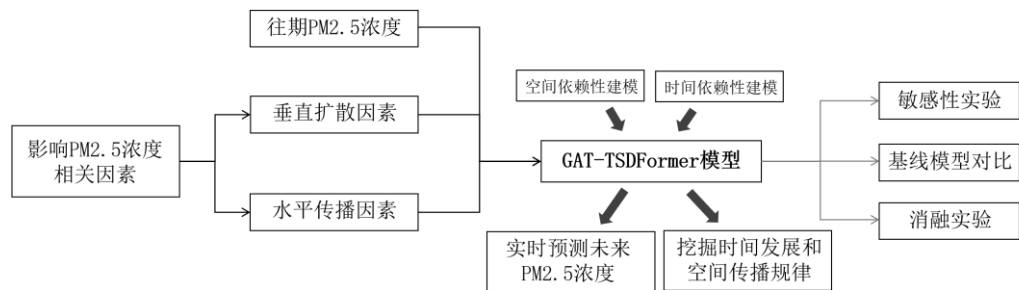


图 2 基本研究思路

本项目首先深化了特征工程：将影响 PM2.5 的相关因素划分为垂直扩散和水平传播因素；接着建立了 GAT-TSDFormer 模型：基于 GAT 模型进行空间依赖性建模、基于 TSDFormer 模型进行时间依赖性建模；然后进行消融实验、敏感性实验、基线模型对比实验，对 GAT-TSDFormer 性能进行全面的分析和对比；接着深入揭示了 PM2.5 的时空发展传播规律：通过模型中的 GAT 模块中的注意力机制组件挖掘 PM2.5 空间传播规律，通过 TSDFormer 中的序列成分分解以及各推理算子组件来分析 PM2.5 的时间发展规律；最终实现 PM2.5 实时预测：通过构建相关数据库，将模型接入实时数据网站，为社会公众、政府部门提供精准的 PM2.5 浓度预测数据。

3.2 研究内容

3.2.1 构建特征工程，深入挖掘 PM2.5 时空发展传播因素

本项目的主要目标是准确预测每小时 PM2.5 的浓度，从前面文献综述可知，存在大量的影响 PM2.5 发展传播的因素，为了提升我们后续模型的预测精度，我们充分挖掘已有的因素信息，深化特征工程，将关于大气环境和污染物的传播信息纳入到输入变量之中。

在本项目中，我们将影响 PM2.5 发展传播的因素分为两类：垂直扩散因素和水平传播因素，其中垂直扩散因素是指 PM2.5 通过复杂的动态或热力学过程在垂直方向上的扩散过程，主要为某城市当地的气象特征状态,在后续模型中我们将它们作为节点属性.通过查阅相关文献,我们整理了 PM2.5 垂直扩散因素如下表所示：

表 1 PM2.5 垂直扩散因素表

相关因素	影响机制
PBL 高度	影响 PM2.5 垂直稀释
K 指数	湿空气分层不稳定（K 指数）影响 PM2.5 聚集
950ha 风速	950ha 处风速直接影响 PM2.5 垂直扩散
2m 温度	通过化学效应和冷锋通风影响 PM2.5 浓度
相对湿度	空气水分会促进硝酸铵（PM2.5 主要成分）的形成

降水	通过拖曳效应产生下沉气流，抑制 PM2.5 垂直扩散
表面压强	较高表面压强意味着稳定大气条件，易导致 PM2.5 累计形成
.....
<p>水平扩散因素是指 PM2.5 通过借助风的动能在水平方向上，从一个城市传播到另外一个城市的传播过程，主要为城市与城市之间的相对位置、风向风速等气象特征状态，在后续模型中我们将它们作为边缘属性。通过查阅相关文献，我们整理了 PM2.5 水平传播因素如下表所示(假设此时研究的是从城市<i>i</i>向城市<i>j</i>)的传播过程：</p>	
表 2 PM2.5 水平扩散因素表	
相关因素	影响机制
城市 <i>i</i> 地面风速	地面风速越大，越容易传播到其它城市
城市 <i>i, j</i> 直线距离	距离越短，越容易传播到其它城市
城市 <i>i</i> 风向与城市 <i>i</i> 至城市 <i>j</i> 方向差值的绝对值	其差值的绝对值越大，借助风的动能传播到 <i>j</i> 城市可能性越大
平流系数 <i>S</i>	公式(1)(2)
.....
<p>平流系数计算公式如下：</p> $S_{j \rightarrow i} = ReLU\left(\frac{ v }{d} \cos(\alpha)\right) \tag{1}$ $\alpha = \gamma - \beta \tag{2}$	
<p>其中：这里 <i>v</i> 是城市<i>i</i>的风速，<i>d</i>是城市<i>i</i>和城市<i>j</i>之间的距离，<i>γ</i>是城市<i>i</i>的风向，<i>β</i>是城市<i>i</i>到城市<i>j</i>的方向，其中ReLU(·)用于捕获从城市<i>i</i>到城市<i>j</i>的单向传输，<i>S</i>_{<i>i</i>→<i>j</i>}方向与气象风向相同。</p>	
<p>在后续项目开展中，我们将查阅更多文献，挖掘影响 PM2.5 垂直扩散和水平传播的相关因素，构建更加全面完善的特征工程。</p>	

3.2.2 构建预测模型，精确预测 PM2.5 每小时浓度

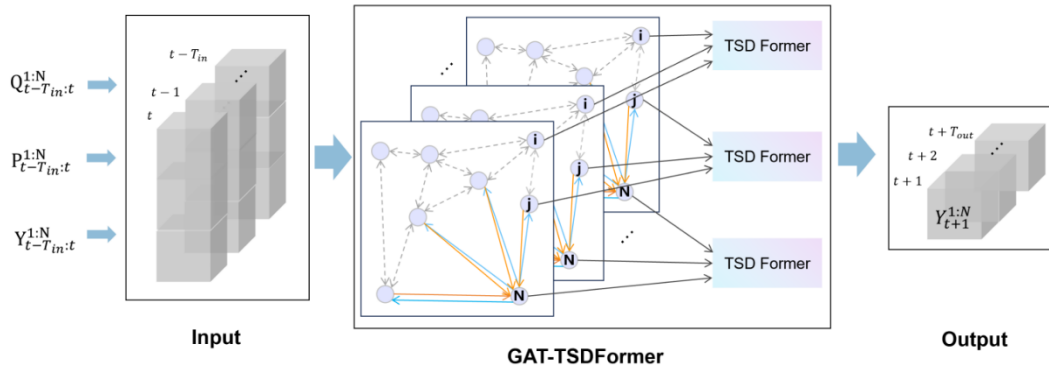


图 3 GAT-TSDFormer 模型

模型主要任务为拟合最优非线性函数 $f(\cdot)$,使得：

$$Y_{t+1:t+T_{out}}^{1:N} = f(Q_{t-T_{in}:t}^{1:N}, P_{t-T_{in}:t}^{1:N}, Y_{t-T_{in}:t}^{1:N}) \quad (3)$$

