

# 基于 FME 的“大数据”分析与应用

## 1. 引言

此次数据分析的目的针对提取的监测图斑围绕永久基本农田、城市开发边界、自然保护区、备案审批数据等进行监测分析，获取土地资源全方位监测、监管的数据。

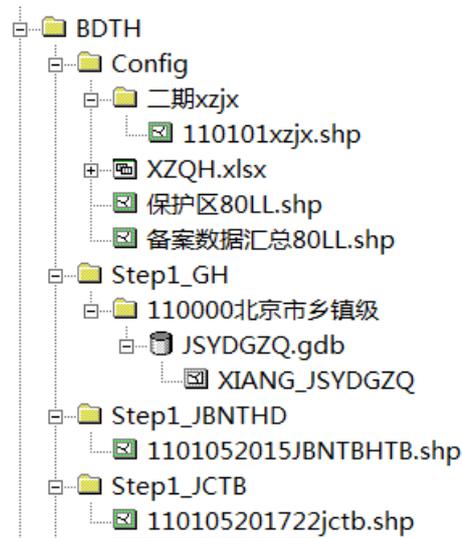
数据总量为 200 个区县，各类数据汇总约 5000 多万条记录，合 60 个 GB。以前的分析方式是逐县、逐类去进行分析，既耗时又耗力。一旦要求有变更，返工工作量相当巨大，很难及时响应甲方的要求。

通过对各类数据与成果要求分析，利用 FME 对工作流程与工作思路做出优化，提高工作效率与成果质量，快速响应甲方的要求。

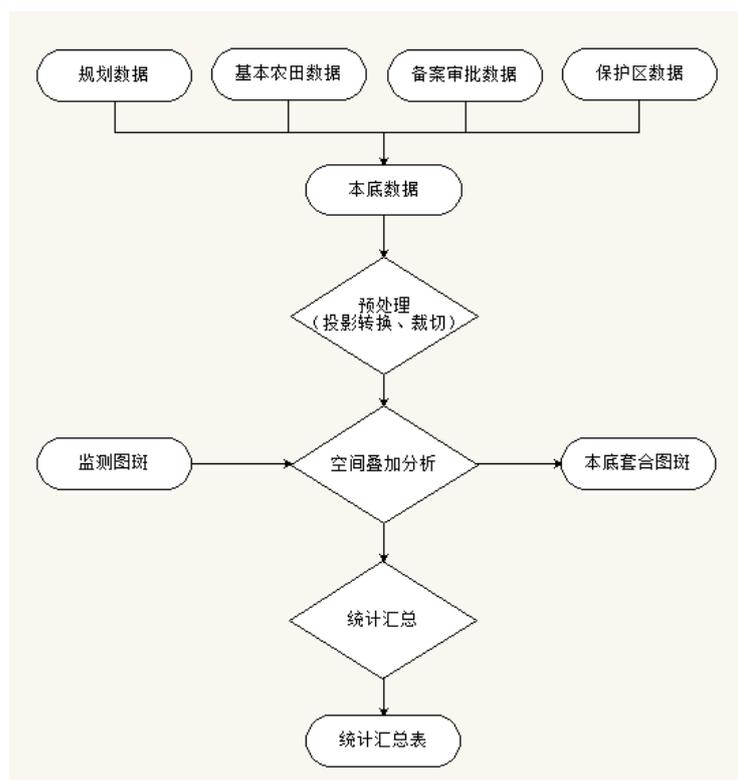
## 2. 本底数据分析

监测成果数据：监测图斑（SHP 数据、80 大地坐标）

- 1) 永久基本农田数据（SHP 数据库、80 大地坐标）
- 2) 备案审批增量数据（SHP 数据、80、54 球面坐标）
- 3) 自然保护区数据（SHP 数据、80 球面坐标）
- 4) 规划数据库（GDB 数据库、80 球面坐标）
- 5) 行政界线（SHP 数据、80 大地坐标）



