

基于能源大数据的碳核算应用与分析方法

我国碳排放总量绝大部分来自能源活动和工业生产过程排放,能源活动碳排放约占我国碳排放的 86.5%,能源数据与能源活动具有物理相关性,并且与碳排放核算密切相关,基于能源数据能准确计算能源生产、传输、消费等各个环节碳排放量,而碳排放核算又是“双碳”管理和分析应用的重要基础。本章主要介绍国内外碳排放核算体系以及碳排放核算方法,然后聚焦基于能源大数据的直接与间接碳排放核算方法。

5.1 国内外碳排放核算体系

5.1.1 国际碳排放核算体系

政府间气候变化专门委员会(Intergovernmental Panel on Climate Change, IPCC)是牵头评估气候变化的国际组织,它是由联合国环境规

划署(United Nations Environment Programme, UNEP)和世界气象组织(World Meteorological Organization, WMO)于1988年建立的,旨在向世界提供当前气候变化及其潜在环境和社会经济影响的科学观点。20世纪90年代末,大多数国家的温室气体排放量只能粗略估计,不能准确测量,因此需要一个普遍认同、适用的测量方法,《IPCC国家温室气体清单指南》应运而生。该指南是IPCC为各国建立国家温室气体清单和减排履约提供的最新方法和规则,是迄今为止接受度最高、应用范围最广的国家层面温室气体排放清单指南。核算范围主要包括能源、工业过程和产品使用、农业林业和其他土地利用、废弃物以及其他5大分类、20个二级排放分类,在数据收集、方法学选择、不确定性、质量保证等方面进行了系统性指导,为《联合国气候变化框架公约》缔约国履行国际义务,依据本国国情开展清单编制提供了依据。《IPCC国家温室气体清单指南》根据核算方法的复杂程度与数据的可获取程度,将方法学分为3个层级,准确性和精度逐层提高。

欧盟、美国是国际上较早开展应对气候变化工作的地区或组织,在碳排放监测核算上积累了丰富的经验,欧洲环境署(European Environment Agency, EEA)及美国环境保护署(U. S. Environmental Protection Agency, EPA)已基于IPCC核算体系发布了符合自身情况的欧洲温室气体清单指南和美国温室气体清单指南。

EPA作为美国国家温室气体清单编制的领导组织,统筹建设了包含政府部门、学术机构、行业协会和环保组织等在内的合作团队。EPA编制的美国国家温室气体清单被认为“所提供的准确和完整的数据,能够在适当情况下向美国国内和国际气候变化政策提供执行依据和文本,并且通过参与联合国气候变化框架公约(United Nations Framework Convention on Climate Change, UNFCCC)和IPCC进程以及通过自身清单编制能力建设来国际化地改进温室气体清单”。美国国家温室气体清单在IPCC体系基础上形成6大分类(农业单独分

类)、25个二级排放分类,主要使用结合 AP-42《空气污染物排放系数汇编》的排放因子法,编制过程有成熟标准的组织模式、系数开发管理模式、数据质量管理模式和不确定管理模式。同时编制工作还得到先进的大气监测技术的支持,并建成了 IPCC 认可的统一碳排放数据库。清单披露周期为每年一次,数据滞后一年半左右。

EEA 是欧盟下属机构,是提供有关环境的可靠独立信息,参与制定、采用、实施和评估环境政策的重要部门。欧盟于 1990 年通过并建立 EEA 的法规,1993 年年底生效的 EMEP/EEA 排放清单指南(以前称为 EMEPCORINAIR 排放清单指南)为估算人为和自然排放源的排放提供了指导(IPCC 只提供人为源排放评估方法),之后转换为气候公约秘书处所要求温室气体的 IPCC 格式。该指南将排放源整体分为自然源和人为源,在 IPCC 体系的基础上将农业单独分类,并纳入火山、林火等,形成 6 大分类、24 个二级排放分类,成员国按国情基于 IPCC 选择方法,以排放因子法为主,该指南在数据收集方法、方法学选择、不确定性等方面进行了系统指导。清单每年更新一次,数据滞后一年半左右。

欧美现有碳排放核算体系较为完善,对于中国碳排放统计核算有一定参考借鉴意义,但并不完全适用。一是因为欧洲已结束了能源为主的高耗能高排放的重工业发展排放时代,取而代之的是以汽油为主的交通运输航空业碳排放,而中国正处于工业化升级转型阶段,煤炭燃烧产生碳排放在一段时间内仍占比很大;二是碳核算的具体过程虽然是客观的测量与计算,但其测量范围的划定及具体标准的制定,仍存在巨大的空间差异性,也带来了不确定性;三是中国各区域发展不一,碳排放差异明显,中国企业生产装备工艺更新改造、产能更新换代对实际碳排放总量的影响较大。中国为履行好《联合国气候变化框架公约》的国际义务,需要基于 IPCC 指南标准,借鉴发达国家的先进经验,结合自身国情编制适配性的方法体系。

5.1.2 中国碳排放核算体系

我国碳排放核算体系也遵循 IPCC 清单指南的整体框架,主要遵循《1996 年指南》、部分参考《2006 年指南》,结合国情制定了区域及行业的温室气体排放核算方法与指南。在区域层面,我国于 2011 年发布了《省级温室气体清单编制指南(试行)》,采用排放因子法,给出了能源活动、工业生产过程、农业、土地利用变化和林业、废弃物处理五大领域的温室气体排放核算方法。当前,31 个一级行政区均已完成 2005 年、2010 年、2012 年和 2014 年省级温室气体清单编制。在行业及企业层面,国家发展改革委组织制定了适用于全国范围的企业碳排放核算技术规范和指南,2013—2015 年先后发布 24 个行业的企业温室气体排放核算方法与报告指南(试行),涵盖发电、电网、钢铁、化工、电解铝、镁冶炼等重点行业。按照《巴黎协定》要求,发展中国家可以两年发布一次温室气体排放信息。我国于 2004 年首次发布 1994 年排放信息,后续又接连发布了 2005 年、2010 年、2012 年和 2014 年的国家温室气体排放信息。

开展省级温室气体清单编制既是提高国家温室气体清单质量的重要环节,也是编制省级应对气候变化规划或低碳发展规划的客观需要。我国编写《省级温室气体清单编制指南(试行)》,旨在加强省级清单编制的科学性、规范性和可操作性,为编制方法科学、数据透明、格式一致的省级温室气体清单提供有益指导。指南从能源活动、工业生产过程、农业、土地利用变化和林业、废弃物处理、不确定性、质量保证和质量控制七方面进行了指导说明。

1. 能源活动

能源生产和消费活动是我国温室气体的重要排放源。《省级温室

气体清单编制指南(试行)》总体上遵循《IPCC 国家温室气体清单指南》的基本方法,指南第1章“能源活动”,一是针对化石燃料燃烧活动排放源、生物质燃料燃烧排放源、煤炭开采和矿后活动逃逸排放源以及石油和天然气系统逃逸排放源进行了界定;二是根据清单编制方法、活动水平数据及其来源和排放因子数据及其确定方法等三方面对化石燃料燃烧活动、煤炭开采和矿后活动逃逸排放以及石油和天然气系统逃逸排放进行了系统说明;三是详细规范了能源部门清单报告格式;四是对电力调入调出二氧化碳间接碳排放量核算方法进行了规范统一。

2. 工业生产过程

工业生产过程温室气体排放指的是工业生产中除能源活动温室气体排放之外的其他化学反应过程或物理变化过程的温室气体排放。《省级温室气体清单编制指南(试行)》第2章“工业生产过程”,一是根据清单编制方法、活动水平数据及其来源和排放因子数据等三方面对水泥、石灰、钢铁以及电石生产过程中二氧化碳排放,己二酸和硝酸生产过程中氧化亚氮排放,一氯二氟甲烷(HCFC-22)生产过程中三氟甲烷(HFC-23)排放以及铝、镁、电力设备、半导体、氢氟烃等五个工业生产过程中其他温室气体排放进行了系统说明,其他生产过程或其他温室气体暂不报告;二是详细规范了工业生产过程清单报告格式。

3. 农业

省级农业温室气体清单包括四部分:一是稻田甲烷排放;二是农用地氧化亚氮排放;三是动物肠道发酵甲烷排放;四是动物粪便管理甲烷和氧化亚氮排放。数据获得的途径优先次序:统计部门数据、行业部门数据、文献发表数据、专家咨询数据。《省级温室气体清单编制

指南(试行)》第3章“农业”,一是针对稻田甲烷排放,从清单编制方法、排放因子确定方法、活动水平数据及相关参数以及排放量计算结果四方面进行系统说明;二是针对省级农用地氧化亚氮排放量,从清单编制方法、活动水平数据及来源、排放因子的确定方法以及排放量估算结果四方面进行系统说明;三是针对动物肠道发酵甲烷排放,从排放源界定、清单编制方法、活动水平数据及来源、排放因子确定方法及需要的数据以及排放量计算结果五方面进行系统说明;四是针对动物粪便管理甲烷和氧化亚氮排放,从甲烷排放、氧化亚氮排放以及温室气体排放量估算结果三方面进行系统说明;五是详细规范了农业部门温室气体清单报告格式。

4. 土地利用变化和林业

土地利用变化和林业温室气体清单,既包括温室气体的排放,也包括温室气体的吸收。在清单编制年份里,如果森林采伐或毁林的生物量损失超过森林生长的生物量增加,则表现为碳排放源,反之则表现为碳吸收汇。《省级温室气体清单编制指南(试行)》第4章“土地利用变化和林业”,一是对我国土地利用进行分类和定义;二是对省级土地利用变化和林业温室气体清单内容与范围进行了系统说明;三是针对森林和其他木质生物质生物量碳储量变化和森林转化温室气体排放,从清单编制方法、活动水平数据与确定方法以及排放因子与确定方法三方面进行系统说明;四是详细规范了土地利用变化和林业清单报告格式。

5. 废弃物处理

城市固体废弃物和生活污水及工业废水处理过程中,可以排放甲烷、二氧化碳和氧化亚氮气体,是温室气体的重要来源。《省级温室气体清单编制指南(试行)》第5章“废弃物处理”,一是对废弃物处理温室

气体排放源进行界定；二是针对固体废弃物处理，从填埋处理甲烷排放和焚烧处理二氧化碳排放两方面进行系统说明；三是针对废水处理，从生活污水处理甲烷排放、工业废水处理甲烷排放以及废水处理氧化亚氮排放三方面进行系统说明；四是详细规范了废弃物处理清单报告格式。