其中 $Q_{t-T_{in}:t}^{1:N}$, $P_{t-T_{in}:t}^{1:N}$, $Y_{t-T_{in}:t}^{1:N}$ 分别表示模型输入从 $t - T_{in}$ 到 t 时刻输入的 N 个城市节点属性、边缘属性以及 PM2.5 浓度, $Y_{t+1:t+T_{out}}^{1:N}$ 表示 $t + 1$ 到 $t + T_{out}$ 时刻模型预测输出的 PM2.5 浓度。由图可以看出, 模型首先将相关变量和 PM2.5 浓度的往期数据作为输入, 通过 GAT 模块进行空间依赖性建模, 然后再通过 TSDFormer 进行时间依赖性建模, 最终输出得到 PM2.5 未来浓度数据, 各部分的详细结构如下：

3.2.2.1 图构建

首先考虑到 PM2.5 浓度受到多种因素影响, 且不同因素对 PM2.5 扩散传播影响机制存在差异, 我们依据 3.2.1 部分中对垂直扩散因素和水平传播因素的分类, 定义有向图的节点属性和边缘属性, 并根据城市距离和是否有高山阻挡来定义邻接矩阵。

首先定义节点属性 P^i ：我们将城市 i 的节点属性 P^i 定义为 PM2.5 垂直扩散因素与城市的经纬度, 具体如下表所示：

表 3 节点属性

变量名	单位
PBL 高度	m
K 指数	K
纬线方向的风速	m/s
经线方向的风速	m/s
2m 处的温度	K
相对湿度	%
总降水量	m
表面压力	Pa
经度	-
维度	-
.....

接着定义边缘属性 $Q^{j \rightarrow i}$ ：我们将城市 j 到城市 i 的边缘属性 $Q^{j \rightarrow i}$ 定义为 PM2.5 水平传播因素，其方向与此时气象风向相同(假设此时方向为从城市 j 到城市 i)，具体如下表所示：

表 4 边缘属性

变量名	单位
城市 i 的风速	km/h
城市 i, j 直线距离	km
城市 i 的风向	-
从城市 i 到城市 j 的方向	-
平流系数 S	%
.....

最后定义邻接矩阵 A ：我们根据以下两个因素来构建邻接矩阵：首先，大多数气溶胶污染物分布在地面以上的一定范围内，因此距离是衡量节点相关性的一个重要因素；其次两个城市沿线的山脉会阻碍污染物的 PM2.5 的水平传播，若山体过高，则会直接阻碍其水平传播，故邻接矩阵 A 定义如下：

$$A_{ij} = H(d_{\theta} - d_{ij}) \cdot H(m_{\theta} - m_{ij}) \quad (4)$$

$$d_{ij} = \|\rho_i - \rho_j\| \quad (5)$$

$$m_{ij} = \sup_{\lambda \in (0,1)} \{h(\lambda \rho_i + (1 - \lambda) \rho_j) - \max\{h(\rho_i), h(\rho_j)\}\} \quad (6)$$

其中： ρ_i 是节点 i 的位置(纬度，经度)， $h(\rho_i)$ 是位置 ρ 的高度， $\|\cdot\|$ 是向量的 L2 范数， $H(\cdot)$ 为阶跃函数。 d_{θ} 和 m_{θ} 分别是距离和高度的阈值。

在本文中，我们距离阈值设置为 $d_{\theta} = 300 \text{ km}$ ，为高度阈值设置 $m_{\theta} = 1200 \text{ m}$ 。这意味着：只有当这两个城市之间的距离小 300km，并且它们之间的山脉高度小于 1200m 时，PM2.5 才能从一个城市传输到另一个城市。

3.2.2.2 空间依赖性建模

在空间依赖性的建模中，我们结合知识增强型图神经网络（GNN）和注意力机制，通过利用邻近信息和更新节点的表示来捕获 PM2.5 的空间传播，从而让模型学习 PM2.5 在空间上即不同城市之间的传播效应。此外考虑到不同节点之间在 PM2.5 传播强度不同，我们引入具有方向的权重参数 $\alpha_{j \rightarrow i}$ 和 $\alpha_{i \rightarrow j}$ ，让模型学习城市之间的传播强度差异，具体结构如下图所示：

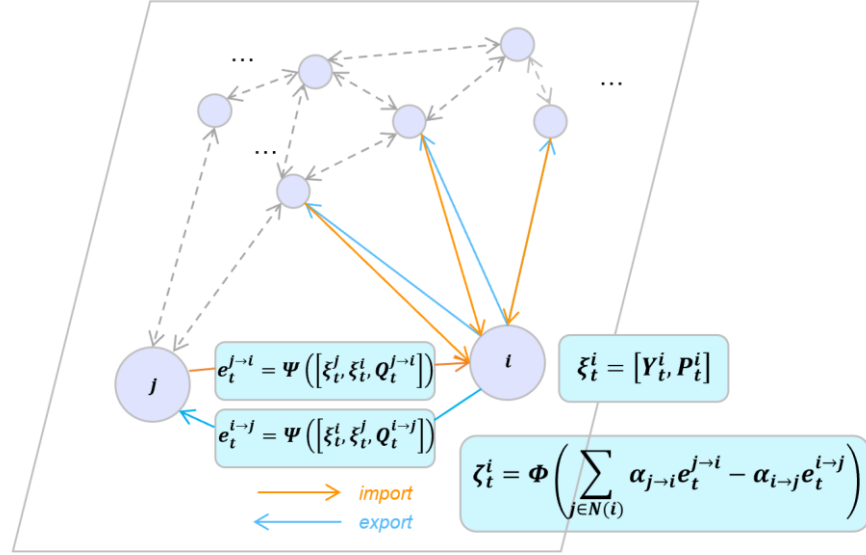


图 4 空间依赖性建模

空间依赖性建模的具体迭代过程的公式如下所示：

$$\xi_t^i = [Y_t^i, P_t^i] \quad (7)$$

$$e_t^{j \rightarrow i} = \Psi([\xi_t^j, \xi_t^i, Q_t^{j \rightarrow i}]) \quad (8)$$

$$\zeta_t^i = \Phi\left(\sum_{j \in N(i)} (\alpha_{j \rightarrow i} e_t^{j \rightarrow i} - \alpha_{i \rightarrow j} e_t^{i \rightarrow j})\right) \quad (9)$$

其中， ξ_t^i 为节点表示； $e_t^{j \rightarrow i}$ 为边缘表示，与边缘属性 $Q_t^{j \rightarrow i}$ 同方向； ζ_t^i 为节点的空间相关性，通过汇总所有邻居的影响来计算，具体而言，给定一个节点 i ，来自其邻居节点集合 $\{j: j \in N(i)\}$ 的传输等于权重参数与边缘表示 $\alpha_{j \rightarrow i} e_t^{j \rightarrow i}$ ， $\alpha_{i \rightarrow j} e_t^{i \rightarrow j}$ 之间的差异，如图 4 中的橙色线和蓝色线所示。

