

# 自然资源领域空间数据批量处理自动化技术探究

沙 漠

重庆市规划和自然资源信息中心 重庆两江 401147

**摘 要：**随着自然资源管理精细化的要求不断深入，空间数据类型日益多样、更新频次不断提高，传统以人工为主或半自动为辅的空间数据处理方式难以满足效率、稳定性和可复用性需求。本文聚焦于自然资源领域的空间数据批量处理自动化技术，提出一种基于ArcGIS Server地理处理服务（GP服务）、FME Desktop模板+脚本、以及开源任务调度工具Snail Job的整体架构。通过设计前端用户上传数据、填写参数、触发任务流程，形成从数据入库、格式转换、标准化、空间分析，到成果输出的批量自动化处理空间数据工具的应用。在研究中结合模块化脚本、模型构建、服务调用与调度机制，系统性探讨关键技术、系统架构与调度策略。结合实践表明，该方案显著提升处理效率、降低人工干预、增强系统可扩展性，为自然资源信息化建设提供了一种可推广路径。

**关键词：**自然资源；空间数据批量处理；自动化；GP服务；FME模板；任务调度

## 一、研究背景与意义

在自然资源管理领域，如：地籍调查、生态修复、规划编制成果等，往往涉及大规模、多时态、多区域的空间数据。数据格式繁多（矢量、栅格、CAD）、来源多样（系统生成、第三方提供）、空间参考、属性字段、精度标准不一，这对数据处理提出了很高要求。此外，随着“一张图”建设，空间数据更新频率变高、批量任务量变大。如果仍主要依赖人工或重复脚本操作，不仅耗时，而且易出错、难以保证标准化和一致性<sup>[1]</sup>。因此，构建一个可复用、可配置、服务化、自动化的空间数据批量处理体系具有显著意义：①提升数据处理效率，缩短周期；②降低人为操作错误，提高数据质量和数据规范性；③支持可扩展、大规模处理，满足自然资源行业的需求；④为用户侧提供易操作界面，减轻技术负担。

## 二、空间数据批量处理自动化需解决的关键技术

### （一）多源异构数据标准化能力

在自然资源领域，空间数据来自不同系统和部门：规划编制数据、审批发件数据、地籍数据、生态监测数据等。它们在数据格式（Shapefile、File GDB、CSV XY、PostGIS等）、空间参考系（WGS84、CGCS2000、多种投影）、字段结构（字段命名、类型、编码）、属性值标准（中文、英文、不同单位）等方面存在差异<sup>[2]</sup>。因此，自动化流程首先必须具备数据标准化能力。包括：①空间参考统一：预定义目标投影并自动重投影。②字段结构统一与映射：建立字段映射规则，如“行政区编码”统

一字段名、“面积”字段单位转换。③数据格式统一：将不同格式数据统一导入系统标准格式（如File GDB或数据库表）。④元数据与属性校核：自动检测缺失字段、异常值、数据类型不匹配。

### （二）参数化、模板化、可复用的处理流程

自动化批量处理要求流程具备高度的复用性和灵活性。具体包括：①模板设计：将常见的处理逻辑（如裁剪、投影转换、要素叠加、字段更新）设计为可重复使用的模板或模型。②参数化脚本/服务：流程入口允许外部传参（输入路径、输出路径、字段名、投影、结果图层名等），从而避免每次重新编写代码。③可配置工作流：用户上传文件+选择参数→调度系统调用模板/服务。④模块化结构：将流程拆分为多个功能模块（数据预处理、格式转换、空间分析、结果入库）以便组合不同任务场景<sup>[3]</sup>。

### （三）服务化处理与接口化调用

批量处理自动化体系中服务化是关键。具体要求：①将数据处理流程封装为服务（如通过GP服务、REST接口、脚本接口），使其可远程调用和调度。②接口必须支持参数输入、状态反馈、日志输出。③支持队列化、并发调用、调用结果的回传。例如使用ArcGIS Server发布GP服务，模型构建器或脚本调用该服务；也可设计脚本接口调用FME模板。