## 3. 思路与流程



(1) 监测图斑与永久基本农田分析。监测图斑与基本农田分析时需要统计所占基本农田图斑面积与基本农田面积。计算永久基本农田图斑中基本农田系数，即用基本农田面积（JBNTMJ）/基本农田图斑面积（TBMJ）。将永久基本农田数据与监测图斑做相交分析，相交面积记录到监测图斑的 FTBMJ 字段中，用 FTBMJ\*刚才算出来的系数，得到图斑所占基本农田面积，记录到监测图斑的 FMJ 字段中。

(2) 监测图斑与备案审批数据分析。将备案审批数据与监测图斑做擦除分析，擦除后的结果为违法图斑部分，违法图斑面积记录到监测图斑的 NMJ 字段中。

(3) 监测图斑与自然保护区数据分析。保护区数据分析分为合法占用与违法占用两类，因此，需要将上一步分析的图斑违法部分与合法部分提取出来，分别于自然保护区数据进行相交分析。将图斑合法占用保护区的面积记录到监测图斑 BHQYMJ 字段中去，将图斑违法占用保护区的面积记录到监测图斑 BHQNMJ 字段中，将占用保护区的保护区名称记录到 BHQMC 字段中，如果图斑同时占用两个保护区，选取占用面积较大的保护区名称做记录。

(4) 监测图斑与规划数据分析。将规划数据进行过滤，筛选出 GHLXDM 为“030”余“040”的图斑，与监测图斑进行相交分析。相交面积记录到监测图斑的 CMJ 字段中。