经过多次递归操作后，图上的每个节点都可获取其他节点的信息实现空间依赖性建模，最终将 ξ_t^i, ζ_t^i 进行连接操作后得到城市 i 在空间依赖性模块的输出：

$$H_t^i = [\xi_t^i, \zeta_t^i] \quad (10)$$

$$H_{t-T_{in}:t}^i = \{H_{t-T_{in}}^i, H_{t-T_{in}+1}^i, \dots, H_t^i\} \quad (11)$$

3.2.2.3 时间依赖性建模

在时间依赖性建模方面，我们通过推理组件和外推组件来有效地对输入的时间序列有效分解为趋势成分、季节成分和短期波动成分并对各个部分进行有效外推。图 5 顶部分支（红虚线上部）说明了用于导出各序列成分的推理过程，而底部分支（红虚线下部）则描述了外推过程，用于预测各成分的潜在演变。最终获得的各成分的分量和残差被组合成一个宽向量，最后输入到多项式回归层以捕获它们复杂的相互关系。

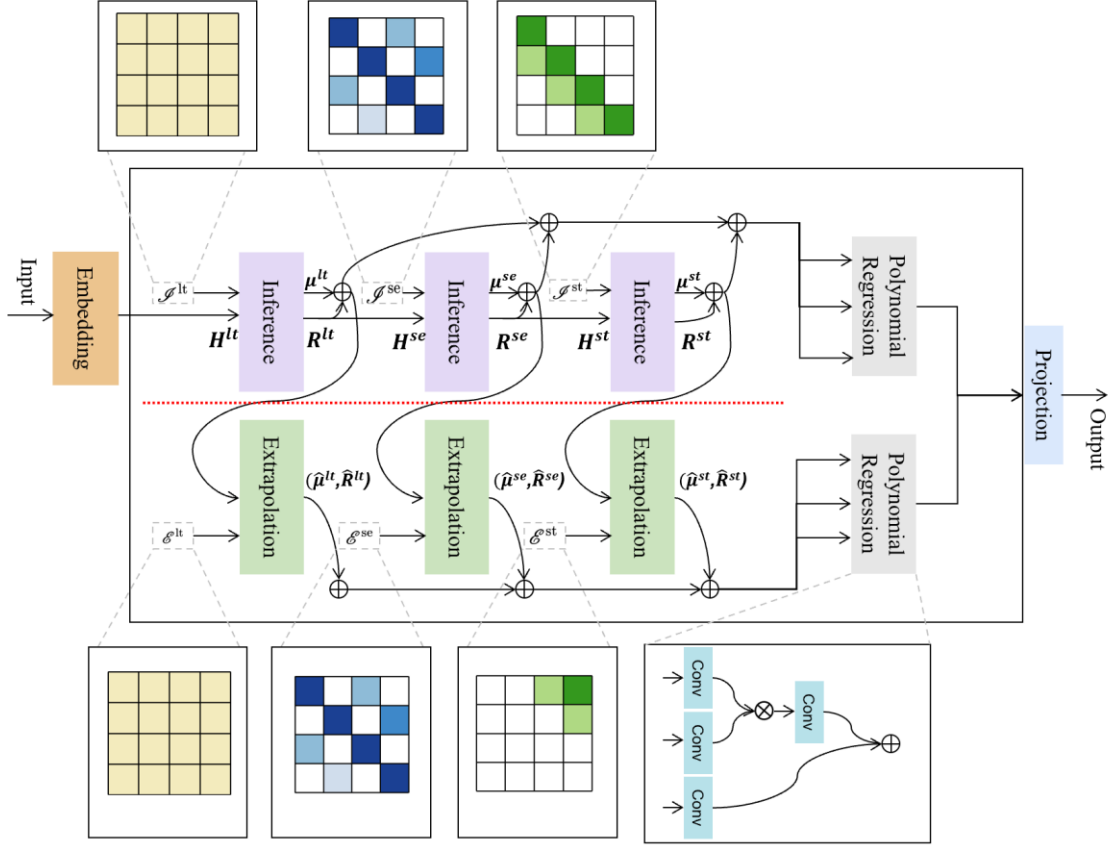


图 5 时间依赖性建模

在推理组件方面，我们通过推理算子 $J^* \in R^{T_{in} \times T_{in}}$ 沿着时间维度分别应用于每个序列，用于推断相应的时间成分，在整个系列中共享：

$$\mu_i^* = J^* H_i^* \quad (12)$$

$$\sigma_i^{*2} = J^* H_i^* - \mu_i^{*2} + \epsilon \quad (13)$$

其中 $* \in \{lt, se, st\}$ 表示趋势、季节、短期变动； $H_i^* \in R^{T_{in} \times d}$ 表示第 i 个序列的输入特征，具体为空间依赖性模型的输出或上个推理模块的残差输出； $\mu_i^*, \sigma_i^{*2} \in R^{N \times T_{in} \times d}$ 为推理第 i 个序列对应成分的均值和标准差；

接着，通过从数据中减去平移效应并将差值除以缩放效应将其余成分从数据中分离出来，得到对应的推理组件的残差 $R_i^* \in R^{N \times T_{in} \times d}$ ：

$$R_i^* = \frac{H_i^* - \mu_i^*}{\sigma_i^*} \quad (14)$$

在外推组件方面，我们通过外推算子 \mathcal{E}^* 对 μ_i^* 与 R_i^* 进行线性映射进行外推：

$$\hat{\mu}_i^* = \mathcal{E}^* \mu_i^* \quad (15)$$

$$\hat{R}_i^* = \mathcal{E}^* R_i^* \quad (16)$$

其中 $\hat{\mu}_i^* \in R^{N \times T_{in} \times d}$ 为外推第 i 个序列对应成分的均值和标准差； \hat{R}_i^* 为对应的外推组件的残差

在时间序列的趋势、季节和短期波动成分的推理和外推中，推理组件和外推组件的主要差别在于推理算子 \mathcal{J}^* 和外推算子 \mathcal{E}^* 的不同，其具体定义如下。

(1) 趋势成分推理和外推算子

趋势成分旨在表征时间序列数据的趋势模式，通过汇总多个季节收集的样本，消除仅影响少数时间步长的短期影响，并获得偏差较小的长期成分的估计。长期分量的推理和外推算子公式定义如下：

$$\mathcal{J}_{t_1, t_2}^{lt} = \frac{1}{T_{in}} \quad (17)$$

$$\mathcal{E}_{t_1, t_2}^{lt} = \frac{1}{T_{in}} \quad (18)$$

(2) 季节成分推理和外推算子

在推理算子方面，在季节持续时间一致的假设下，定义 τ_{in} 为输入序列所包含的季节数， c 为单个季节的长度，则有 $T_{in} = \tau_{in} * c$ 。与趋势成分推理和算子相反，季节性成分采用扩大窗口策略，该窗口由 τ_{in} 大小和 c 膨胀因子定义，捕获跨越多个季节的数据点以反映数据的周期性特征。