6. 不确定性

不确定性分析是一个完整温室气体清单的基本组成之一。应确定清单中单个变量的不确定性；将单个变量的不确定性合并为清单的总不确定性，进一步识别清单不确定性的主要来源，以帮助确定清单数据收集和清单质量改进的优先顺序。同时还要认识到统计方面也可能存在不确定性。《省级温室气体清单编制指南(试行)》第6章“不确定性”，一是对不确定性产生的原因和降低不确定性的方法进行系统说明；二是对量化和合并不确定性的方法进行系统说明。

7. 质量保证和质量控制

质量控制是一个常规技术活动，用于评估和保证温室气体清单质量，由清单编制人员执行。质量保证是一套规划好的评审规则系统，由未直接涉及清单编制过程的人员进行。在执行质量控制程序后，最好由独立的第三方对完成的清单进行评审。质量保证/质量控制过程和不确定性分析彼此间提供了有价值的反馈信息。清单估算和数据来源作为影响不确定水平和清单质量的关键部分，本身也是清单改进的工作重点。《省级温室气体清单编制指南(试行)》第7章“质量保证和质量控制”，一是对质量控制程序中一般质量控制程序和特定类别质量控制程序以及质量保证程序进行解释说明；二是对验证、归档、存档和报告四项工作进行规范指导。

值得注意的是，根据我国递交的国家温室气体清单显示，我国目前

农业活动产生的二氧化碳暂不记录,土地利用变化和林业为碳汇过程,废弃物处理产生的二氧化碳小于全社会碳排放的1%。

目前,我国也基于IPCC清单指南发布了发电、电网、钢铁、化工、电解铝、镁冶炼、平板玻璃生产、水泥生产、陶瓷生产、民航、石油和天然气生产、石油化工、独立焦化、煤炭生产、造纸和纸制品生产企业、其他有色金属冶炼和压延加工业、电子设备制造、机械设备制造、矿山、食品、烟草及酒、饮料和精制茶、公共建筑运营、陆上交通运输、氟化工,以及工业其他行业等24个行业、企业温室气体排放核算方法与报告指南(试行)。

相关指南从适用范围、引用文件和参考文献、术语和定义、核算边界、核算方法、质量保证和文件存档以及报告内容和格式规范七方面对各行业温室气体排放进行了指导说明。对行业、企业碳排放核算有着非常重要的意义,具体排放源和核算方法与区域碳排放核算较为类似,针对不同行业的生产过程碳排放核算会根据企业实际的生产工艺进行细化,在此不再赘述。

5.2 碳排放核算方法

总体来看,国内外碳排放核算标准涉及的通用碳排放核算方法可分为三类,一类是实测法,通过测量仪器直接针对二氧化碳的浓度、流量等进行实时监测,包括宏观层面的卫星监测和微观层面的烟气排放连续监测;另一类是计算法,基于排放活动数据或物质平衡关系,间接计算出二氧化碳排放量,分为物料平衡法和排放因子法两种方法;另外随着大数据技术不断发展,依托机器学习等技术,挖掘碳排放与其他解释变量的关联关系从而构建回归分析方程的方法也逐渐兴起,本书称为大数据计算法。

5.2.1 卫星监测法

卫星监测法是依赖卫星监测技术,通过大气二氧化碳浓度观测溯源碳排放的方法,主要应用在宏观大气二氧化碳浓度监测领域,如图 5-1 所示。

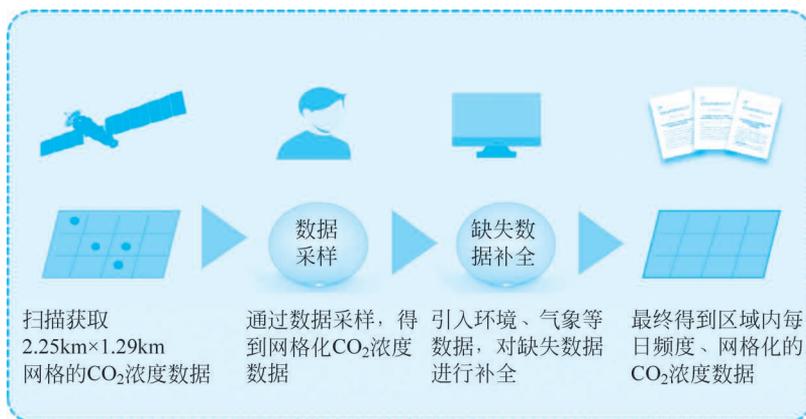


图 5-1 卫星监测法

卫星监测技术主要是运用了温室气体的太阳光反射,其遥感观测是基于光谱吸收特征,采用傅里叶变换红外光谱仪从干涉强度信号中提取光源辐射的发射光谱或物质的吸收光谱。卫星通过观测温室气体羽流吸收部分反射光谱的方式,完成对温室气体的监测工作。关键仪器为广角法布里-珀罗标准具,它由间隔几微米的 2 块反射板制成干涉仪。当光通过这个干涉仪时,不同波长的光会在不同的位置产生干涉,形成干涉图案。当太阳光穿过含温室气体的大气层时,红外光谱中某些波段的光波会被吸收,然后呈现出其特征光谱。卫星监测法的优点为:视野广阔,可以覆盖全球或大范围的区域,获取的信息量多,适合宏观温室气体的监测;瞬时成像,实时传输,快速处理,可以迅速获取

信息和实时动态监测；受地面影响小，可以探测一些人力难以到达或受干扰的区域；可以使用多种波段或多种传感器，获取不同类型或不同层次的信息。

卫星监测法的缺点为：其空间分辨率较低，可能限制了对小范围或点源碳排放的精细监测；观测存在周期性，对瞬时变化或短期波动不够敏感；气象条件、地表反射和高碳排放负荷区域可能导致观测困难，且监测成本相对较高。此外，该方法只能监测碳排放浓度状态，难以直接量化大气中的碳排放量，通常更适用于碳排放的校核，而不太适用于碳排放的具体治理工作。该方法适用部分区域碳排放宏观监测。

5.2.2 烟气排放连续监测法

烟气排放连续监测法主要采用连续排放监测系统或装置，通过直接测量烟气流速和烟气中二氧化碳浓度来计算温室气体的排放量，主要应用在企业的排放源设施（排口），相关监测数据和信息传送到环保主管部门，以确保排放企业排放物浓度和排放总量达标，如图 5-2 所示。



图 5-2 烟气排放连续监测法

针对气态污染物,监测系统共涉及3种方法,即直接测量法、稀释取样法和直接抽取法。直接抽取法,也称完全抽取法,适用范围广,一般来讲,只要气体分析仪本身的量程覆盖所测气体的测量范围,就可以采用直接抽取法。稀释取样法,即抽取少量烟气经过滤后,使用无污染的干燥空气,按照一定比例(如1:100)稀释,使烟气达到常规空气状态下不结露,再用普通环境空气监测仪分析。直接测量法又称现场连续监测法,不抽取烟气,对烟气的污染物进行直接测量,在烟道两侧分别安装发射装置和接收装置,发射装置发出一束红外线或紫外线穿过烟道到达接收装置,利用烟道作为样品气室吸收特征光谱进行测量。该方法的优点是计量结果准确性高,且对仪器要求低;计量结果实时性强,且可使用多组分同时测量。

烟气排放连续监测法存在一定的局限性,主要是该系统或装置体积大,构成复杂,设备投资高,且不利于实现实验室的校准;工作环境恶劣,工作气体多为高尘、高温、高湿、易腐蚀,维修成本较高。且根据监控系统流程,任一环节发生故障均会导致测量数据异常,会使估算结果存在较大误差。该方法通常设计用于监测特定工业点源的碳排放,在宏观层面上难以涵盖广阔的区域。该方法适用于火力发电、化工等工业领域碳排放监测。

5.2.3 物料平衡法

物料平衡法是以物质守恒和转化定律为基础,对其化学反应过程进行物料平衡计算的方法,即输入的物质质量必定等于输出的物质质量。该方法既可以用于整个生产过程,也可以用于局部生产过程。物料平衡法可以系统地、全面地计算和研究碳排放,其计算方法为取一定时期内燃料的碳平均含量和灰烬中的碳平均含量,根据差值计算碳排放量。该方法一般程序为先通过质量基准、物料组合和未知量等来制

作生产流程图,之后列出独立方程并校验方程数量与未知量是否一致,最后解方程组求出未知量。该方法主要应用于企业碳排放核算,其优点是微观层面计算更加准确,可以有效针对生产过程碳排放进行系统、全面地分析。

该方法的缺点主要体现在应用范围有限,难以应用于宏观层面的碳排放核算,计算结果受限于收集数据的设备,准确性难以保障。碳排放核算时计算过程复杂,数据要求高,数据需求多,工作量大,需要投入的人力成本高,需要详细了解工业生产过程数据以及生产工艺等情况。且针对具体环节和工艺的能源消耗和含碳量数据测量困难,测量成本高,限制了物料平衡法的使用。该方法适用于数据基础较好的企业,具体公式见式(5-1)

$$E = \sum_{i \in N} (Q_i \times C_i) - \sum_{j \in M} (Q_j \times C_j) \times \frac{44}{12} \quad (5-1)$$

式中, N 为能源种类集合; M 为燃烧输出产品和燃烧剩余物集合; Q_i 为第*i*种能源的消耗量; C_i 为第*i*种能源的含碳量; Q_j 为第*j*种输出产品和燃烧剩余物的产生量; C_j 第*j*种输出产品和燃烧剩余物的含碳量; $44/12$ 是C转换成 CO_2 的转换系数(即 CO_2/C 的相对原子质量)。

5.2.4 排放因子法

排放因子法是IPCC提出的一种碳排放核算方法。原理是依照《IPCC国家温室气体清单》,针对每一种排放源的活动数据与排放因子,以活动数据和排放因子的乘积作为碳排放量。主要应用在国家、区域、行业等不同维度对象碳排放核算。排放因子数值可以通过IPCC、EPA、EEA等权威机构获得。排放因子法的一般程序为先确定目标的系统流程,之后通过调研、分析和整理资料来收集目标活动

水平数据,然后基于前两步工作进行清单数据分析,再根据清单分析结果标定碳排放因子数据,最后对碳排放因子核算结果进行质量控制。针对能源活动、工业生产过程等每种排放源将活动水平数据乘以排放因子得到碳排放量,如图 5-3 所示。该方法的优点在于原理简单易于理解、应用范围广、成本低、有成熟的官方数据作为计算输入。