(5) 以上分析结果中，FTBMJ、FMJ、NMJ、BHQYMJ、BHQNMJ、CMJ 最小统计面积为 1 亩，不足统计面积的记录为 0。

以上相交分析中会存在监测图斑与多个图斑相交的情况，统计相交面积时，需要求和计算，特别是与多个基本农田图斑相交的，需要分别计算相交图斑的系数再求和。

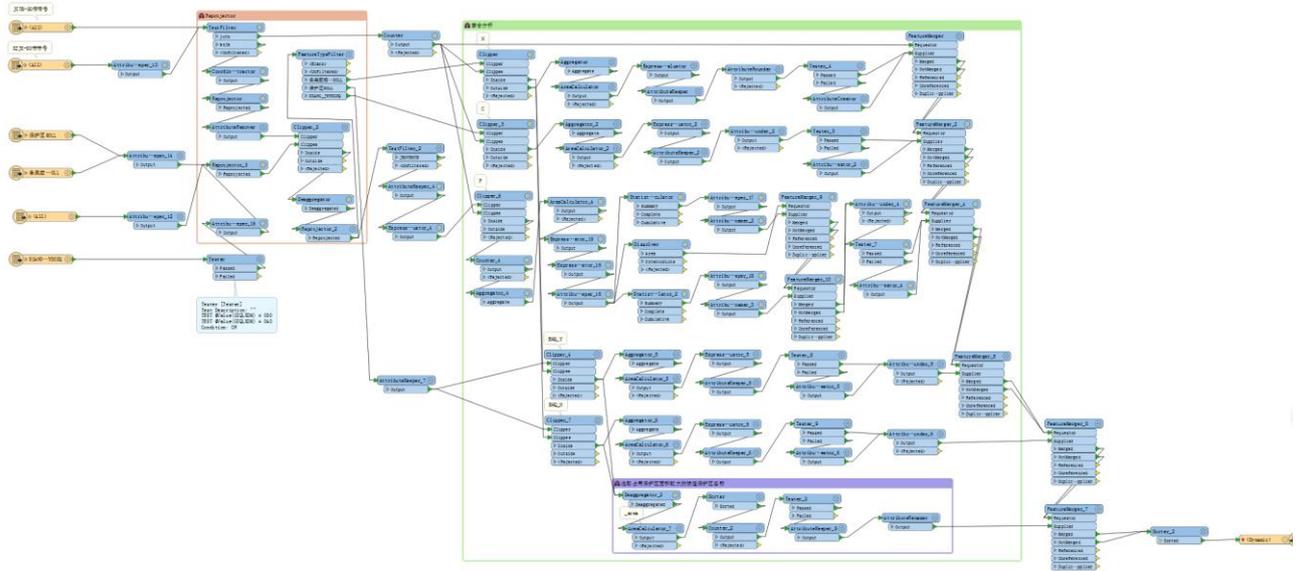


图 1 空间分析

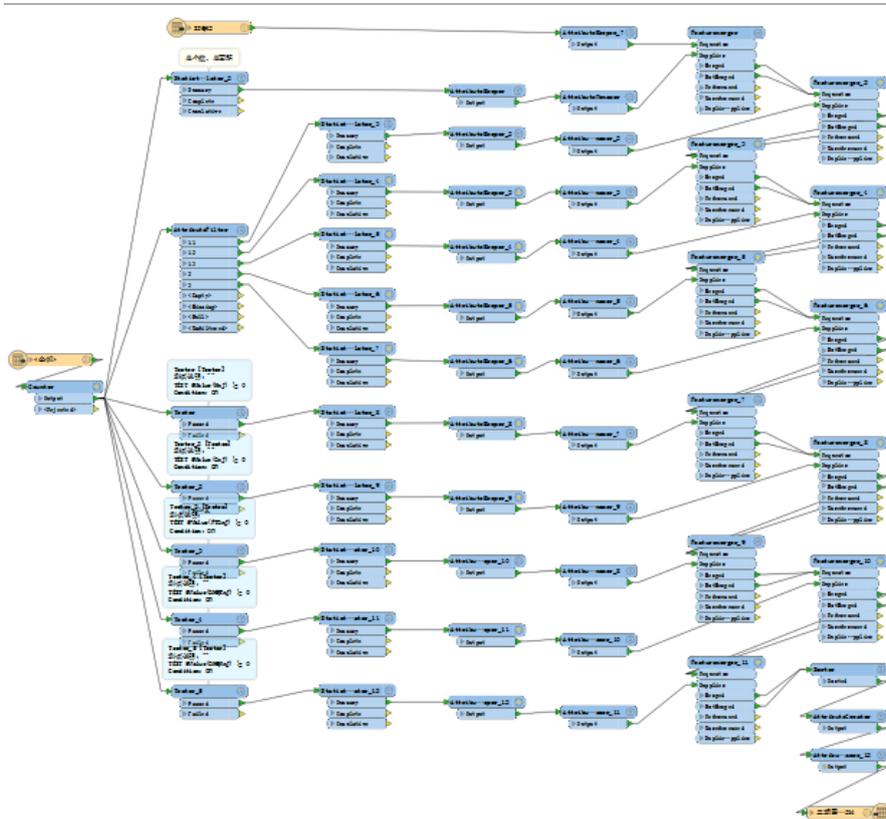


图 2 统计汇总

#### 4. 经验总结

FME 成了日常工作中的首选工具。通过此次用 FME 的优化，对甲方的要求作出了积极高效的响应，同时得到了许多惊喜，也对 FME 的局限性作出了一些调整。

- (1) 时间短效率高。此次统计分析只需要一台电脑运行了 13 个小时，数十倍于传统的工作模式。
- (2) 正确率高。人为干预较少，数据成果正确率 100%。
- (3) 一体化的处理模式，一键执行即可，不用过多操作。（前期需对模版做反复测试，选取数据要有典型性）
- (4) 大事化小，小事化了。对各类问题进行拆解，拆解后逐个击破。
- (5) FME 在大批量空间数据叠加时效率不是很高。这一点在将数据按省进行拆解处理后得到了很大的提升。
- (6) FME 处理数据时总是读取全部数据后才执行后续操作的。前期有效的优化数据内容与质量会有显著的提升。
- (7) 后续可以利用 FME Server 进行进一步的优化。