此外，我们利用注意力机制，根据输入序列中的顺序为不同季节分配不同的权重，从而可以准确地衡量它们与季节成分的相关性。权重定义为参数矩阵 $W^{se} \in R^{\tau_{in} \times \tau_{in}}$ ，接着通过进行归一化得到 $A^{se} \in R^{\tau_{in} \times \tau_{in}}$ 。其中 A^{se} 代表季节相关性，用于评估季节之间存在的内在相互联系，季节成分推理算子定义如下：

$$\mathcal{J}_{t_1, t_2}^{se}(A) = \begin{cases} A_{i,j}^{se} & t_1 - t_2 \equiv 0 \pmod{c} \\ 0 & \text{其它} \end{cases} \quad (19)$$

其中：

$$i = \left\lfloor \frac{t_1}{c} \right\rfloor, j = \left\lfloor \frac{t_2}{c} \right\rfloor \quad (20)$$

$$A_{i,j}^{se} = \frac{\exp(W_{i,j}^{se})}{\sum_{k=0}^{\tau_{in}-1} \exp(W_{i,k}^{se})} \quad (21)$$

在外推算子方面，若推理算子所得的季节性成分稳定，则可直接重复进行直接外推，但这一情况并不多见，因此为适应不断变化的季节性成分，我们构建了动态的外推算子。首先我们定义输出序列的最小季节数表示为 τ_{out} ，定义 $\hat{W}^{se} \in R^{\tau_{out} \times \tau_{in}}$ 为参数矩阵来捕获输入和输出序列中每对季节之间的季节相关性，与推理算子类似，我们对 \hat{W}^{se} 进行归一化，得到 $\hat{A}^{se} \in R^{\tau_{out} \times \tau_{in}}$ ，故**季节成分外推算子**定义如下：

$$\mathcal{E}_{t_1, t_2}^{se}(A) = \begin{cases} \hat{A}_{i,j} & t_1 - t_2 \equiv 0 \pmod{c} \\ 0 & \text{其它} \end{cases} \quad (22)$$

其中：

$$i = \left\lfloor \frac{t_1}{c} \right\rfloor, j = \left\lfloor \frac{t_2}{c} \right\rfloor \quad (23)$$

$$\hat{A}_{i,j}^{se} = \frac{\exp(\hat{W}_{i,j}^{se})}{\sum_{k=0}^{\tau_{in}-1} \exp(\hat{W}_{i,k}^{se})} \quad (24)$$

(3) 短期波动成分推理和外推算子

短期成分识别季节性和长期成分未解释的不规则现象和短暂现象，与季节成分推理外推相比，它需要一个较小的窗口大小 δ ，以包含具有直接相关性的最近观察结果。这些观察结果受当前短期因素的影响，根据相关滞后表现出不同程度的相关性。为捕获不同滞后阶数的相关性，我们定义参数向量 $w^{st} \in R^\delta$ 来构建注意力机制来利用了注意力加权的观察结果，最终**短期波动成分外推算子**定义如下：

$$\mathcal{J}_{t_1, t_2}^{st}(A) = \begin{cases} A_{t_1-t_2}^{st} & (t_1 - t_2 \geq 0) \wedge (t_1 - t_2 < \delta) \\ 0 & \text{其它} \end{cases} \quad (25)$$

$$A_i^{st} = \frac{\exp(w_i^{st})}{\sum_{i=0}^{\delta-1} \exp(w_i^{st})} \quad (26)$$

与季节成分的外推算子类似，我们首先定义 $\hat{w}_{st} \in R^{\delta \times \delta}$ 为参数矩阵来捕获窗口大小中各个时刻的相关性，并对 \hat{w}_{st} 进行归一化，故**短期波动成分外推算子**定义如下：

$$\mathcal{E}_{t_1, t_2}^{st}(A) = \begin{cases} \hat{A}_{t_1, T_{in}-1-t_2}^{st} & (t_1 < \delta) \wedge (t_2 > T_{in} - \delta - 1) \\ 0 & \text{其它} \end{cases} \quad (27)$$

$$\hat{A}_{i,j}^{st} = \frac{\exp(\hat{w}_{i,j}^{st})}{\sum_{k=0}^{\delta-1} \exp(\hat{w}_{i,k}^{st})} \quad (28)$$

3.2.3 编写代码，进行相关实验

3.2.3.1 数据来源

对于数据集的查找，我们首先使用了 PM2.5-GNN 模型论文里面构建的一个完整的 **4 年数据集 KnowAir**，该数据集共覆盖 184 个城市（节点），由于这些城市主要是在**京津冀地区**，所以我们将使用该数据集 KnowAir 用于我们完成模型代码后对京津冀地区的 PM2.5 浓度预测的研究，并通过该数据集进行进一步完善我们的代码。在模型改进和主体代码调试工作结束后，我们还将继续寻找合适的全国范围内的数据集来进行全国 PM2.5 浓度的预测。

3.2.3.2 消融实验

消融实验是指通过系统地移除模型的某一部分来评估该部分对模型整体性能的影响，从而来验证模型某组件的合理性。在本文中，我们的主要组件分别为空间依赖性建模中的注意力机制组件和时间依赖性建模中的推理外推组件，为此本文以 GAT-TSDFormer 模型作为完整模型，分别以单独移除空间依赖性建模中的注意力组件、单独移除时间依赖性建模中的推理外推组件以及两者皆移除作为移除组件后模型，通过对比来了解被移除组件的作用来验证我们的观点，并根据结果对模型做进一步的调整和实验。消融实验的实验步骤如下：

（1）运行完整的 GAT-TSDFormer 模型代码：

（2）运行移除空间依赖性建模中的注意力组件后的模型代码： 通过从 GAT-TSDFormer 模型中移除了空间依赖性建模的注意力机制，保留了时间依赖性建模的时间序列推理和外推组件，可以评估注意力机制在模型中的作用以及其对整体性能的影响。

（3）运行移除时间依赖性建模中的推理外推组件后的模型代码： 在这个实验中，我们从 GAT-TSDFormer 模型中移除了时间依赖性建模中的推理外推组件，保留了空间依赖性建模的注意力机制，这样便可以评估时间序列推理和外推组件在模型中的作用以及其对整体性能的影响。

（4）运行两者皆移除后的模型代码： 在这个实验中，我们从 GAT-TSDFormer 模型中同时移除了空间依赖性建模的注意力组件，和时间依赖性建模中的推理外推组件来评估两个组件的整体作用以及它们对整体性能的综合影响。