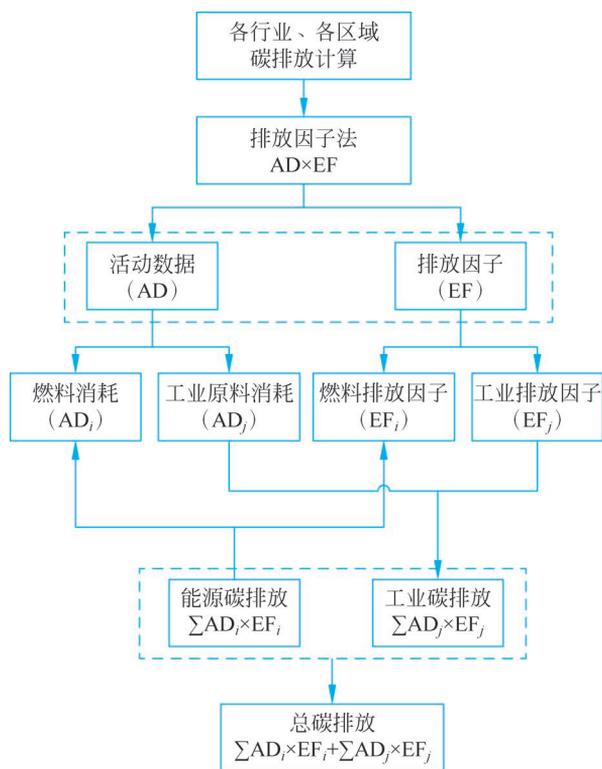


图 5-3 排放因子法

该方法存在一定的局限性,主要体现在依赖大量统计数据,导致其时效性和分辨率不高,且受其他因素影响导致不确定性较大。计

算时活动水平数据获取的准确度和精度会影响碳排放核算的精度。且由于能源类型、燃烧设备、燃烧技术、制作工艺等差异,采用缺省值或区域平均水平排放因子可能导致碳排放核算结果与实际情况存在偏差。因此该方法适用于排放源较为稳定的宏观及中观层面碳排放计算。

5.2.5 大数据计算法

大数据计算法是基于机器学习等人工智能方法或计量经济学等统计方法,通过算法模型构建碳排放与其他变量的关联关系,利用分析模型计算碳排放数据。目前比较主流的方法是 DMSP/OLS(美国国防气象卫星计划/卫星运行的线性扫描系统)夜间灯光数据分析法和 NPP/VIIRS(对地观测卫星/红外辐射成像仪)夜间灯光数据分析法。郭忻怡等基于国家温室气体排放清单指南提出的核算方法,利用 DMSP/OLS 夜间灯光数据和美国国家极轨卫星数据,构建了碳排放空间回归模型,对原煤、原油、天然气等 9 种能源消费的碳排放量进行了测算。张永年等基于 DMSP/OLS 夜间灯光数据,通过 Theil-Sen Median 趋势分析法与 Mann-Kendall 检验法构建了碳排放拟合模型,对我国 14 年间能源碳排放量进行了预估分析。许燕燕等基于我国 DMSP/OLS 夜间灯光数据和能源统计数据,对成渝地区碳排放进行了研究。以 DMSP/OLS 夜间灯光数据分析碳排放时空动态为例。先对夜晚灯光数据和碳排放数据进行多重检验,之后输入面板数据模型得到碳排放估计值,再通过碳排放统计数据与估计值进行验证,最后得到碳排放的时空动态。瞿植等基于 DMSP-OLS 与 NPP-VIIRS 夜间灯光数据,通过构建转换模型,将能源消费碳排放量分配到像元尺度,并进行时空变化分析。对比普通最小二乘法(Ordinary

Least Squares, OLS)模型、地理加权回归(Geographically Weighted Regression, GWR)模型计算能源消费碳排放对景观生态指数、遥感生态指数(Risk Screening Enviromental Indicators, RSEI),以及代表性的生态系统服务功能的影响,以探测能源消费碳排放的生态效应。关伟等对东北三省碳排放的演变趋势进行了动态分析,基于 DMSP/OLS 夜间灯光数据,构建东北三省的碳排放模型,从栅格级、市级及县级模拟分析 1994—2013 年碳排放的时空演变特征,其详细流程见图 5-4。该方法,优势在于可以让各种数据相互之间建立数学关联,之后可以根据其他维度数据与目标数据的映射关系,用质量更高、时效性更好的数据来替换原本的输入数据,且数据规模越大,数据集越复杂,分析效果越好。

但大数据算法也存在局限性,通常大数据算法输入的特征数据没有物理层面的实际含义,导致结果的可解释性较差。除此之外,相关模型在训练过程中需要足够的样本空间容量,如果训练数据量不足,模型可能无法捕捉到数据的变化趋势,从而影响它在新数据上的泛化能力和准确性。此外,模型可能无法捕捉到数据中的潜在结构或特征,从而影响它对问题的全面理解和分析能力。该方法多应用于宏观和中观层面的碳排放核算。

将上述五种主流碳排放核算方法从优势、局限性和适用尺度三方面进行了归纳总结,详情见表 5-1。其中卫星监测法、烟气排放连续监测法和物料平衡法适用于企业和产品的微观层面核算,对分地区、分行业的宏观层面核算并不适用。排放因子法虽适用于分地区、分行业碳排放核算,但容易受到其他因素影响导致不确定性较大。而大数据算法不仅适用于分地区、分行业碳排放核算,还可以多维度、高频次地进行分析,且更容易和统计学相关算法模型结合,可以更好地推广应用。

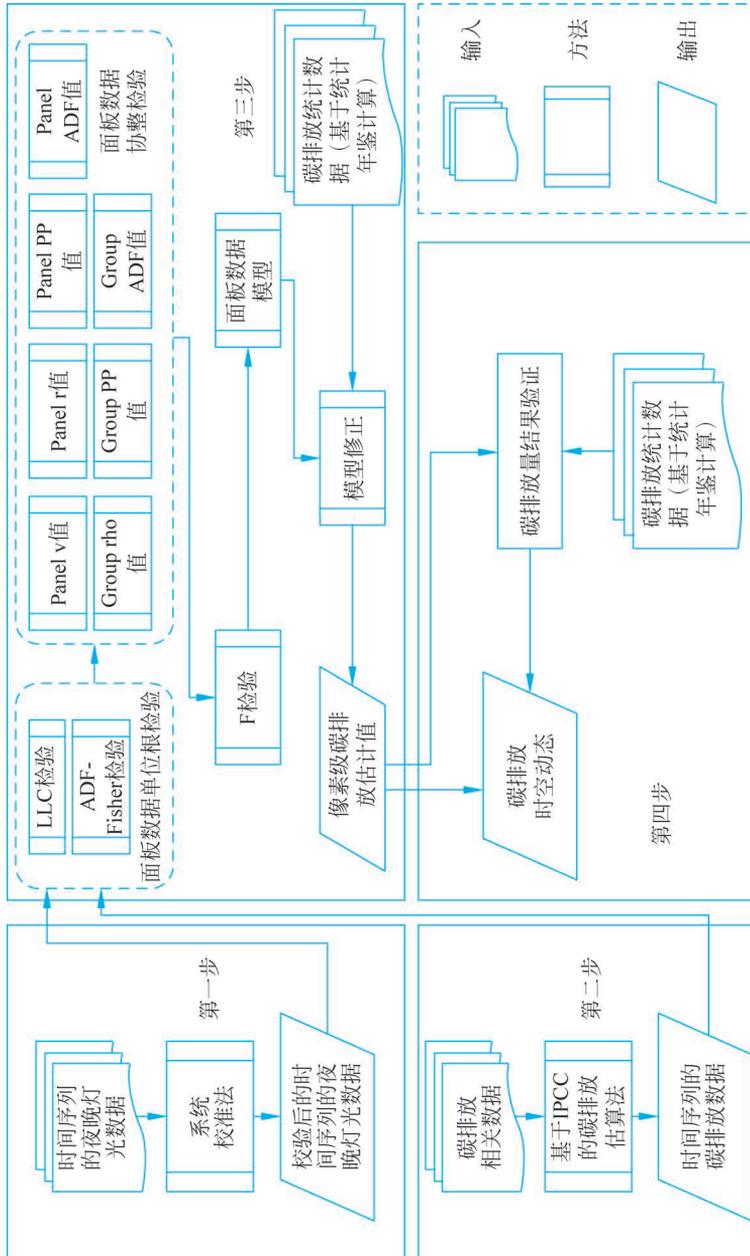


表 5-1 碳排放核算方法比较

方 法	优 势	局 限 性	适用尺度
卫星监测法	监测范围广；实时传输；受地面影响小；获得信息较多	空间分辨率较低、观测存在周期性、监测成本相对较高、不太适用于碳排放的具体治理工作	宏观
烟气排放连续监测法	计量结果准确性高；对仪器要求低；计量结果实时性强；可使用多组分同时测量	设备投资高、不便于实现实验室的校准、维修成本较高、估算结果存在较大误差、难以涵盖广阔的区域	微观
物料平衡法	微观层面计算更加准确；可以有效针对生产过程碳排放进行系统、全面的分析	难以应用于宏观碳排放核算,受限于收集数据的设备;计算过程复杂,数据要求高,人力成本大,测量成本高	中观、微观
碳排放因子法	原理简单易于理解、应用范围广、成本低；有成熟的官方数据做支撑	依赖大量的统计数据；受其他因素影响导致不确定性较大，且时效性和分辨率不高	宏观、中观、微观
大数据计算法	数据是实时数据，质量高且统计口径一致；可以进行多维度、高频次的分析	结果的可解释性较差；需要足够的样本空间容量，否则容易影响泛化能力和准确性	宏观、中观

5.3 基于能源大数据的直接碳排放核算方法

能源数据是碳排放核算核心的活动水平数据，但目前全量能源数据采集获取的时效性还存在不足，快速准确的碳排放核算需要具有更时效性、分辨率、准确性的数据进行支撑。电力数据作为一种能源数据，和煤、油、气等其他数据关联性强，且电力行业良好的信息化基础，保证了电力数据在时效性、准确性等方面具有显著优势。因

此构建电力数据与能源消费、产品产量的回归分析模型,实现能源活动和工业生产过程等活动水平月度数据的快速计算,并与碳排放因子结合计算碳排放量,是一条结合大数据计算法和排放因子法切实可行的技术路线。

5.3.1 理论框架

将大数据分析技术与机器算法模型相结合来计算碳排放具有充分的研究基础和良好的应用前景。本章在 IPCC 碳核算指南的基础上,结合机器学习算法,建立基于能源大数据的碳排放核算方法学体系,并对电力数据及其他碳排放相关数据进行数据清洗、数据填补、数据平滑降噪等预处理工作。随后分解碳排放关键驱动因素,加以训练构建回归分析模型,并与国内外权威数据库进行对比验证,不断调优参数、更新模型,如图 5-5 所示。