### （四）调度执行与资源管理

批量任务通常数据量大、运行周期长、资源占用高，

需要一个调度机制来合理管理、控制、监控。关键技术包括：①任务队列与优先级管理：保证任务按照预定规则执行、避免冲突。②并发、分布式执行能力：在服务器集群或多实例环境中，任务可并发执行，提高效率。③资源分配与限流：防止任务抢占所有资源导致系统崩溃<sup>[4]</sup>。④执行监控、日志管理、异常重试机制：任务失败或超时应自动处理或告警。⑤前端反馈机制：用户可查看任务状态、日志、耗时、结果链接。

### （五）前端交互与用户友好性

为了使用户能够操作自动化流程，而无需深度编程背景，系统提供友好的前端交互界面：①上传数据：用户上传文件、选择图层、指定参数。②参数填写与验证：用户填写必要参数（如字段名、投影、图层名称），系统进行输入验证。③任务提交与进度反馈：提交任务后可查看进度条、日志、完成状态。④结果下载/查看：提供结果数据链接。这样，自动化批量处理从“技术后台”封装为用户可操作的“前端入口”。

## 三、自动化处理流程设计—基于GP服务+FME模板+调度工具

我们以实际项目为例详细说明系统设计：使用FME Desktop构建的模板+外部脚本传参，使用开源调度工具Snail Job作为触发机制，再调用ArcGIS Server发布的GP服务或模型构建器模型服务，全流程自动化运行。

### （一）系统总体架构

系统采用“三层架构”：①前端层：用户通过网页上传数据、填写参数、提交任务。②调度层：由Snail Job开源调度工具负责监控任务、触发执行、状态管理。③服务层：包括FME Desktop模板运行脚本、ArcGIS Server发布的GP服务或通过模型构建器封装流程、数据库入库模块。

流程如下：①用户上传空间数据（如：Shapefile、CSV XY、GDB）并填写参数（例如：目标GDB路径、字段映射、投影目标、输出图层名等）<sup>[5]</sup>。②前端将任务信息写入任务队列或任务表。③调度系统（Snail Job）检测到待执行任务，按照预设规则触发任务执行。④调度模块调用脚本：首先运行FME模板，传入用户参数，完成格式转换、标准化、初步预处理然后调用ArcGIS Server上的GP服务/模型服务，执行空间分析、字段更新、要素入库等操作。⑤完成后，结果反馈至前端：日志、耗时、成功/失败状态、输出链接。⑥若任务失败或出现异常，调度系统触发告警或自动重试机制。

### （二）FME模板+脚本设计

使用FME Desktop构建一个或多个模板（Workspace），每个模板负责一个或多个常见的空间数据预处理任务，例如：坐标转换（从XY文本、CSV、Shapefile等转换为指定投影）；批量裁剪（按行政区划、预设边界）；字段标准化与映射（如重命名字段、类型转换、单位变换）；数据检测与清理（如空属性检测、重复几何剔除）。为了实现自动化，编写了一个脚本，该脚本接收参数，然后调用FME Desktop命令行模式执行对应Workspace。模板运行结束后，将输出文件或库交给下游GP服务处理。这种“模板+参数脚本”的方式有效提高了流程的可配置性与可复用性，并减少了手动操作<sup>[6]</sup>。

### （三）GP服务/模型构建器模型服务设计

使用模型构建器（ModelBuilder）将常用的空间分析流程（如图层叠加、字段匹配更新、合并统计等）封装为模型。然后将模型发布至ArcGIS Server为GP服务。GP服务具备以下特性：远程调用支持（REST接口）；可定义输入参数（例如目标图层、字段名、图层路径、输出路径）；支持脚本化调用（如通过arcpy调用、HTTP请求调用）；可配置为批量运行实例，以支持并发任务。在案例中，例如“目标图层与a图层、b图层相交，并在目标图层的指定字段中记录a、b图层值”的流程可封装为一个GP服务，用户上传图层+参数（目标图层、a图层、b图层、字段名等），调度系统触发后直接调用该GP服务执行。