通过对比这些移除组件后的模型和完整的 GAT-TSDFormer 模型，我们可以了解被移除组件的作用，并验证我们的观点：移除某个组件后模型性能显著下降，那么说明该组件对整体性能的贡献很大；此外在消融实验的基础上，我们还可以改进某个组件以进一步调整和优化模型，达到更好的预测效果。

3.2.3.3 敏感性实验

(1) 超参数敏感性实验

对于 GAT-TSDFormer 以及各个基线模型的超参数如学习率、优化器、窗口大小，我们将通过调整其数值，选择能够使对应模型达到最佳预测性能的超参数组合，保证基线模型对比的合理性和科学性；此外，我们也将通过系统地调整 GAT-TSDFormer 模型超参数，来科学观察分析各个超参数对模型性能的影响，并以此评估模型的稳定性。

(2) 序列长度敏感性实验

在对 PM2.5 进行时空预测时，GAT-TSDFormer 模型以及对应的基线模型可能对于不同的输入序列长度，呈现出不同的预测性能，因此有必要对于输入序列长度进行敏感性实验，对比不同的输入序列长度下各个模型的预测精度，更加全面地比较分析各个模型地预测性能。

3.2.3.4 基线模型对比实验

为更加准确地衡量模型的预测精度，我们将 GAT-TSDFormer 模型与以下三类基准模型进行比较：

(1) 历史平均值 (HA)：通过相应时期的历史读数的平均值来预测空气质量。它不需要复杂的算法和大量的计算资源，只依赖于历史平均值。但由于没有考虑其他因素的影响，对于复杂的时空关系和变化模式预测效果有限。

(2) 多层感知机 (MLP)：它通过多个全连接层组成，来构建无时序性的模型来预测空气质量。MLP 的基本结构包括输入层、隐藏层和输出层。输入层接收原始数据或特征向量作为输入。它在处理非线性关系时具有较强的表达能力，但对输入数据的尺度和分布比较敏感，需要进行适当的预处理。

(3) STGNN 的变体模型：STGNN 及其变体是目前图论在 PM2.5 时空预测的重要模型，我们选择 MTGNN (Wu et al 2020) 和 GAGNN (Chen et al. 2021) 作为基线模型。这些模型利用图神经网络的能力来捕捉空气质量数据中的时空依赖关系，通过建模节点之间的关系和时间上的演化，实现对未来空气质量的预测，其优点在于能够处理复杂的时空数据，具有较高的预测准确性，但缺点在于计算复杂度高，以及对稀疏矩阵敏感。

(4) Transformer 的变体模型：目前基于 Transformer 变体模型在时空预测上获得了很大成功，我们选取 iTransformer (Q. Liu. et al.2022) 和 Airformer (D. Liang et al. 2023) 作为基线模型。这些模型利用 Transformer 的自注意力机制和多头注意力机制来捕捉时空序列数据中的依赖关系和模式，并进行预测。相比于传统的循环神经网络，Transformer 模型具有并行计算的优势，并且能够处理长距离依赖关系，但其缺点在于模型复杂度高，需要较长的训练时间以及计算资源。

3.2.3.5 评价指标

在评价指标方面，我们使用四组评价指标来评估模型的表现：

(1) 训练集、验证和测试集的损失：显示模型的泛化能力；

(2) 平均绝对误差 (MAE)、均方根误差 (RMSE) 与平均绝对百分比误差 (MAPE)：直接检验预报精度；

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (29)$$

$$MAE = \frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \quad (30)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (31)$$

其中 y_i , \hat{y}_i 分别是地面实况和预测, n 是观测数据样本的总数。

(3) 常用气象指标测量污染阈值附近的性能：临界成功指数(CSI)、检测概率(POD)、虚警率(FAR)，其中 CSI 和 POD 越高，FAR 越小，模型的性能就越好

为了计算 CSI、POD、FAR，需要利用 PM2.5 污染阈值，我们根据中国《环境空气质量标准》建议的空气质量良好的分界点 75pg/m3 作为污染阈值，接着将预测值和地面真实值二值化为 0/1 矩阵，表示污染与否，然后计算正确命中(预测值=1，真值=1) 错过目标(预测=0，真值=1)和假警报(预测=1，真值=0)的数量，最后，使用公式 (31) (32) (33) 计算 CSI、FAR、FOD：

$$CSI = \frac{\text{正确命中数}}{\text{正确命中数} + \text{错过目标数} + \text{假警报数}} \quad (32)$$

$$FAR = \frac{\text{假警报数}}{\text{正确命中数} + \text{假警报数}} \quad (33)$$

$$POD = \frac{\text{正确命中数}}{\text{正确命中数} + \text{错过目标数}} \quad (34)$$

(4) 模型效率对比指标：为了比较 GAT-TSDFormer 与各个基线模型在相同情况下的训练预测效率，我们采用以下六个指标来进行对比分析：①模型参数量；②模型参数文件；③训练过程占用 GPU 内存大小；④测试过程占用 GPU 内存大小；⑤训练时间；⑥推理时间。

3.2.3.6 部分结果

由于 GAT-TSDFormer 的具体模型代码还未编写，因此我们仅根据上述的基线模型对应的开源代码进行实验。此外由于我们设备性能有限，面临着运行速度过慢，内存有限的问题，因此仅在京津冀地区进行了各个基线模型实验（不包括 GAT-TSDFormer 模型），且参数是我们事先指定，没有经过调参工作，下图是各个基线模型对北京市 PM2.5 浓度预测效果图。

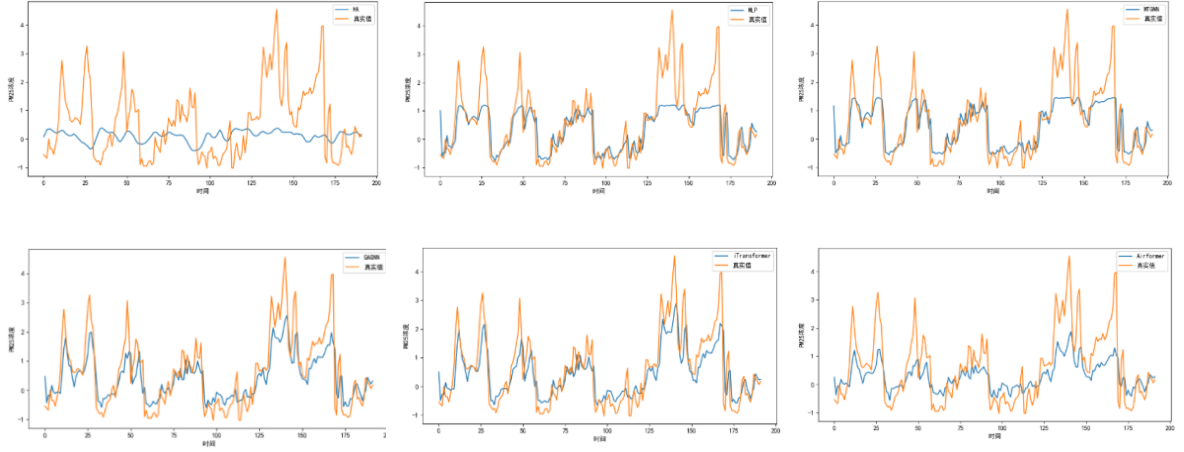


图 6 基线模型简单对比结果

3.2.4 挖掘模型组件，深入分析 PM2.5 时空发展传播机制

3.2.4.1 空间传播规律

从空间视角来看，一个城市的 PM2.5 的质量的水平，不仅仅与自身城市的工业排放，热岛效应等有关，还在风力的传输下受到其他城市 PM2.5 浓度等气象因素的影响。分析 PM2.5 的空间传播规律对各级政府在污染物治理和环境保护政策上达成区域合作，高效治理方面具有十分重要的意义。