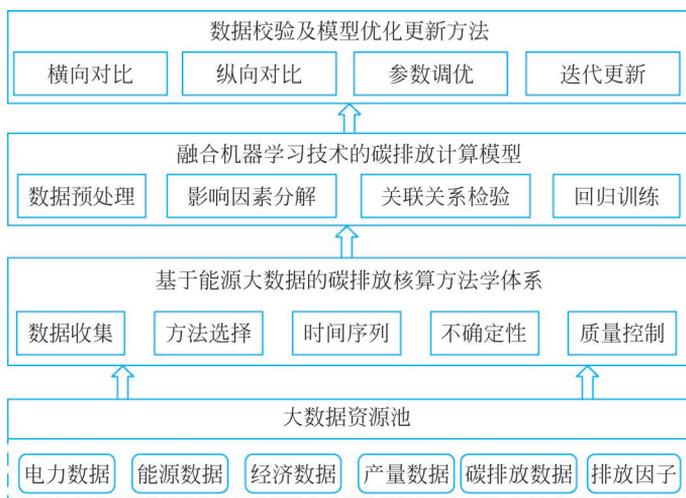


图 5-5 基于能源大数据的碳排放核算理论框架

5.3.2 方法体系

由于能源大数据碳排放核算方法采用充分结合大数据计算法及排放因子法的技术路线,因此在开展计算时需要遵循碳排放核算指南及有关标准,本章介绍的能源大数据碳排放核算方法在 IPCC 碳排放核算体系基础上进行了继承和扩充,并引入机器学习、大数据分析等技术,在方法选择和时间序列上进行了创新,形成了能源大数据碳排放核算方法学。

1. 数据收集

IPCC 碳排放核算方法学中数据收集步骤主要运用数据库系统理论,使用数据库采集、系统日志采集等方法收集现有数据、生成新数据和调整数据库清单。IPCC 碳排放核算方法学的数据来源包括:国家统计局、部门专家及利益相关组织、其他国际国内专家、IPCC 排放因子数据库、出版统计资料的国际组织(如联合国、欧盟统计局或国际能源署、经济合作与发展组织、国际货币基金组织)、环境类著作、期刊和报告中的科技论文、《联合国气候变化框架公约》缔约国提交的国家清单报告等。

能源大数据碳排放核算方法学的数据收集在 IPCC 碳排放核算方法学原有理论和方法的基础上进行了扩充,主要引入了基于网络理论的 Web 数据挖掘法。Web 数据挖掘是一项结合了数据处理、信息处理、可视化、数理统计等领域的综合技术,可以对 Web 数据资源中未知且有潜在应用价值的信息进行提取。能源大数据碳排放核算方法学应用 Web 数据挖掘技术收集外部公开数据集。

2. 方法选择

IPCC 方法学主要分三类:排放清单、剩余源和吸收源。排放清单是

指用于收集和记录温室气体排放信息的工具。能源大数据碳排放核算方法学选择的方法是在 IPCC 碳排放核算方法学原有理论和方法的基础上进行了创新,具体包括运用机器学习理论、统计学、计量经济学、大数据分析等理论的基于传统拟合回归的线性回归(Linear Regression, LR)模型、基于时间序列的自回归分布滞后(Auto-Regressive Distributed Lagged, ARDL)模型、基于深度学习的人工神经网络(Artificial Neural Network, ANN)模型、基于机器学习的梯度提升迭代决策树(Gradient Boosting Decision Tree, GBDT)模型。能源大数据碳排放核算方法学应用这些方法对能源活动和工业产品产量进行测算。

1) LR 模型

LR 模型是利用数理统计中的回归分析,来确定两种或两种以上变量间相互依赖的定量关系的一种统计分析方法。若两个或者多个变量之间存在“线性关系”,则可以通过历史数据,厘清变量之间的联系,建立一个有效的模型,通过一个或多个自变量来预测因变量结果,其数学表达式可描述为

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon \quad (5-2)$$

式中, p 为样本个数; x_i 是第 i 个样本; β 为系数; y 为因变量; ε 为随机误差项。

2) ARDL 模型

ARDL 模型是一种较新的协整检验模型,更加适用于小样本集的模型训练,可解释性强。ARDL 模型使用因变量的滞后项和自变量的当期及滞后项进行构建,直接估计出变量间的长期效应和短期的平衡关系,其数学表达可描述为

$$y_t = c_0 + c_1 t + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{i=1}^q \beta'_i X_{t-i} + u_t \quad (5-3)$$

式中, $p \geq 1, q \geq 0, X_{t-i}$ 为自变量; y_t 与 y_{t-1} 为因变量; u_t 为误差项; β'_i 为对应自变量的系数; ϕ_i 表示对应因变量滞后项的系数; t 表

示因变量 y_t 取值的时间。该模型包含了自回归和分布滞后两种模型，因此同时考虑了序列相关性和动态影响。

3) ANN 模型

深度学习是一种通过简单模型的逐层堆叠获取高层非线性特征的深度神经网络方法，它的基本特点是试图模仿人类大脑神经元之间传递和处理信息的模式。目前应用比较广泛的人工神经网络方法有反向传播(Back Propagation, BP)神经网络、Hopfield 神经网络、Kohonen 神经网络等。ANN 模型是对人类神经活动的一种模拟，其结构由多个人工神经元相互连接生成，包含 3 部分：输入、激活函数与输出。其中，输入一般为一个 n 维的向量，输出可表示为

$$\mathbf{N}_{\text{out}} = f(\omega \mathbf{N}_{\text{in}} + b) \quad (5-4)$$

式中， \mathbf{N}_{in} 表示人工神经元的输入； ω 与 b 分别为人工神经元的连接权重与偏置； f 为该神经元的激活函数，当该人工神经元接收外来输入，经处理大于阈值后会进行响应并输出。

4) GBDT 模型

GBDT 模型是通过采用加法模型(即基函数的线性组合)，以及不断减小训练过程产生的残差来达到将数据分类或者回归的方法。GBDT 模型可对多个弱分类器进行不断迭代，形成一个强分类器。GBDT 模型的基础为决策树，通过计算负梯度值不断地对模型进行迭代，形成多个决策树的级联，使残差变化不断减小，从而提升模型的计算速度与准确率。GBDT 模型通过梯度提升方法实现多个决策树的集成，从而解决过拟合的问题，有效提高了训练速度。GBDT 模型可以表示为决策树的加法模型，即映射 $F(x)$ 若干分类器组成，其表达式如下：

$$F(x, P) = \sum_{m=0}^M \beta_m h(x, a_m) \quad (5-5)$$

式中， P 为回归数的参数，其表达式为 $F(x, P) = \{\beta_m, a_m\}_0^M$ ； x 为输入样本； a_m 为第 m 棵树的参数； β_m 为第 m 棵树的权重；函数 $h(x, a_m)$ 表示具有参数 a_m 和 x 的决策树； M 表示决策树的数量。

3. 时间序列

能源大数据碳排放核算方法学的时间序列在 IPCC 碳排放核算方法学原有理论和方法的基础上进行了创新,具体包括运用机器学习理论、统计学、计量经济学、大数据分析等理论的差分整合移动平均自回归模型(Autoregressive Integrated Moving Average Model, ARIMA 模型)、季节性差分自回归滑动平均(Seasonal Autoregressive Integrated Moving Average, SARIMA)模型、遗传算法(Genetic Algorithm, GA)、蚁群优化(Ant Colony Optimization, ACO)算法、粒子群优化(Particle Swarm Optimization, PSO)算法、二次规划(Quadratic Programming, QP)算法等方法。能源大数据碳排放核算方法学应用 ARIMA 模型、SARIMA 模型等时间序列预测算法对时序数据进行填补,应用 QP 算法、GA 算法、ACO 算法、PSO 算法等对年度数据进行月度拆分。

1) ARIMA 模型

ARIMA 模型又称整合移动平均自回归模型(移动也可称作滑动),是时间序列预测分析方法之一。ARIMA(p, d, q)中,AR 是“自回归”, p 为自回归项数,MA 为“滑动平均”, q 为滑动平均项数, d 为使之成为平稳序列所做的差分次数(阶数)。ARIMA(p, d, q)模型是 ARMA(p, q)模型的扩展。ARIMA(p, d, q)模型可以表示为

$$\Phi(B)(1-B)^d y_t = \theta(B)\epsilon_t \quad (5-6)$$

式中, B 是延迟算子(Lag Operator)。

2) SARIMA 模型

SARIMA 模型是基于 ARIMA 模型的扩展,也是时间序列预测分析方法之一。SARIMA 模型在 ARIMA(p, d, q)模型基础上又增加了 3 个超参数(P, D, Q),以及一个额外的季节性周期参数 s 。SARIMA(p, d, q)(P, D, Q, s)总共 7 个参数,可以分成 2 类,3 个非季节参数(p, d, q)和 4 个季节参数(P, D, Q, s)。SARIMA 模型的一般形式为

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^D y_t = c + \theta(B)\Theta(B) \quad (5-7)$$

式中, S 和 D 分别表示季节周期的长度和季节差分的阶数; B^S 表示季节后移算子。

3) GA 算法

GA 算法是一种通过模拟自然进化过程搜索最优解的方法,将染色体的一些性质如“选择、交叉、变异”用在了求解过程中,其原理可概述为: 概率性保留当前解,或者对当前解进行交叉和变异处理,更新为新的解,然后用一个评价函数(适应度)评价解的好坏,进而确定这些解在下一代解中出现的概率,往复迭代,直到满足终止条件。

4) ACO 算法

ACO 算法是一种用来寻找优化路径的概率型算法。其灵感来源于蚂蚁在寻找食物过程中发现路径的行为。基本思路为: 用蚂蚁的行走路径表示待优化问题的可行解,整个蚂蚁群体的所有路径构成待优化问题的解空间。路径较短的蚂蚁释放的信息素量较多,随着时间的推进,较短路径上累积的信息素浓度逐渐增高,选择该路径的蚂蚁个数也越来越多。最终,整个蚂蚁群体会在正反馈的作用下集中到最佳的路径上,此时对应的便是待优化问题的最优解。

5) PSO 算法

PSO 算法是一种进化计算技术,源于对鸟群捕食的行为研究。该算法最初是受到飞鸟集群活动的规律性启发,进而利用群体智能建立的一个简化模型。PSO 算法在对动物集群活动行为观察基础上,利用群体中的个体对信息的共享使整个群体的运动在问题求解空间中产生从无序到有序的演化过程,从而获得最优解。PSO 算法的基本思想是通过群体中个体之间的协作和信息共享来寻找最优解。

6) QP 算法

QP 算法的对象是现代控制理论中以状态空间形式给出的线性系统,而目标函数为对象状态和控制输入的二次型函数,其数学表达式可通常描述为

$$y_{m,r} = \underset{y_{t,r}}{\operatorname{argmin}} \{ (x_{m,r} - y_{m,r})^T \mathbf{A} (x_{t,r} - y_{m,r}) \} \quad (5-8)$$