### （四）调度系统—Snail Job的应用

在系统中使用了开源任务调度工具Snail Job来管理任务。其主要特点包括：①支持任务分区、广播、静态分片、DAG工作流、实时日志监控。②提供Web界面管理任务，如任务创建、调度、执行日志、重试管理。③可与Java环境深度整合（Spring Boot等）但也可以用于跨平台调度<sup>[7]</sup>。在系统中，前端提交的任务会写入任务队列表，Snail Job作为中心调度平台检测新任务、分配执行流程、触发脚本调用（FME模板）或调用GP服务。

### （五）整体流程梳理与关键技术细节

在整体流程的梳理及关键技术的应用中，涉及到用户界面、任务入队、调度触发、预处理阶段及核心分析等多个环节，以用户界面的实践操作为例，其通过用户登录平台→上传数据（文件或目录）→选择任务类型（如图层入库、叠加分析、字段更新等）→填写参数（输入路径/输出路径、目标GDB、字段映射、投影信

息),直至完成登录。而核心分析即为GP服务,具体实践以脚本或调度系统调用ArcGIS Server上的GP服务,传入中间结果路径及分析参数。服务执行空间叠加、字段更新、统计汇总、成果入库至目标GDB/数据库。过程中输出日志信息。

#### 四、讨论与结语

##### (一) 方法优势与适用性

本研究设计并实现的自动化批量处理体系具有以下优势:①高效性:从上传到结果输出流程高度自动化,减少人工干预,处理速度大幅提升<sup>[8]</sup>。②复用性强:借助FME模板+参数脚本、GP服务模板,处理流程模块化、可复用、适应不同任务场景。③服务化与接口化:通过GP服务和脚本接口,使处理流程可集成到更大系统或前端平台中<sup>[9]</sup>。④用户友好:提供前端操作界面,降低用户技术门槛,支持上传数据、填写参数、查看状态、获取结果。⑤可扩展性:调度系统支持并发、多任务队列、资源管理,可应对大规模、多批次数据处理需求。

##### (二) 结语

本文从自然资源领域空间数据批量处理的角度出发,提出了一种基于GP服务、FME模板+脚本、任务调度系统(SnailJob)构建的自动化处理体系。通过设计可配置模板、服务化流程、调度机制与用户交互界面,实现了数据上传、格式转换、空间分析、结果输出等批量环节的自动衔接。实践表明,该体系在提升效率、降低人工、增强可复用性、支持大规模任务方面具有显著优势。

#### 参考文献

- [1]肖孔力.基于FME的空间信息数据处理技术研究[J].现代信息技术,2025,9(15):163-167.
- [2]张琦,玉捷.基于GIS的日常变更自动化分析方法及其在城市管理中的应用[J].城市建设理论研究(电子版),2025,(08):172-174.
- [3]黄华东,马宏阳,樊川.地理空间数据服务实时发布系统的设计与实现[J].现代测绘,2024,47(01):75-80.
- [4]李桂城,蔡忠亮,李伯钊,等.新型专题地图制图自动化系统框架与关键技术[J].测绘科学,2024,49(01):181-189.
- [5]张雯雯,缪爱兰,俞晓路.自然资源“一张图”数据处理方法研究[J].测绘标准化,2023,39(03):32-36.
- [6]任恒一,史文中.基于工作流的空间数据质量检查方法[J].测绘地理信息,2023,48(04):137-140.
- [7]林澍哲,梁梅芳,谢萍,等.自然资源资产管理信息系统设计与实现[J].国土资源信息化,2021,(01):43-48.
- [8]梁哲恒,吴宝佑,谢刚生.利用分布式自动化并行处理框架提升空间数据分析效率[J].测绘通报,2020,(10):143-147.
- [9]王晓龙.基于Python的空间矢量数据与超链接文件一致性的检查和自动化处理[J].北京测绘,2020,34(05):724-727.