从定量角度，在 GAT-TSDFormer 模型中的空间依赖性建模中，我们在 GNN 的基础上引入了具有方向的权重参数 $\alpha_{j \rightarrow i}$ 来反映不同城市之间 PM2.5 的空间传播效应。为此对于节点城市 i ，将其邻居节点城市 $j \in N(i)$ 按照 $\alpha_{j \rightarrow i}$ 的大小降序排序，便可得到序列 $z_{in} : \{z_{in_1}, \dots, z_{in_{N(i)}}\}$ ，同理按照 $\alpha_{i \rightarrow j}$ 的大小降序排序，便可得到序列 $z_{out} : \{z_{out_1}, \dots, z_{out_{N(i)}}\}$ ，其中 z_{in_j}, z_{out_j} 分别表示对节点城市 i ，其空间输入效应和输出效应第 j 强的节点城市。

对城市 i 而言，在序列 z_{in} 的基础上，便可清晰地得知自身受周边城市的 PM2.5 水平输入情况，对于城市 i 周边的城市而言，序列 z_{out} 的基础上，便可清楚知道城市 i 对其 PM2.5 的水平输出情况，因此对于某一个城市群，便可清晰得知 PM2.5 的空间传播规律，其上级政府部门便可对有强水平输出的城市进行重点治理，并对有强水平输入的城市进行一定的损失补贴，实现人力、物力、财力的合理配置，并以此构建 PM2.5 污染物传播的城市间对话机制，建立城市区域间空气质量的联合治理机制，更加科学高效地应对空气污染问题。

从定性角度，我们可以通过热力图将权重参数 $\alpha_{j \rightarrow i}, \alpha_{i \rightarrow j}$ 进行可视化，首先定义城市 i 到城市 j 的净输入效应为：

$$\alpha_{ij} = \begin{cases} \alpha_{i \rightarrow j} - \alpha_{j \rightarrow i}, & \alpha_{i \rightarrow j} - \alpha_{j \rightarrow i} > 0 \\ 0, & \text{其它} \end{cases} \quad (35)$$

接着根据 α_{ij} 大小绘制空间传播热力图，由于代码还未编码， α_{ij} 为随机数，不代表真实情况。

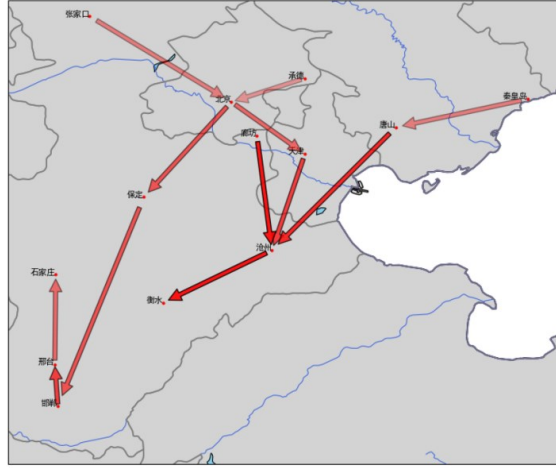


图 7 京津冀地区空间传播热力图(非真正结果)

3.2.4.2 时间发展规律

从时间视角来看，PM2.5 浓度不仅受到往期 PM2.5 数据的影响，并且受到往期各种气象因素的影响，而这些影响往往存在着时滞性，这对社会公众和政府部门的相关行动形成了巨大挑战。因此从数据驱动的角度阐明其时间发展规律，对于社会公众采取防护措施、政府部门颁布相关气象警报、实施相关政策具有非常重要的意义。

首先，我们可以通过序列分解来分析 PM2.5 时间发展规律：在 GAT-STDFormer 的时间依赖性建模中，我们引入了趋势、季节、短期波动成分的推理和外推组件，通过推理组件对输入变量序列进行分解从而得到各个成分序列，为此，我们可以对每一个输入序列分析其在趋势、季节、短期波动成分的波动情况，从数据驱动的角度去分析各个输入变量的时间波动情况。

其次，我们可以通过相关算子来分析 PM2.5 时间发展规律：在季节推理算子中，我们引入了**季节相关性矩阵** $A^{se} \in R^{\tau_{in} \times \tau_{in}}$ 来刻画不同季节之间的相关性，在得到季节相关性矩阵 A^{se} 之后，我们便可以去分析季节之间的内在的联系；在短期波动推理算子中，我们也引入了**短期波动相关性矩阵** $A^{st} \in R^{\delta \times \delta}$ 来刻画系统收到某种短期冲击后不同时刻的相关性，同样的在得到短期波动相关性矩阵 A^{st} 之后，我们便可以去分析收到短期冲击之后序列的趋势状况。

除了上述两个方法之外，我们也将探索如何从多项式回归层挖掘相关组件，来更好地分析 PM2.5 的时间发展规律，这是我们项目接下来进行模型优化的一个重点方向。

3.2.5 接入实时数据库，进行 PM2.5 浓度实时预测

在日常的公众生活，政府部门的预警发布和政策实施，都需要对当前和未来一段时间的 PM2.5 浓度的准确预测，因此基于前文中训练好的的最终模型之后，我们需要将相关的输入变量对应的实时数据库接入到模型之中，以此来进行一个实施的预测，其具体的实现步骤如下：

Step1 导入相关网站数据到数据库中：将与能够实时获取相关输入变量的网站例如中国环境监测总站等网站合作，通过 Python 的 requests 库，调用提供相关输入变量数据的 API，并将返回的响应解析为 JSON 格式，以便进一步处理，然后利用 Mysql-Connector-Python 或 PyMySQL 等数据库连接库，将解析后的 JSON 数据或结构化数据插入到 MySQL 数据库中，确保数据的持久化存储；

Step2 导入 MySQL 数据库中相关数据到 GAT-TSDFormer 模型中：从数据库中读取输入变量的相关数据，并进行必要的预处理工作如数据清洗和格式转换等，以适应模型的输入要求，处理后的数据随后被导入到模型中，用于训练或进行预测；