式中, $x_{t,r}$ 以及 $x_{m,r}$ 为待优化变量; $y_{t,r}$ 以及 $y_{m,r}$ 为因变量; \mathbf{A} 为约束矩阵。

4. 不确定性

IPCC 碳排放核算方法学在不确定性过程中运用概率论和统计学等理论,使用误差传播、蒙特卡洛模拟等方法,为估算报告与年排放、清除量、排放和清除随时间变化的趋势有关的不确定性提供指导。

1) 误差传播

误差传播用来估算整个清单中和关注年份与基年趋势中单个类别的不确定性。其通过使用误差传播公式来计算源类别的不确定性,具体分为两类,一是由于今年和关注年份的某一特定源类别和气体的排放增加引起的不确定性;二是由于关注年份的某一特定源类别和气体排放增加引起的不确定性。通过源类别不确定性的简单合并从而估算出一年中总体的不确定性以及趋势的不确定性。

2) 蒙特卡洛模拟

蒙特卡洛模拟适用于详细的分类不确定性估算,尤其是不确定性大、分布非正态的情况。蒙特卡洛模拟的主要理论基础是概率统计理论,主要手段是随机抽样、统计试验。它的基本思想是:为了求解问题,首先建立一个概率模型或随机过程,使它的参数或数字特征等于问题的解,然后通过对模型或过程的观察或抽样试验来计算这些参数或数字特征,最后给出所求解的近似值,解的精确度用估计值的标准误差来表示。

能源大数据碳排放核算方法学的不确定性估算继承了 IPCC 碳排放核算方法学原有理论和方法。

5. 质量控制

IPCC 碳排放核算方法学在时间序列步骤中运用统计学和管理学等理论,使用质量保证(Quality Assurance, QA)/质量控制(Quality Control, QC)、验证活动等方法提高国家温室气体清单的透明性、一致

性、可比较性及准确性。

QA/QC 与验证活动应该是清单编制过程中的重要组成部分。QA/QC 与验证的结果可能会引起对清单或类别不确定性估算的重新评估,以及排放或清除估算的后续改进。例如,QA/QC 的结果可能会指出应该成为改进工作重点的某个类别估算方法学中的特定变量。

QC 是一个常规技术活动系统,用于在编制清单时评估和保持质量。它由清单编制人员执行。QC 系统旨在:

- (1) 提供定期和一致检验来确保数据的内在一致性、正确性和完整性;
- (2) 确认和解决误差及疏漏问题;
- (3) 将清单材料归档并存档,记录所有 QC 活动。

QC 活动包括一般方法,如对数据采集和计算进行准确性检验,对排放和清除计算、测量、估算不确定性、信息存档和报告等使用业已批准的标准化规则。QC 活动还包括对类别、活动平均数据、排放因子、其他估算参数及方法的技术评审。

QA 是一套规划好的评审规则系统,由未直接涉足清单编制/制定过程的人员进行评审。评审确认可测量目标已实现,确保清单代表了在目前科学知识水平和数据获取情况下排放和清除的最佳估算,而且支持 QC 计划的有效性。

能源大数据碳排放核算方法学的 QC 在 IPCC 碳排放核算方法学原有理论和方法的基础上进行了扩充,具体包括运用偏差分析理论的智能偏差分析等方法。

5.3.3 技术路线

基于能源大数据碳排放核算方法学,结合电力大数据,本章探索构建了能源大数据碳排放核算模型,包括两大步骤:一是以电算能(产量),将历史电量、GDP 等数据与能源活动、工业产量数据进行训练,构

建分析模型,通过当月电力等数据计算出当月能源活动和工业产量数据;二是以能(产量)算碳,将能源活动和工业产量数据乘以相关碳排放因子得到当月碳排放量,如图 5-6 所示。

1. 以电算能(产量)

以电算能(产量)是模型的主要部分,包括两个步骤。步骤一,回归分析计算,使用电力等数据回归分析计算能源活动及工业生产过程产量数据;步骤二,频度转换,将年度数据拆分到月度,如图 5-7 所示。

1) 步骤一: 回归分析

回归分析计算,主要是利用输入和输出量之间的关联关系,选用基于时间序列的 ARDL 模型进行计算、预测未来趋势,是整体模型的核心。是将历史电量、GDP 等数据与能源活动、工业产量数据进行训练,构建回归分析计算模型。

2) 步骤二: 频度转换

由步骤一计算出来的能源活动和工业生产过程产量数据是年度数据,需要通过步骤二将年度的低频数据转换为月度的高频数据。通过二次规划算法,将能源活动、工业生产过程产量与用电量数据之间的抽象距离作为目标函数,求取每月的能源活动和工业生产过程产量数据,使目标函数取最小值(即抽象距离最短),如图 5-8 所示,以能源消耗数据为例,通过二次规划算法,实现年度能源消耗数据拆分为月度能源消耗数据,其变化趋势和月度电量的变化趋势相吻合。

2. 以能(产量)算碳

以能(产量)算碳是将以电算能(产量)输出的能源消费和产品产量数据,乘以对应的碳排放因子得到碳排放量。碳排放因子包括《省级温室气体清单编制指南(试行)》《企业温室气体排放核算方法与报告指南(试行)》等定义的通用因子,见图 5-9。

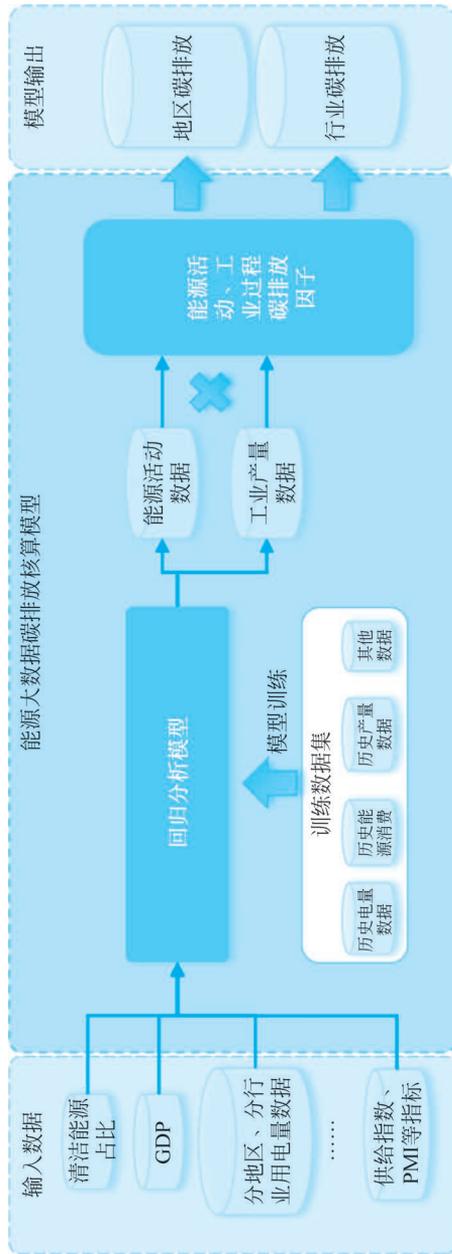


图 5-6 “能源大数据碳排放核算模型”技术路线



图 5-7 以电算能(产量)架构

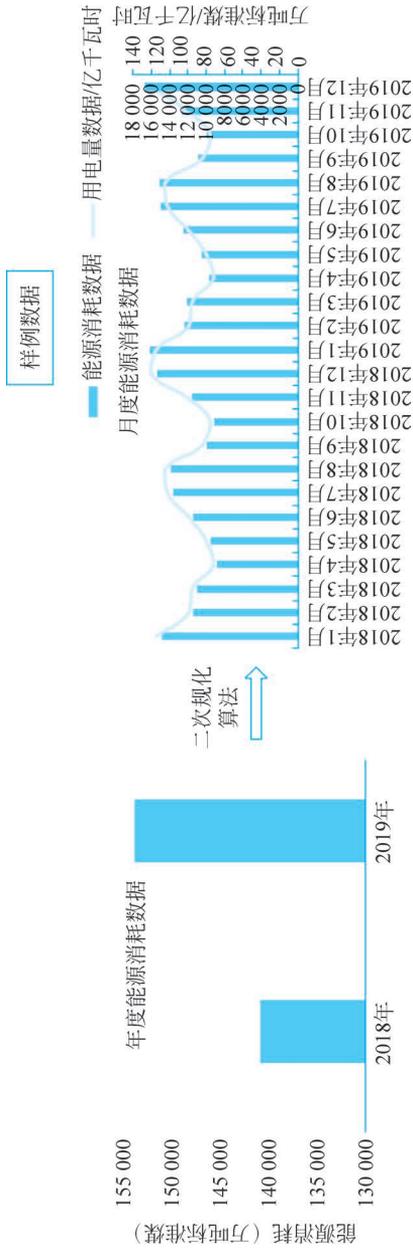


图 5-8 频率转换例图

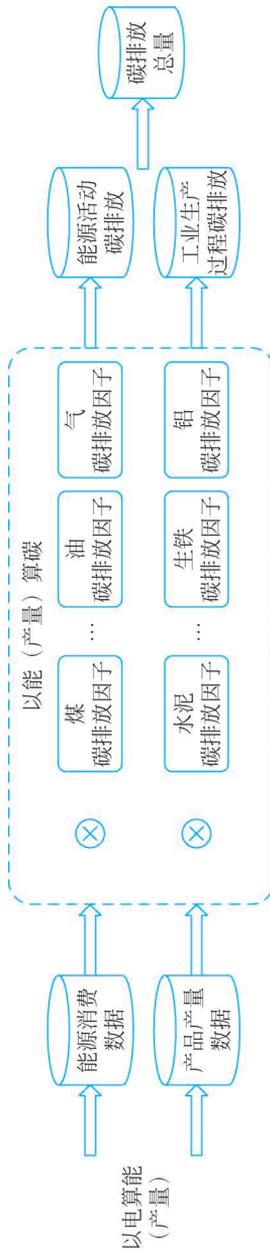


图 5-9 以能(产量)算碳架构

5.4 基于能源大数据的间接碳排放核算方法

根据国际通用做法,电力碳排放核算需要获得电力碳排放因子。目前各国碳排放因子的选取主要参考《IPCC 国家温室气体清单指南》,对于全国电力碳排放因子,我国迄今共发布过三次,其中在2017年国家发展改革委首次公布全国电网碳排放因子,随后生态环境部分别于2022年和2023年更新了全国电网平均碳排放因子值。对于区域电网碳排放因子,目前处于长时间的停滞状态,国家发展改革委于2013年和2014年分别发布了2010年区域碳排放因子和2013年及2014年区域碳排放因子,在此之后已有近10年未更新。区域电力碳排放因子的更新缺乏时效性,导致已有电力碳排放因子计算结果不够精细,在计算空间维度方面,分时分区的计算颗粒度不够,地理尺度上计算覆盖区域大、计算缺乏时效性,未考虑可再生能源发电类型和数量在不同区域的发展差异;在计算时间维度方面,计算时长一年内只有一个时段,未考虑可再生能源发电的时变特性等问题。当前针对碳排放强度的应用方式单一,难以直观反映电力碳排放强度的时空演化特征。

为了实现电力碳排放准确计量,急需开展电碳关联分析,研究能反映绿色属性的电力碳排放因子计算方法,提升电力碳排放因子的时空分辨率,获得动态电力碳排放因子。

本节提出动态电力碳排放因子计算方法,从电碳耦合的理论框架进行介绍,对电网节点和分区域动态电力碳排放因子计算原理进行详细说明,并给出应用领域。

5.4.1 理论框架

电力碳排放因子计算方法针对两类主体，一类是节点，另一类是区域。通常电力碳排放因子也可以称为电力碳强度，对节点与区域又分别衍生出发电侧、用电侧等不同角度的电力碳强度计算方法，最终聚合并等效为节点电力碳强度的碳量平衡计算方程与区域电力碳强度计算方法。对于区域电力碳强度计算方法，本书提出了考虑特高压绿电输送的区域电力碳强度计算方法，以及区域电网与省间分层联合的电力碳强度计算方法，如图 5-10 所示。

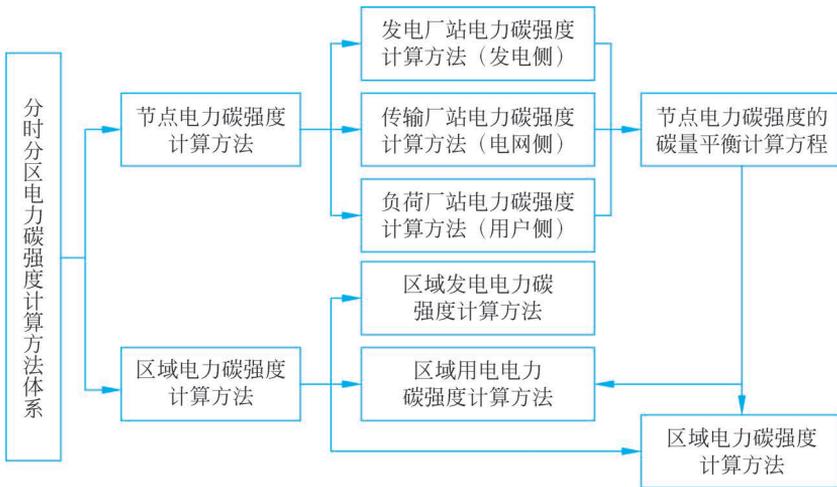


图 5-10 动态电力碳排放因子计算理论框架

5.4.2 节点电力碳强度计算方法

变电站作为电能生产、传输与消费的基本单元，也是电力碳强度计算的基本单元。变电站可分为发电厂站、传输厂站和负荷厂站，各类型

节点碳强度计算方式也有所不同。

1. 发电站节点

定义变电站电力碳强度(Substation Carbon Intensity of Electricity, SCIE),用来表示在该厂站消耗单位电量时,在发电侧所对应产生的碳排放量,其量纲为 $\text{kgCO}_2/(\text{kW} \cdot \text{h})$,或者吨 $\text{CO}_2/(\text{MW} \cdot \text{h})$ 。

对于发电站 k ,其 SCIE 等于发电厂的发电碳强度,对于一个连接多台发电机组的发电站,其碳排放量和电力碳强度分别为

$$F_{k,t} = \sum_{k \in G_k} G_{k,t} \cdot e_k \quad (5-9)$$

$$\text{SCIE}_{k,t} = \frac{\sum_{k \in G_k} G_{k,t} \cdot e_k}{\sum_{k \in G_k} G_{k,t}} \quad (5-10)$$

式中, $F_{k,t}$ 为发电站 k 在时段 t 的总碳排放量; $G_{k,t}$ 为发电机组 k 扣除厂站用电后,在时段 t 的总发电量; e_k 为发电机组 k 的碳排放因子。

2. 传输厂站节点碳输入量与电力碳强度

线路作为变电站之间传输电能的载体,在传输电能的同时,也传输对应的碳量。定义线路 (i, j) 的始端和终端有功潮流、线损分别用 $P_{i,j}, P_{j,i}, \Delta L_{ij}$ 表示,变电站 i 的电力碳强度用 SCIE_i 表示,则有如下的线路碳量平衡方程:

$$P_{i,j,t} \cdot \text{SCIE}_{i,t} = -P_{(j,i),t} \cdot \text{SCIE}_{i,t} + \Delta L_{(i,j),t} \cdot \text{SCIE}_{i,t} \quad (5-11)$$

式中, $P_{i,j,t} \cdot \text{SCIE}_{i,t}$ 为从变电站 i 在线路 (i, j) 的始端流入的碳量; $P_{(j,i),t} \cdot \text{SCIE}_{i,t}$ 为从变电站 i 在线路 (i, j) 的终端流入变电站 j 的碳量; $\Delta L_{(i,j),t} \cdot \text{SCIE}_{i,t}$ 为与线路 (i, j) 的线损所对应的碳量。

对于传输厂站 k , 其输入的总碳量等于与其相连的输入线路碳量之和, 其 SCIE 等于其总输入碳量除以其总输入有功功率, 或者总输出有功功率:

$$F_{k,t} = \sum_{i \in B_{1k}} (-P_{(k,i),t}) \cdot \text{SCIE}_{i,t} \quad (5-12)$$

$$\begin{aligned} \text{SCIE}_{k,t} &= \frac{\sum_{i \in B_{1k}} (-P_{(k,i),t}) \cdot \text{SCIE}_{i,t}}{\sum_{i \in B_{1k}} P_{(k,i),t}} \\ &= \frac{\sum_{i \in B_{1k}} (-P_{(k,i),t}) \cdot \text{SCIE}_{i,t}}{\sum_{j \in B_{0k}} P_{(k,j),t}} \end{aligned} \quad (5-13)$$

式中, B_{1k}, B_{0k} 分别为传输厂 k 输入和输出线路的集合。

3. 负荷厂站碳输入量与电力碳强度

对于负荷厂站 k , 其输入的总碳量等于与其相连的输入线路碳量之和, 其 SCIE 等于其总输入碳量除以其总输入有功功率, 或者总线路输出有功功率与总负荷之和:

$$F_{k,t} = \sum_{i \in B_{1k}} (-P_{(k,i),t}) \cdot \text{SCIE}_{i,t} \quad (5-14)$$

$$\begin{aligned} \text{SCIE}_{k,t} &= \frac{\sum_{i \in B_{1k}} (-P_{(k,i),t}) \cdot \text{SCIE}_{i,t}}{\sum_{i \in B_{1k}} P_{(k,i),t}} \\ &= \frac{\sum_{i \in B_{1k}} (-P_{(k,i),t}) \cdot \text{SCIE}_{i,t}}{\sum_{j \in B_{0k}} P_{(k,j),t} + L_{k,t}} \end{aligned} \quad (5-15)$$

式中, B_{1k}, B_{0k} 分别为负荷厂 k 输入和输出线路的集合; $L_{k,t}$ 为负荷厂 k 的总负荷。

对于拥有分布式发电量 $G_{k,t}$ 的负荷厂站,有负荷内部平衡法和厂站总体平衡法两种处理方式。

负荷内部平衡法:其总输入碳量保持不变,将分布式电源发电量从总负荷中扣除,这种情况下,分布式电源只影响拥有分布式发电的那部分负荷的碳强度,其总输入碳量和碳强度计算公式不变,把式(5-15)中的 $L_{k,t}$ 的计算公式改为: $L_{k,t} = L'_{k,t} - G_{k,t}$,这种情况适合于低碳园区。

厂站总体平衡法:将分布式发电电量及其碳量计入厂站总输入碳量,据此计算厂站电量碳强度:

$$F_{k,t} = \sum_{i \in B_{lk}} (-P_{(k,i),t}) \cdot \text{SCIE}_{i,t} + G_{k,t} \cdot e_k \quad (5-16)$$

$$\begin{aligned} \text{SCIE}_{k,t} &= \frac{F_{k,t}}{\sum_{i \in B_{lk}} (-P_{(k,i),t}) + G_{k,t}} \\ &= \frac{F_{k,t}}{\sum_{j \in B_{Ok}} P_{(k,j),t} + L_{k,t}} \end{aligned} \quad (5-17)$$

4. 基于厂站节点电力碳强度的碳量平衡方程

深入探索以上变电站节点碳强度计算方法,可得到通用变电站节点碳量平衡方程。

对于任何类型的变电站 k ,其输出电量与输入电量一定是平衡的:

$$\sum_{i \in B_{lk}} (-P_{(k,i),t}) + G_{k,t} = \sum_{j \in B_{Ok}} P_{(k,j),t} + L_{k,t} \quad (5-18)$$

相应地,其输出碳量和输入碳量也一定是平衡的,考虑到式(5-17),产生两种碳量平衡方程:

$$\sum_{i \in B_{lk}} (-P_{(k,i),t}) \cdot \text{SCIE}_{i,t} + G_{k,t} \cdot e_k$$

$$= \left(\sum_{j \in B_{Ok}} P_{(k,j),t} + L_{k,t} \right) \cdot \text{SCIE}_{k,t} \quad (5-19)$$

$$\begin{aligned} & \sum_{i \in B_{Ik}} (-P_{(k,i),t}) \cdot \text{SCIE}_{i,t} + G_{k,t} \cdot e_k \\ & = \left(\sum_{j \in B_{Ik}} (-P_{(k,i),t}) \right) + G_{k,t} \cdot \text{SCIE}_{k,t} \end{aligned} \quad (5-20)$$

对于一个具有 S 个变电站的系统, SCIE 的计算公式可以用一个 S 维的矩阵方程表示, 定义:

$$\begin{cases} \mathbf{H}_t = \text{diag} \left(\sum_{j \in B_{Ok}} P_{(k,j),t} + L_{k,t} \right) - \mathbf{P}_t \\ \mathbf{F}_t = [G_{1,t} \cdot e_1, \dots, G_{s,t} \cdot e_s, \dots, G_{S,t} \cdot e_S]^T \end{cases} \quad (5-21)$$