Step3 导入模型预测结果到 MySQL 数据库与相关网站中：将模型的预测结果再通过 API 存储进 MySQL 数据库，最后通过连接网站与数据库的 API 将模型预测结果反馈给网站，为网站对相关数据进行可视化等操作。



图 8 实时预测模型实现架构图

3.3 研究目标

本文以轻量化和解释性为视角，以城市群 PM2.5 浓度时空预测问题作为研究对象，首先通过论证美丽中国建设、绿色发展理念、区域联防联控联治下**高精度、轻量化、高可解释性的 PM2.5 时空预测模型**的重要性，接着查阅国内外相关领域的研究文献，指出了目前学者对该方面研究的上存在的三点不足，然后通过**深化特征工程、综合目前各模型优劣势以及时间序列分解思想**，设计了**基于 GNN 与注意力机制的 GAT 模型捕获空间依赖性以及基于时间序列分解的 TSDFormer 模型捕获时间依赖性**，最终构建了 **GAT-TSDFormer 模型**，力图实现以下预期研究目标：

(1) 理论目标

首先，通过将 PM2.5 相关影响因素划分为**垂直扩散因素和水平传播因素**，**深化特征工程**，为后续模型构建和模型解释性做基本准备；

其次，充分利用目前时空预测研究领域**中各个模型的优势**，结合**时间序列分解**这一重要思想，创新性地构建 **GAT-TSDFormer 时空预测模型**，在保证模型预测准确性地条件下**减低模型复杂度，减少模型参数量**，实现**轻量化、高精度的 PM2.5 时空预测模型**；

最后，通过**空间依赖性建模 GAT 模型中的注意力组件来挖掘空间传播规律**，通过**时间依赖性建模 TSDFormer 模型中的各个推理算子来挖掘时间发展规律**，进而综合分析 **PM2.5 时间发展和水平传播规律**，实现**高可解释性的 PM2.5 时空预测模型**。

(2) 社会目标

首先，**实现精准的 PM2.5 浓度预测**，帮助社会公众、政府部门及时掌握未来空气质量的走向，从而使得政府及时发布预警、公众及时能够及时采取防护，来有效减少 PM2.5 对公众健康和社会发展的负面影响；

其次，**挖掘 PM2.5 时间发展和水平传播规律**，助力政府部门从**预警预报转变到预防为主**。在**空间角度上实现重点区域重点治理、重点城市重点出击**：通过挖掘 PM2.5 空间传播规律，分析 PM2.5 以及相关因素在不同城市之间传播规律，来帮助政府将有限的人力、物力和财力投入到关键区域、关键城市，并且为城市群环境协同治理，区域合作提供科学依据；**在时间角度上实现重点实践重点作为、重点窗口重点把握**：通过挖掘 PM2.5 时间发展规律，了解各种相关因素与 PM2.5 之间复杂的作用关系，为政府及时采取相应行政措施如关停工厂、限制排放提供有效依据。

3.4 技术路线

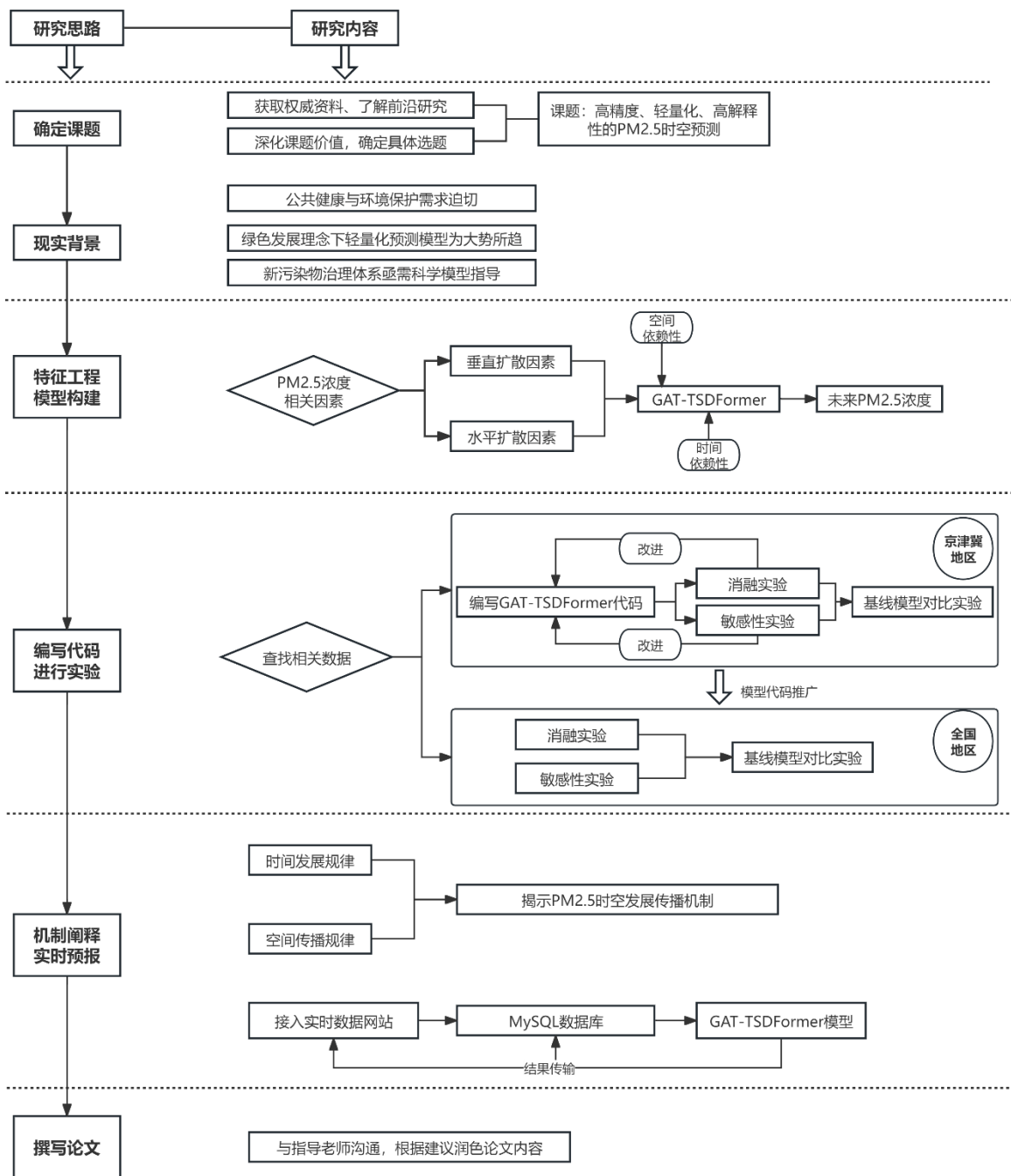


图 9 技术路线图

四、项目特色与创新点

4.1 项目特色

(1) 选题具有前沿性

在城市群污染物预测领域，基于深度学习的研究方法正逐渐成为研究的热点。本项目以时空预测为切入点，通过对空间和时间依赖性进行建模，不仅可以保证模型精度、实现轻量预测，还能够提供对 PM2.5 时间发展和空间传播规律的解释，研究方向不仅符合学界的前沿需求，同时也具有实际应用的重要意义；