\mathbf{P}_t 是由 $-P_{(k,i),t}$ 构成的、对角线元素为 0 的 $S \times S$ 的电能输入下三角稀疏矩阵, SCIE_t 为 S 维用电侧电力碳强度向量, 则式(5-18)可写成矩阵方程式(5-22), 则方程式(5-19)可写成矩阵方程式(5-24):

$$\mathbf{H}_t \cdot \text{SCIE}_t = \mathbf{F}_t \quad (5-22)$$

同理, 对应于方程式(5-21), 定义:

$$\begin{cases} \mathbf{H}'_t = \text{diag} \left(\sum_{i \in B_{Ik}} (-P_{(k,i),t}) + G_{k,t} \right) - \mathbf{P}_t \\ \mathbf{F}_t = [G_{1,t} \cdot e_1, \dots, G_{s,t} \cdot e_s, \dots, G_{S,t} \cdot e_S]^T \end{cases} \quad (5-23)$$

$$\mathbf{H}'_t \cdot \text{SCIE}_t = \mathbf{F}_t \quad (5-24)$$

对于方程式(5-21)和式(5-23), 其左端矩阵中的值由线路上的潮流、发电机组发电量或者负荷厂站的负荷量确定, 因而进行 SCIE 计算的先决条件是已知电网的状态估计或潮流计算结果。

在计算出 SCIE 后, 可以计算出对应不同空间区域的区域用电侧电力碳强度(Carbon Intensity of Electricity Sector, CIES):

$$\text{CIES}_{a,t} = \frac{\sum_{s \in A_s} L_{s,t} \cdot \text{SCIE}_{s,t}}{\sum_{s \in A_s} L_{s,t}} \quad (5-25)$$

通过以上公式也可以计算出各个电网区域的电能交换量,以及交换碳量。

5.4.3 区域电力碳强度计算方法

1. 区域发电侧

区域发电侧电力碳强度与其燃料类型和发电效率有关,煤电机组最高,天然气机组次之,风、光、水等可再生能源发电效率较小,一般取0,核电机组为低碳机组。可见,区域发电侧电力碳强度只取决于区域内的发电量,以及随之产生的碳排放量。区域内,可再生能源发电和低碳发电比例越大,区域的电力碳强度越低。

$$\text{CIEG}_{a,t} = \frac{\sum_{k \in G_a} G_{k,t} \cdot e_k}{\sum_{k \in G_a} G_{k,t}} \quad (5-26)$$

式中,CIEG_{a,t}为区域a在时段t的发电电力碳排放强度;G_a为区域a内所有发电机的集合;G_{k,t}为发电机k在时段t的发电量;e_k为发电机k的碳排放因子。

2. 区域用电侧

对于电能纯输出地区,其发电侧电力碳强度与用电侧电力碳强度相同。对于电能输入区域,其用电侧电力碳强度需要考虑邻近电网输入的电量,即与邻近区域电网之间的电力交易,以及随之带来的碳排放量,因而邻近电网的用电侧电力碳强度会影响区域的用电侧电力碳强度。

$$\text{CIEC}_{a,t} = \frac{\sum_{k \in G_a} G_{k,t} \cdot e_k + \sum_{n \in N_a} I_{n,a,t} \text{CIEC}_{n,t}}{\sum_{k \in G_a} G_{k,t} + \sum_{n \in N_a} I_{n,a,t}} \quad (5-27)$$

式中, $I_{n,a,t}$ 为在时段 t 从邻近区域 n 输入区域 a 的电量; N_a 是区域 a 输入电能的区域集合; $CIEC_{n,t}$ 为邻近区域 n 的用电电力碳强度。

3. 电网区域电能平衡方程

电网区域电能平衡方程用来描述电网中各个区域之间的电能平衡关系,通常包含了发电量、传输损耗、区域间电能交换几部分,用于确保电网中电能供需平衡,保证电力系统的稳定运行。随着可再生能源和分布式发电的增加,电网区域电能平衡的管理和优化将变得更加复杂,同时也为提高能源效率和减少温室气体排放提供了新的机遇。

$$\begin{aligned} & \sum_{k \in G_a} G_{k,t} + \sum_{n \in N_a} I_{n,a,t} \\ & = \sum_{l \in L_a} D_{l,t} + NL_{a,t} + \sum_{m \in M_a} O_{a,m,t} \end{aligned} \quad (5-28)$$

式中, $I_{n,a,t}$ 为在时段 t 从邻近区域 n 输入区域 a 的电量; N_a 是区域 a 输入电能的区域集合; L_a 是区域 a 负荷的集合; $D_{l,t}$ 是负荷 l 在时段 t 的用电量; $NL_{a,t}$ 是区域 a 在时段 t 的网损; M_a 是从区域 a 输出电能的区域集合; $O_{a,m,t}$ 是在时段 t 从区域 a 输出到区域 m 的电能。

4. 区域碳排放量输入输出平衡方程

电网区域碳排放量输入输出方程表征的是区域内输入电能所产生的碳排放与输出及网损电能所产生的碳排放的平衡关系。电力碳强度作为方程中的关键部分,用于衡量发电过程中的环境负荷,与区域能源结构、能源效率、技术水平等息息相关。一般地,可以通过加大可再生能源的发电比例降低区域电力碳强度,从而使区域碳排放量维持在一个较低的水平。

$$\begin{aligned} & CIEC_{a,t} \left(\sum_{l \in L_a} D_{l,t} + NL_{a,t} + \sum_{m \in M_a} O_{a,m,t} \right) \\ & = \sum_{k \in G_a} G_{k,t} \cdot e_k + \sum_{n \in N_a} I_{n,a,t} \cdot CIEC_{n,t} \end{aligned} \quad (5-29)$$

式中, $CIEC_{n,t}$ 为区域 n 的输入侧电力碳强度。

5. 考虑特高压绿电输送的区域电力碳排放因子计算方法

特高压作为跨区域大容量传输线路,其中可再生能源输送比例较高,目前电网区域电力碳排放因子未考虑绿电跨区域传输的碳流量传导影响,导致区域电力碳排放因子计算准确度不高。为解决这一问题,本章提出了考虑特高压绿电输送的区域电力碳排放因子计算方法。考虑到实际特高压线路跨区域大容量绿电传输对区域间碳流量计算的影响因素,基于通用区域电力碳强度计算公式,引入特高压绿电传导因子,进一步提升区域电力碳强度计算结果精度。

在应用层面,依照数据需求从电力统计年鉴、能源统计年鉴等数据源获取发电厂碳排放因子数据,并筛选 2018 年各省各月数据作为测试样本,对数据进行统一审查梳理。

数据导入后,基于区域电网内不同类型发电机组的上网发电量数据计算总碳排放量,而后,得出计及特高压线路在区域间的绿电量传输的碳流量传导计算。最后,建立区域电网电力碳排放因子计算方程,并计算求解,如图 5-11 所示。

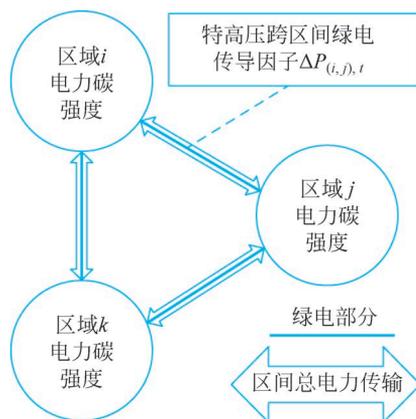


图 5-11 考虑特高压绿电输送的区域电力碳强度示意图

6. 区域电网与省间分层联合的电力碳强度计算方法

在实际电网拓扑结构中,采用区域电网调度与省级电网调度的分层结构运行方式,但在获取电网计算数据时,区域电网与省级电

网密切依赖,无法构建相对独立的区域电力碳强度单元计算,如图 5-12 所示。由于区域间和省间的电量交换数据是按照两个调度层级获取,无法直接获取全国各个省间的电量交换数据来直接计算省级电力碳排放因子,需要先获取区域电网间的电量交换数据,再获取区域内的省间电量交换数据进行计算,所以,需要建立一种区域与省间电量交换的两级联合电力碳排放因子的计算方法以解决以上问题。

基于区域电力碳强度计算方法,提出区域电网与省间分层联合的电力碳强度计算方法。考虑到实际区域电网调度控制区域与省级调度控制区域的电网拓扑结构交错复杂,基于通用区域电力碳强度的计算理论难以适用,利用创新的区域电网与省间分层联合的电力碳强度计算方法可更好地解决区域电网与省级电网间的耦合依赖计算问题。区域间分层是将电网划分为电网区域层级和省级,先在电网区域级计算其碳排放因子,再计算省级碳排放因子。通过分层算法,将复杂的电网层级关系考虑到计算过程中,保证计算结果的准确性。

5.4.4 节点碳排放因子的并行计算

目前全国电网因为特高压和省级联络线,已经形成跨区输送、整体成网的特点,如果计算全国细颗粒度(到地市-县-园区)级别的电力碳强度需要进行维度非常大的矩阵运算,如果只计算主网节点碳排放因子,矩阵维度数量级为百余级,计算量庞大,计算时间长,无法利用电网高频数据的特点(目前潮流有功采集间隔达到 5min 一次),计算更短时间断面的数据。

采用划分区域等效计算的方法可以实现局部精确计算,在考虑跨区输入影响方面,无法准确衡量输入、输出的相互影响。

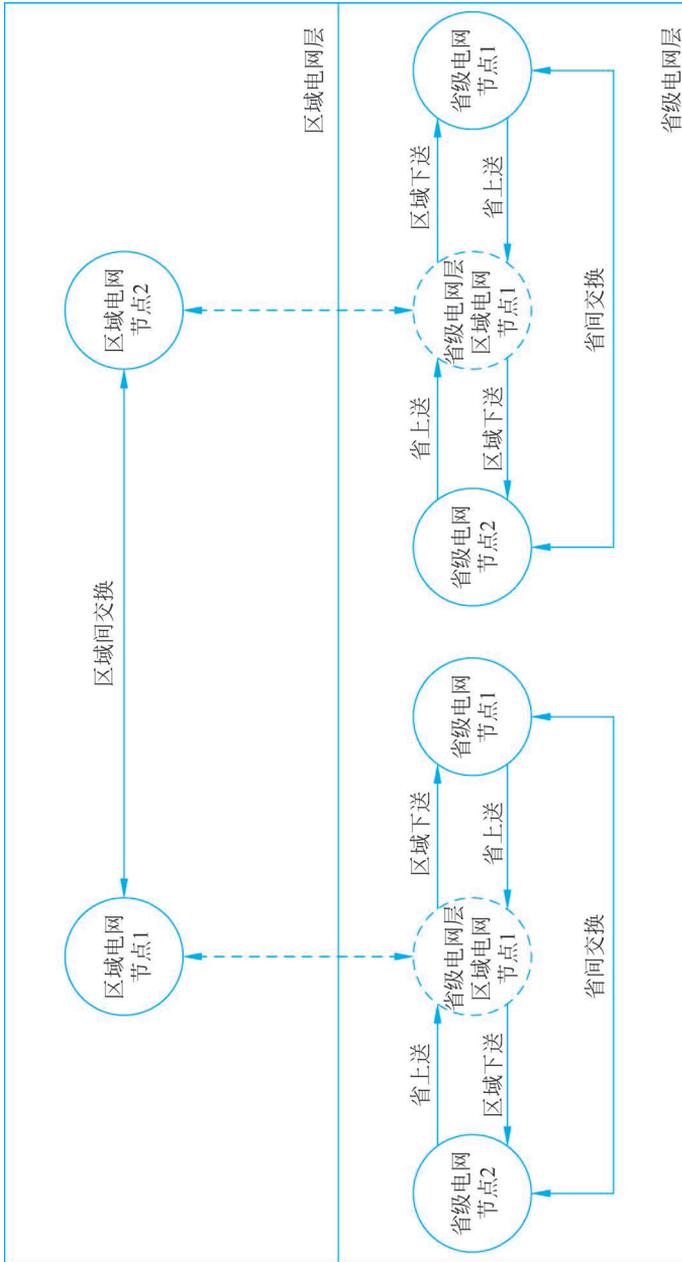


图 5-12 区域电网与省级电网逻辑关系图

区域电力碳排放因子计算过程涉及的明细数据和计量数据具有战略意义, 汇聚计算容易造成风险, 需要一种分布式的计算方法, 通过参数传递的方式实现协同计算, 且分布式计算结果与集中计算结果保持一致。

通过使用分布式的方法构建全国范围内的电力网络模型, 实现空间维度精确到厂站及用电企业级别、时间维度上达到小时级别的电力碳排放因子核算。

传统的节点电力碳排放因子需要构建平衡方程:

$$E_i \epsilon_i + \sum e_{j,i} \lambda_j = \lambda_i (E_i + \sum e_{j,i}) \quad (5-30)$$

然后各节点联立得到

$$\begin{pmatrix} E_1 + \sum e_{j,1} & -e_{2,1} & \cdots & -e_{n,1} \\ -e_{1,2} & E_2 + \sum e_{j,2} & \cdots & -e_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ -e_{1,n} & -e_{2,n} & \cdots & E_n + \sum e_{j,n} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} = \begin{pmatrix} E_1 \epsilon_1 \\ E_2 \epsilon_2 \\ \vdots \\ E_n \epsilon_n \end{pmatrix} \quad (5-31)$$

矩阵化

$$\mathbf{P}_n \boldsymbol{\lambda}_n = \mathbf{C}_n \quad (5-32)$$

则各节点的碳排放因子为

$$\boldsymbol{\lambda}_n = \mathbf{P}_n^{-1} \mathbf{C}_n \quad (5-33)$$

根据计算公式

$$\boldsymbol{\lambda}_n = \begin{pmatrix} \lambda_x \\ \lambda_e \end{pmatrix} = \mathbf{P}_n^{-1} \mathbf{C}_n$$

$$= \begin{pmatrix} \mathbf{P}_{\mathbf{X}}^{-1} + \mathbf{P}_{\mathbf{X}}^{-1} \mathbf{T}_{e\mathbf{X}} \mathbf{P}_e \mathbf{T}_{\mathbf{X}e} \mathbf{P}_{\mathbf{X}}^{-1} & -\mathbf{P}_{\mathbf{X}}^{-1} \mathbf{T}_{e\mathbf{X}} \mathbf{P}_e \\ -\mathbf{P}_e \mathbf{T}_{\mathbf{X}e} \mathbf{P}_{\mathbf{X}}^{-1} & \mathbf{P}_e \end{pmatrix} \begin{pmatrix} \mathbf{C}_{\mathbf{X}} \\ \mathbf{C}_e \end{pmatrix} \quad (5-34)$$

则有

$$\lambda_e = \mathbf{P}_e \mathbf{C}_e - \mathbf{P}_e \sum_{\chi \in \mathbf{X}} \mathbf{P}_{\chi e} (\mathbf{P}_{\chi} + \mathbf{I}_{e\chi})^{-1} \mathbf{C}_{\chi} \quad (5-35)$$

最终

$$(\mathbf{P}_{\chi} + \mathbf{I}_{e\chi}) \lambda_{\chi} + \mathbf{P}_{e\chi} \lambda_e = \mathbf{C}_{\chi} \quad (5-36)$$

$$\lambda_{\chi} = (\mathbf{P}_{\chi} + \mathbf{I}_{e\chi})^{-1} \mathbf{C}_{\chi} - (\mathbf{P}_{\chi} + \mathbf{I}_{e\chi})^{-1} \mathbf{P}_{e\chi} \lambda_e \quad (5-37)$$

整个计算过程需要求取 $(\mathbf{P}_{\chi} + \mathbf{I}_{e\chi})^{-1}$ 的逆矩阵和 $(\mathbf{P}_e - \mathbf{T}_{\mathbf{X}e} \mathbf{P}_{\mathbf{X}}^{-1} \mathbf{T}_{e\mathbf{X}})^{-1}$ 的逆矩阵,如果平均每个省有 m 个节点,边缘点有 l 个节点,共有 N 个区域需要求 N 次 m 阶矩阵的逆运算和 1 次 l 阶矩阵的拟运算,因为边缘点数量不多,一般 $l \ll m$ 。

求逆矩阵运算时间复杂度为 $O(M^3)$,全国如果有 M 个节点,考虑 $M = Nm + l$,则全国整体时间复杂度为 $O((Nm + l)^3)$,使用分块法计算时间复杂度为 $O(m^3) \times N + O(l^3)$ 。

从现实意义上讲, $(\mathbf{P}_{\chi} + \mathbf{I}_{e\chi})^{-1}$ 可以由 N 个节点并行计算后,再计算总节点 $(\mathbf{P}_e - \mathbf{T}_{\mathbf{X}e} \mathbf{P}_{\mathbf{X}}^{-1} \mathbf{T}_{e\mathbf{X}})^{-1}$,最后计算出 λ_n 。节点并行计算示意图见图 5-13。

从图 5-13 可以看出,节点并行计算中,各节点对自己内部及输入内部的有功和拓扑是已知的,对于边缘点之间拓扑和有功是部分已知,对其他节点是未知的,在联合计算过程中,各节点可以提供 $\mathbf{P}_{\chi e}$ $(\mathbf{P}_{\chi} + \mathbf{I}_{e\chi})^{-1} \mathbf{P}_{e\chi}$ 和 $\mathbf{P}_{\chi e} (\mathbf{P}_{\chi} + \mathbf{I}_{e\chi})^{-1} \mathbf{C}_{\chi}$ 数据, $\mathbf{P}_{\chi} + \mathbf{I}_{e\chi}$ 所代表的节点内拓扑经过上述变换是无法复原的,可以看作一种隐私保护,总节点只需要知道 \mathbf{P}_e ,通过各节点数据求出 λ_e 之后分发给各省,求出各节点

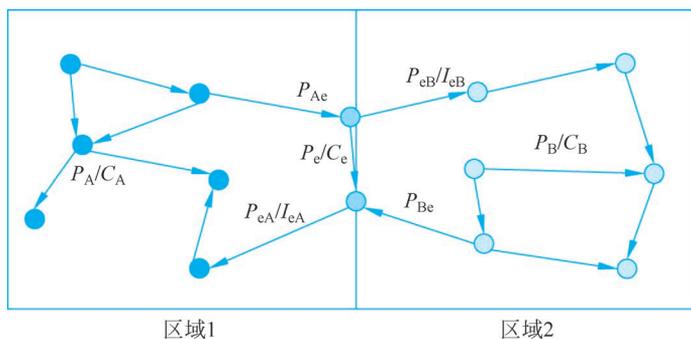


图 5-13 节点并行计算示意图

λ_χ , 最后统一和总节点校对。计算步骤如下:

步骤	各节点	总节点
1	计算 $(\mathbf{P}_\chi + \mathbf{I}_{e\chi})^{-1}$	
2	计算 $(\mathbf{P}_\chi + \mathbf{I}_{e\chi})^{-1} \mathbf{P}_{e\chi}$ ①	
3	计算 $(\mathbf{P}_\chi + \mathbf{I}_{e\chi})^{-1} \mathbf{C}_\chi$ ②	
4	发送参数①、②以及 $\mathbf{P}_{\chi e}$	接收参数①、②和 $\mathbf{P}_{\chi e}$
5		用各节点参数①、 $\mathbf{P}_{\chi e}$ 计算 $\mathbf{P}_\epsilon = (\mathbf{P}_e - \sum_{\chi \in \mathbf{X}} \mathbf{P}_{\chi e} (\mathbf{P}_\chi + \mathbf{I}_{e\chi})^{-1} \mathbf{P}_{e\chi})^{-1}$
6		用参数②、 $\mathbf{P}_{\chi e}$ 计算 $\mathbf{R} = \mathbf{P}_\epsilon \sum_{\chi \in \mathbf{X}} \mathbf{P}_{\chi e} (\mathbf{P}_\chi + \mathbf{I}_{e\chi})^{-1} \mathbf{C}_\chi$
7		计算 $\lambda_e = \mathbf{P}_\epsilon \mathbf{C}_e - \mathbf{R}$
8	接收 λ_e	下发 λ_e
9	计算 λ_χ	计算 $\lambda_A, \lambda_B, \dots, \lambda_\chi$
10	发送各节点 λ_χ , 接收总节点 λ_χ	发送总节点 λ_χ , 接收各节点 λ_χ
11	比对	比对

通过采用 3200 点(1000×3 内部点加 200 边缘点)、6300 点(2000×3

内部点加 300 边缘点)和 9600 点(3000×3 内部点加 600 边缘点)的数据进行测试,可以看到随着数据集合点数的增加,集中算法与分布式计算方法的差距越来越大,但二者变化趋势是一致的,见图 5-14。

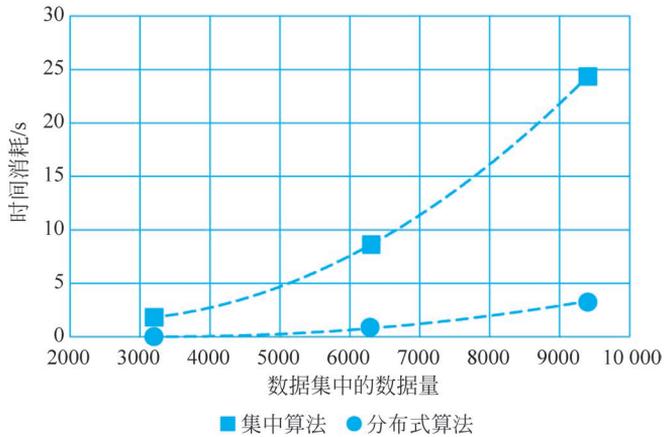


图 5-14 集中算法与分布式计算方法对比