(2) 方法具有推广性

本文采用的时空预测模型不仅适用于 PM2.5 的预测，还可广泛应用于其他污染物和气象预测，通过稍微修改该模型还可以应用于诸如交通流量预测、区域传染病预测等多个领域，这种通用性使得所提方法在环境科学以及相关领域具有广泛的应用前景；

(3) 结果具有实用性

本文构建的模型在接入相关的实时数据库后，能够实现对未来 72 小时内各个城市 PM2.5 浓度的实时预测，使得该模型不仅具有科研意义，更具实际应用的实用性。通过对实时数据的准确预测，政策制定者和相关群众可以及时采取措施，有效减少以 PM2.5 为主的雾霾带来的社会经济损失。

4.2 创新点

(1) 在特征工程方面：本文与以往研究采用的将 PM2.5 相关因素直接输入的处理方法不同，**创新性地**将相关因素划分为垂直扩散因素和水平因素，为后续模型构建提供了具有方向性的图结构以及进行空间传播规律揭示奠定基础；

(2) 在模型轻量化方面：本文基于**时间序列分解**的朴素思想，创新性地构建各序列成分推理和外推组件，替代了现有研究中基于注意力机制和局部全连接的模型组件，实现了**保证模型精度的前提下减少了参数个数，降低了模型复杂度**；

(3) 在模型解释性方面：不同于以往 PM2.5 时空预测模型仅关注预测精度，本文通过创新模型架构保留了高度可解释性：通过空间依赖性建模 **GAT 模型中的注意力组件来挖掘空间传播规律**，通过时间依赖性建模 **TSDFormer 模型中的各个推理算子来挖掘时间发展规律**，进而综合分析 **PM2.5 时间发展传播规律**，实现了高可解释性的 PM2.5 时空预测模型。

五、项目进度安排

5.1 研究计划



图 10 研究计划图

第一步：前期准备（2024.05-2024.07）

根据专家评审和指导老师提出的改进意见，对研究总体框架和研究思路与方法进行修改完善，最终确定模型的相关输入变量。在查阅了相关时空预测模型的大量文献后，确定最终的时空预测模型架构，并确定课题研究实施方案；

第二步：编写代码（2024.07-2024.09）

使用 Python 语言，查找网上现有 GNN 和 Transformer 变体的代码作为借鉴，将模型对应的组件进行具体代码实现，最终组合成我们的 GAT-TSDFormer 模型对应的代码；

第三步：京津冀地区实验（2024.09-2024.11）

在 KnowAir 网站上收集所需的变量数据，在京津冀地区来训练测试 GAT-TSDFormer 模型，并通过消融实验来验证各组件的合理性并进一步改进模型以及调试代码，并确定最终的模型架构以及代码主体，接着进行基线模型对比实验，最后基于相关模型组件对京津冀地区的时间发展和空间传播规律进行揭示；

第四步：全国范围推广（2024.11-2025.01）

寻找全国范围的模型输入变量的数据集，将在京津冀地区实验改进后的 GAT-TSDFormer 模型在参数设置等方面进行改进，以达到能够处理全国范围内 PM2.5 浓度预测的功能，最终再进行消融实验和基线模型对比实验，最后基于相关模型组件对京津冀地区的时间发展和空间传播规律进行揭示并可视化。

第五步：实现实时预测（2025.01-2025.02）

首先，寻找所需的实时数据网站，并导入到 MySQL 数据库中，接着将数据进行预处理后导入已经训练好的 GAT-TSDFormer 模型中，然后将模型的预测结果导出到 MySQL 数据库与相对应的 PM2.5 浓度预测可视化网站中；

第六步：论文撰写（2025.02-2025.04）

得到相对成熟的研究成果之后，我们将开始撰写以 GAT-TSDFormer 为核心的 PM2.5 浓度时空预测的论文，同时选择合适的刊物，在指导老师的建议下进行论文的修改，最终完成我们的论文并投稿。

5.2 人员分工

表 5 人员分工表

人员	分工内容
刘威腾	负责制定基本研究框架及方案，撰写申报书研究内容等部分 负责协调组内分工，整合研究资料 运用 Python 实现模型构建的代码实现 负责主要模型代码的编写
招铸峰	负责确定预测思路、模型的选择 负责实现所选模型对应组件的具体代码编写 负责训练测试 GAT-TSDFormer 模型 负责改进模型以及调试代码
王国杰	负责确定预测思路、模型的选择 负责进行各个实验 负责改进模型参数以推广至全国地区 负责预测结果可视化
杨添欣	负责收集国内外相关文献 负责申报书的立项依据等部分的撰写 负责图表美化，描绘思维导图
王萌	负责收集国内外相关文献 负责申报书的立项依据等部分的撰写 负责申报书及论文的排版工作

六、经费预算及依据

表 6 经费预算表

项目	用途	费用
书籍购买	Python 语言等相关书籍的购买、相关期刊文献的订购	500 元
数据购买	各气象监测站的 PM2.5、相关空气指标的数据	500 元
服务器租赁	华为云服务器租用，按需购买计算资源	15322 元
交通费用	跨校区咨询导师往返交通费用、实验交通费用	500 元
应急备用	紧急事件应急处理等	500 元
论文出版费	版面费用	2000 元
查重费用	知网大学生论文查重	300 元
总费用		19622 元

其中服务器租赁用于模型实验，其购买和费用截图如下：



图 1 服务器租赁详情图

七、项目预期成果

- 1.构建全国范围各城市的实时预测模型
- 2.完成《基于 GAT-TSDFormer 的 PM2.5 时空预测报告》
- 3.撰写“基于 GAT-TSDFormer 的 PM2.5 时空预测——基于轻量化与解释性视角”的学术论文

八、申请人承诺

我保证上述填报内容的真实性。如果获得资助，我与本项目组成员将严格遵守有关规定，切实保障研究工作的进行，按时报送有关材料。

项目负责人（签名）：刘威腾

时 间：2024 年 3 月 5 日

九、指导教师意见：

项目从 PM2.5 时空预测模型的高精度、轻量化、高可解释性等方向进行研究，提出了 PM2.5 时空预测新思路，研究方案及技术路线可行，具有一定的创新性，同意申报。

签 名：谢贤芳 王国长

时 间：2024 年 3 月 5 日

十、学生所在学院意见:

签字盖章:

时间:

十一、学校意见:

签字盖章:

时间: