基于 FME 的 DOM 元数据生产技术研究

—以“三调”高分二号的 1:1000DOM 成果为例

湖南省第二测绘院 尹梁波

0 引言

在数字化测绘生产日益发展的今天，基础地理信息作为各种信息的空间基础支持信息，在全球范围内受到了越来越多的关注，社会各领域对空间信息的需求也越来越大。元数据作为基础测绘成果的重要阐释性文档，从多角度出发，对测绘成果的平面基准、高程基准、投影方式等信息做出了详细的说明，在基础数据的分析、一图多用等方面具有十分重要的现实意义。然而，元数据的生产是一个非常繁琐的过程，其涉及不同类型数据信息的录入，工作量庞大。仅仅依靠人工生产，成本高、周期长、效率低。事实上，作为最终测绘成果的附件，元数据质量的好坏直接影响到整批成果的合格与否。

FME 作为一个用于空间数据提取、转换、处理的强大工具，提供了多源数据处理和应用的解决方案，用户可以根据项目实际需求建立具有特色的空间数据处理模型，实现数据之间的转换、处理。为进一步提高元数据生产的作业效率，改善目前人工录入数据属性所存在的不足，缩短作业周期，提高作业效率和生产质量，本文基于 FME 对“三调”项目中 1:1000DOM 成果元数据的生产展开研究探讨。

1 元数据生产的要点分析

元数据的内容随着其描述的主题内容、应用范围不同而不同。以某省第三次全国土地调查项目为例，基于高分二号遥感影像的 1:10000DOM 成果的元数据内容主要包括：成果数据基本信息(MBII)、数据源情况(MPID)、数据生产过程信息(MPPI)、一级质量检查情况(MQC1)、二级质量检查情况(MQC2)、成果验收情况(MDAC)等方面 6 个图层。根据省《第三次全国土地调查正射影像生产技术规程》的相关要求，元数据成果按县域分幅，即一个县域分幅正射影像对应一个元数据文件，以 MDB 格式存储，数据文件本身采用 CGCS2000 国家大地坐标系，投影信息与正射影像成果保持一致，命名与县域正摄影像成果存储文件夹相同。

根据项目技术规程的要求，数据生产过程信息(MPPI)、一级质量检查情况(MQC1)、二级质量检查情况(MQC2)、成果验收情况(MDAC)等四个图层的属性信息涉及的数据类型相对单一，可直接在创建的模板上录入相应属性信息。而成果数据基本信息(MBII)图层涉及的元数据文件名、县级行政辖区名称、县级行政辖区名称、数据量大小、中央子午线、等字段需要视实际情况的而录入，数据源情况(MPID)图层涉及的对应影像文件名、景号、数据源时间、侧视角、面积等字段信息依赖于高分二号遥感影像元数据 XML 文件中的相关属性信息。故而，基于高分二号的 1:10000DOM 成果元数据生产的要点主要有：

1） 创建元数据 MDB 文件模板；

2） 根据县域 SHP 矢量数据创建成果数据基本信息 MBII；

3） 根据高分二号遥感影像 XML 元数据及镶嵌 SHP 数据补充数据源情况图层信息 MPID；

根据上述要点和相关规程的要求，基于高分二号的 1:10000DOM 成果元数据生产所涉及的关键数据主要有如下几项：

1） 县域 SHP 矢量数据（XZQDM、XMC、经度等）；

主要包含县级行政区辖区代码、县级行政辖区名称、县级行政辖区名称拼音缩写、以及县级行政区的经纬度等字段。

2） MDB 数据（MBII、MPID、MPPI、MQC1、MQC2、MDAC 等）；

3） 镶嵌面 SHP 矢量数据（ID 等）；

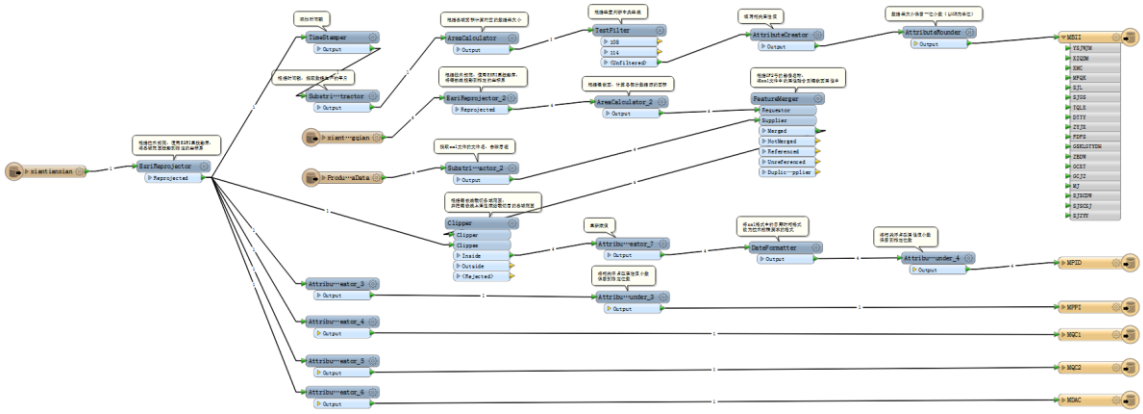
主要包含在实际生产过程中，县域 DOM 成果数据所采用的高分二号源数据。不同镶嵌面要素的 ID 对应生产采用的那一景卫星数据文件名；该文件范围覆盖全省县域 DOM 成果数据镶嵌面

4） 高分二号 XML 元数据

主要包含了源高分二号卫星影像的相关信息；

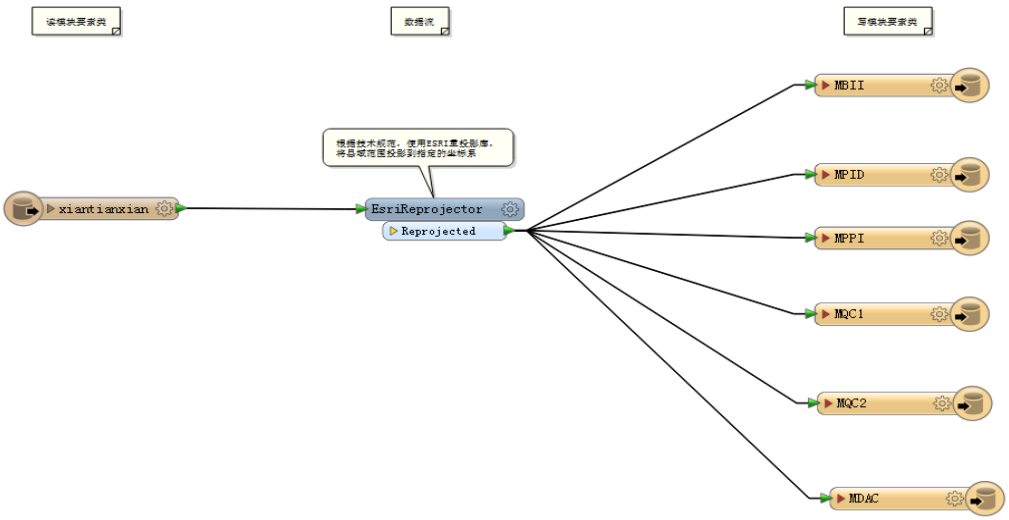
2 基于 FME 的元数据生产流程

基于 FME 的元数据生产流程主要包括：元数据 MDB 文件模板的制作、成果数据基本信息(MBII)图层的制作、数据源情况(MPID)图层的制作。具体生产模板流程图如下所示：



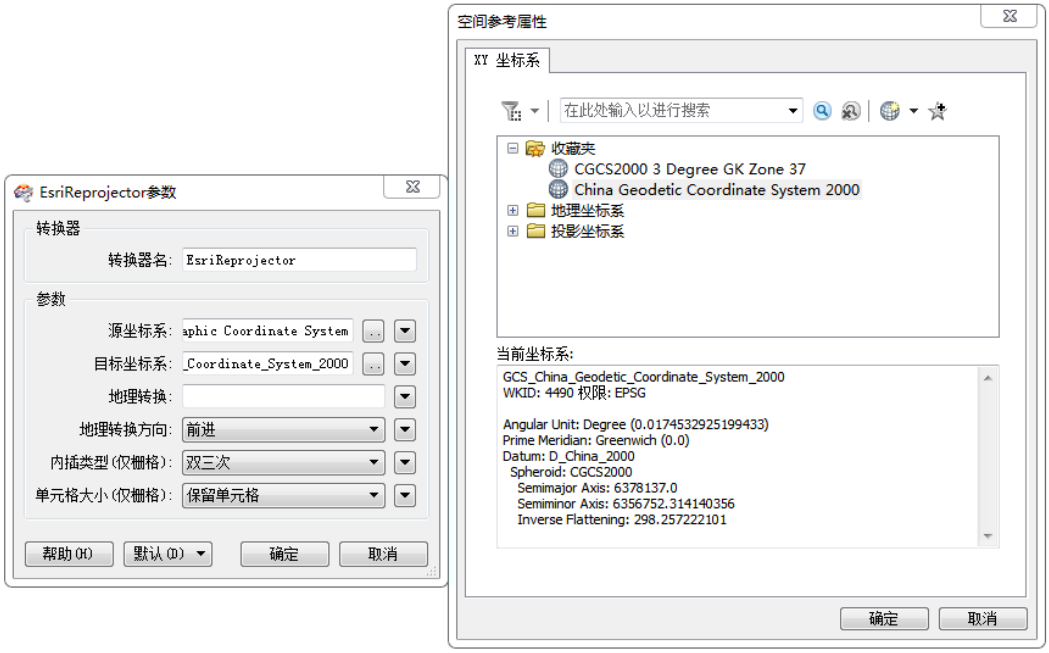
2.1 元数据 MDB 文件模板的制作

在 FME Workbench 中创建工作空间，读模块格式选择 Esri Shape，添加县域 SHP 文件，写模块格式选择 Esri Geodatabase (Personal Geodb)，命名 MDB 数据集名称 MDB。根据省《第三次全国土地调查正射影像生产技术规程》中对各个图层中字段的要求，分别创建 MBII、MDAC、MPID、MPPI、MQC1、MQC2 这六个要素类：



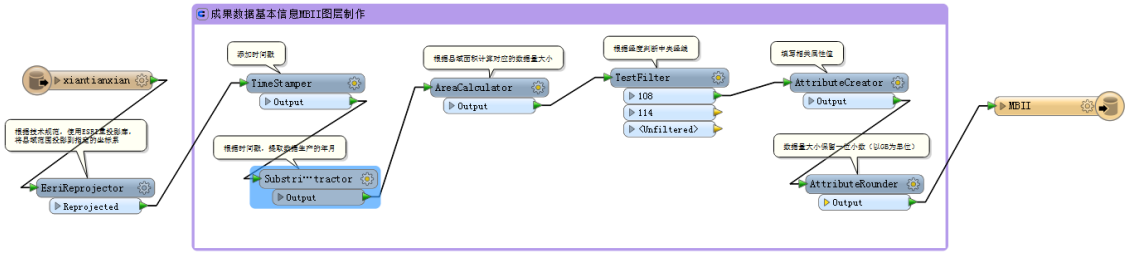
涉及的转换器及其功能介绍：

EsriReprojector 转换器：使用 ESRI 重投影库，把县域 SHP 要素坐标重投影到技术规程规定的坐标系统。该转换器输入的 SHP 要素上关于坐标系统的任何设置都会被忽略。对于不熟悉 FME 坐标系统库而习惯了 Arcgis 坐标系统的用户来说，该转换器显得非常方便。具体的参数设置如下：



2.2 成果数据基本信息 MBII 图层的制作

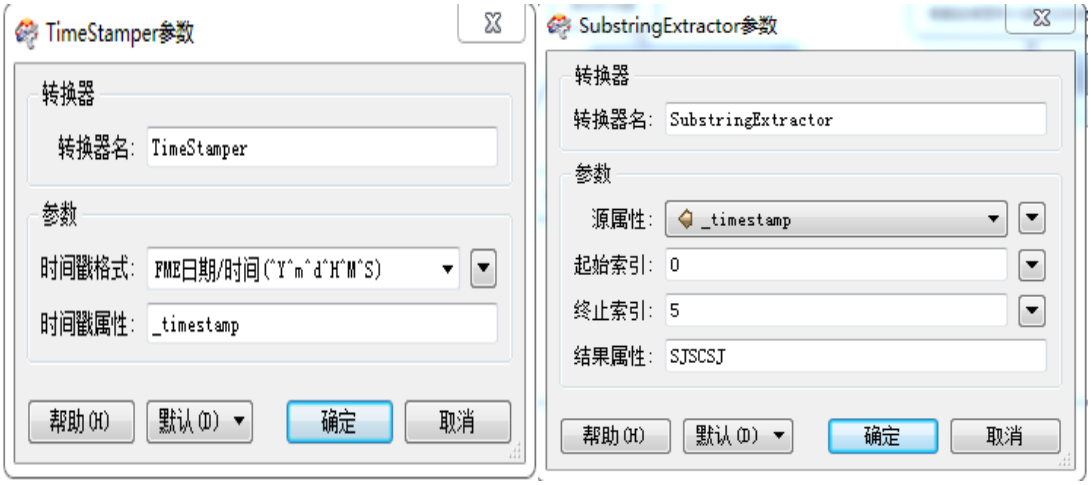
根据制定好的元数据 MDB 模板文件和县域 SHP 文件，添加 MBII 要素类的属性内容，制作出成果数据基本信息图层。其中，县域 SHP 矢量数据包含县域的行政辖区代码、名称及其拼音缩写、经纬度等字段。MBII 图层的关键字段主要包括数据量大小 SJL、中央子午线 ZYJX、数据生产时间 SJSCSJ 等。具体的生产模板流程图如下：



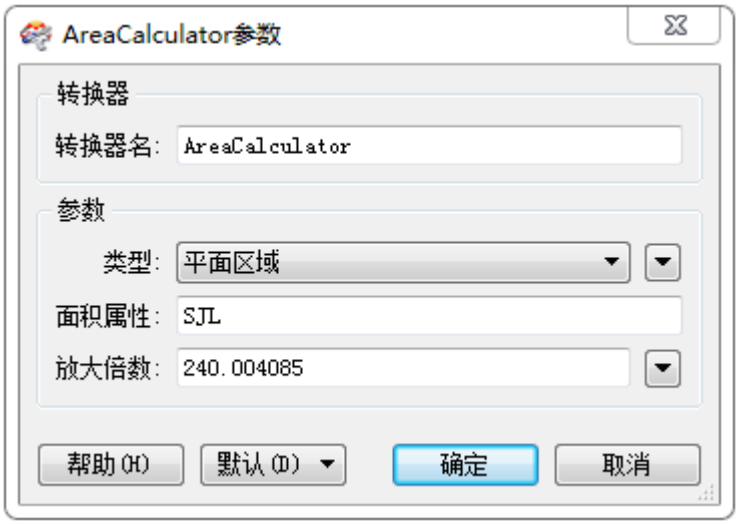
涉及的转换器及其功能介绍如下：

1） TimeStamper 转换器和 SubstringExtractor 转换器：利用 TimeStamper 转换器给县域 SHP 矢量要素添加一个时间戳为新属性，时间戳格式为年月日时分秒，再通过 SubstringExtractor

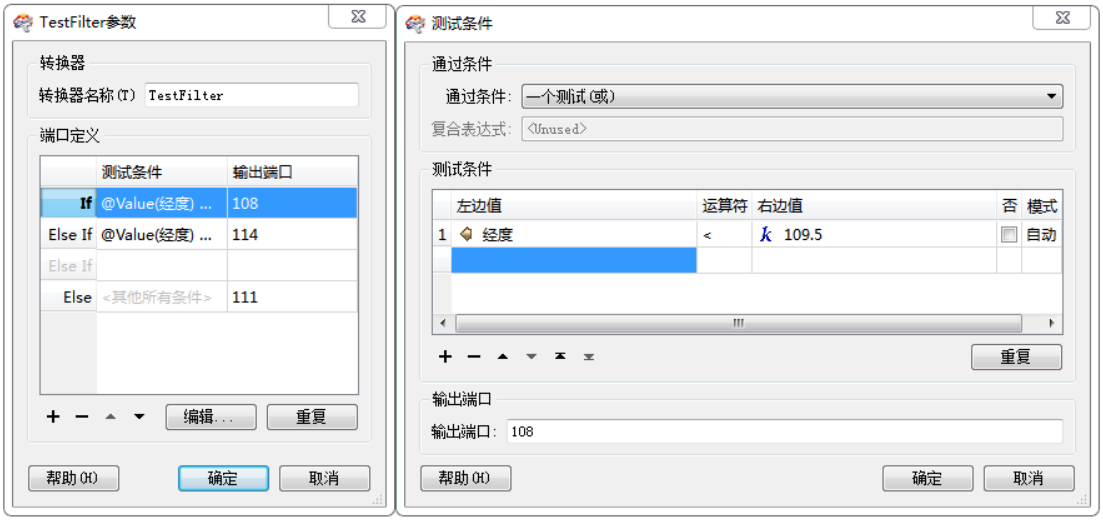
转换器提取出时间戳属性中的年月，最终结果属性为数据生产时间 SJSCSJ。具体的参数属性设置界面如下：

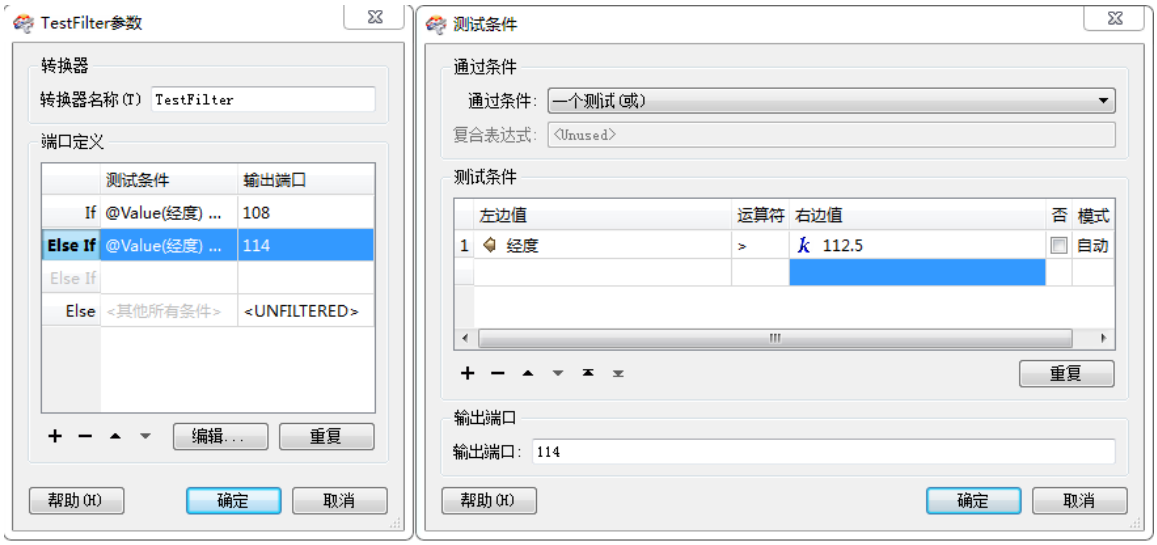


2） AreaCalculator 转换器：计算当前县域的 SHP 矢量数据的面积。基于高分二号的 1:1000DOM 成果数据在按县域出图时，所设置的重采样方法、采样间距等参数一致，故县域面积和实际成果数据量大小存在固定比例系数。将 AreaCalculator 转换器中的方法倍数设置成其固定的比例系数，最终得出的面积属性即为数据量大小。具体的参数设置界面如下：

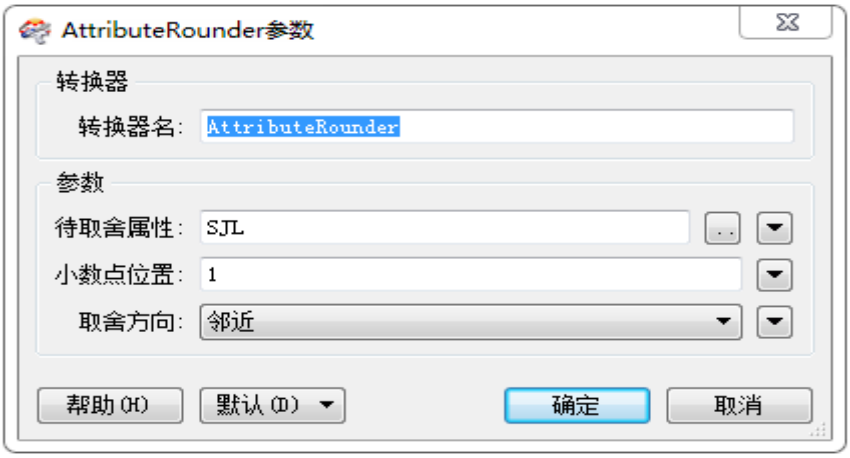


3） TestFilter 转换器：根据县域 SHP 矢量数据中的经度字段属性，设置测试条件。根据本省实际情况，结合省域界线，经度在 109.5-112.5 之间的中央经线为 111，经度为 109.5 及其以下的中央经线为 108，经度在 112.5 及其以上的，中央经线为 114。由此便可判断出县域所处的中央子午线。具体的参数设置界面如下:





4） AttributeCreator 转换器和 AttributeRounder 转换器：通过 AttributeCreator 转换器创建 MBII 要素的相关属性，并对属性设置常量、属性值和表达式，最后通过 AttributeRounder 转换器将 MBII 要素中的 SJL 属性近似到指定的一位小数位数。相关参数属性设置界面如下：



AttributeRounder

将一个属性近似到指定的小数位数。

输入端

Input

输出端

Output

参数

要近似处理的属性

小数点位置

取舍方向

该参数用来指定进行什么样的近似操作：

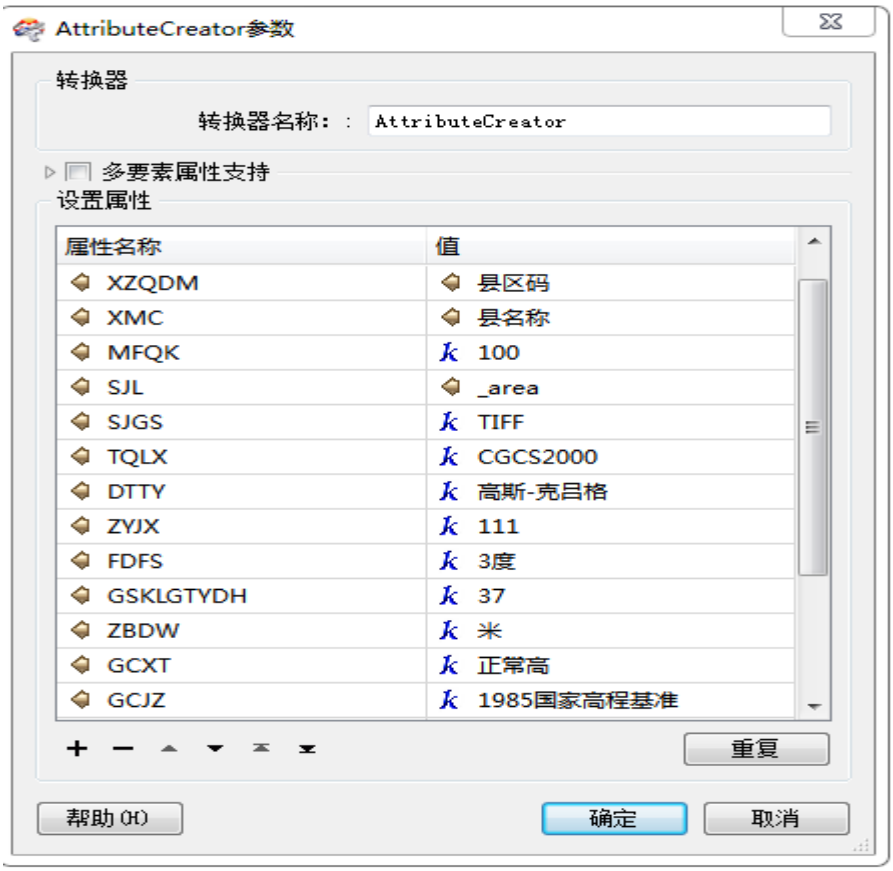
- nearest:数值四舍五入到最接近的值；
- up:数值总是向上取值；
- down:数值总是向下舍值。

编辑转换器参数

关于转换器参数选项

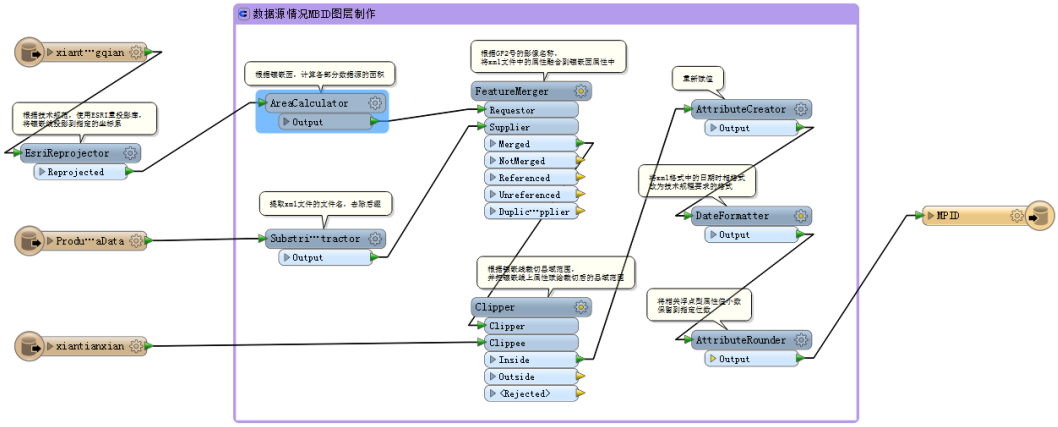
转换器类别

Calculators

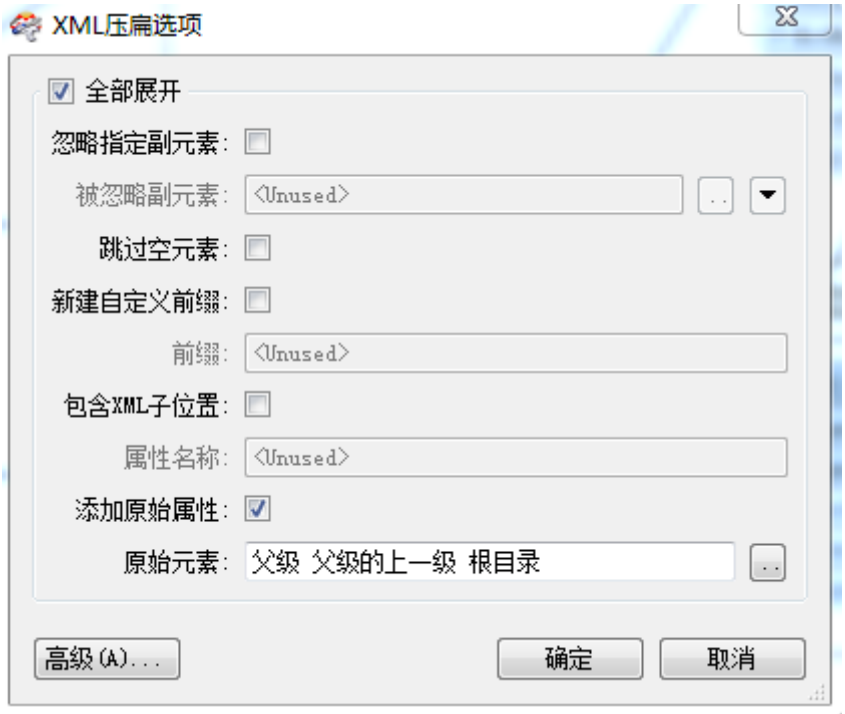


2.3 数据源情况 MPID 图层的制作

MPID 要素的制作涉及元数据 MDB 模板文件、县域 SHP 矢量文件、DOM 镶嵌矢量面文件、高分二号元数据 XML 文件等不同类型文件。其具体的生产流程如下图所示：



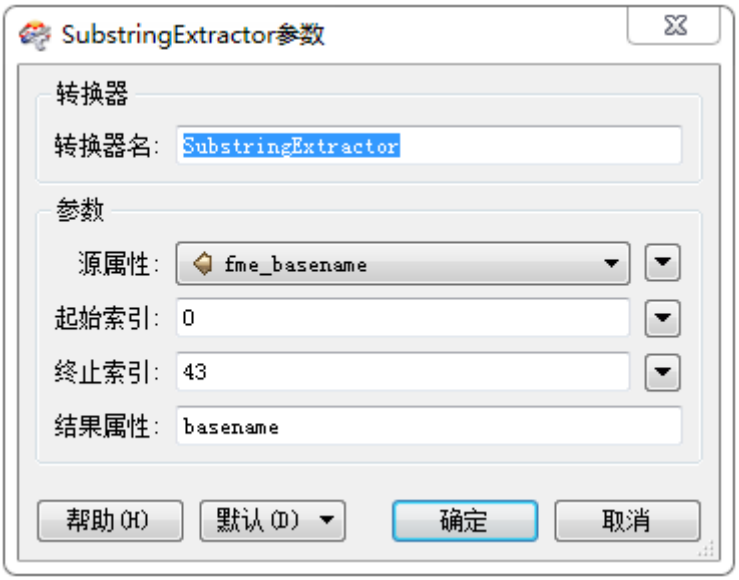
其中，镶嵌面矢量文件包含了高分二号遥感影像名称字段；添加高分二号元数据 XML 文件读模块时，匹配元素设置为 ProductMetaData，添加原始属性以便获取高分二号文件名，源文件为指定路径下所有 XML 后缀的文件。具体的参数设置如下：



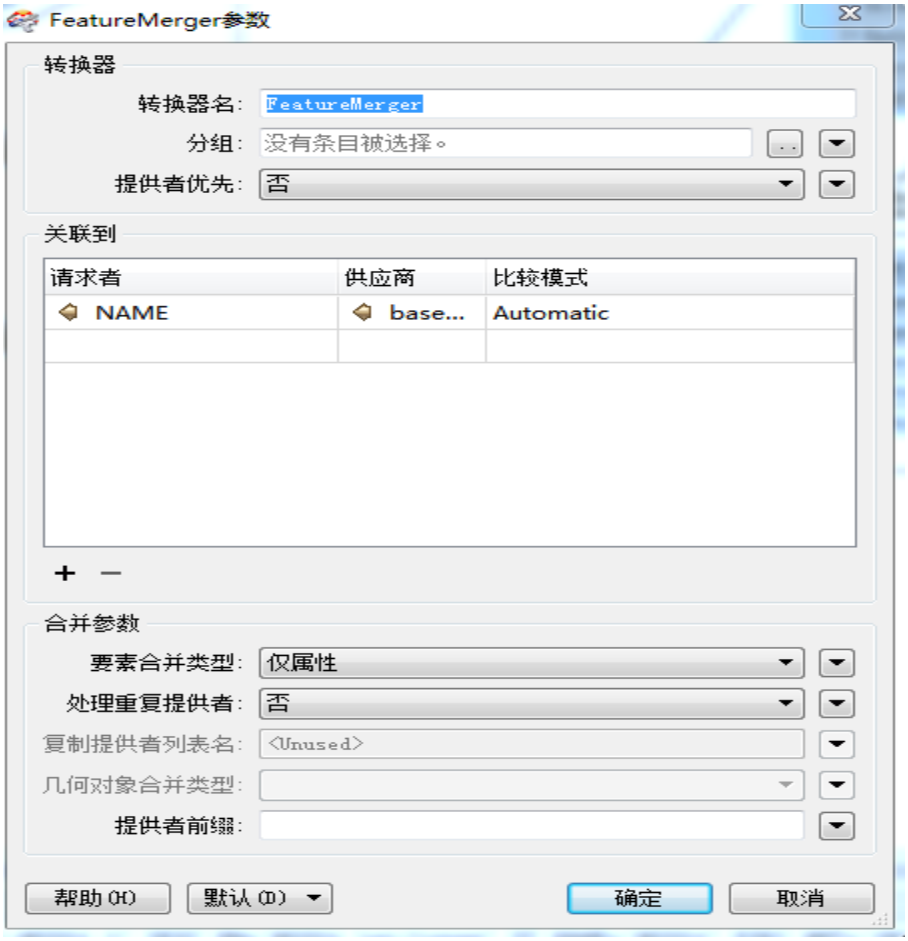
涉及的转换器及其功能介绍如下：

1） AreaCalculator 转换器：计算出各个镶嵌面的面积

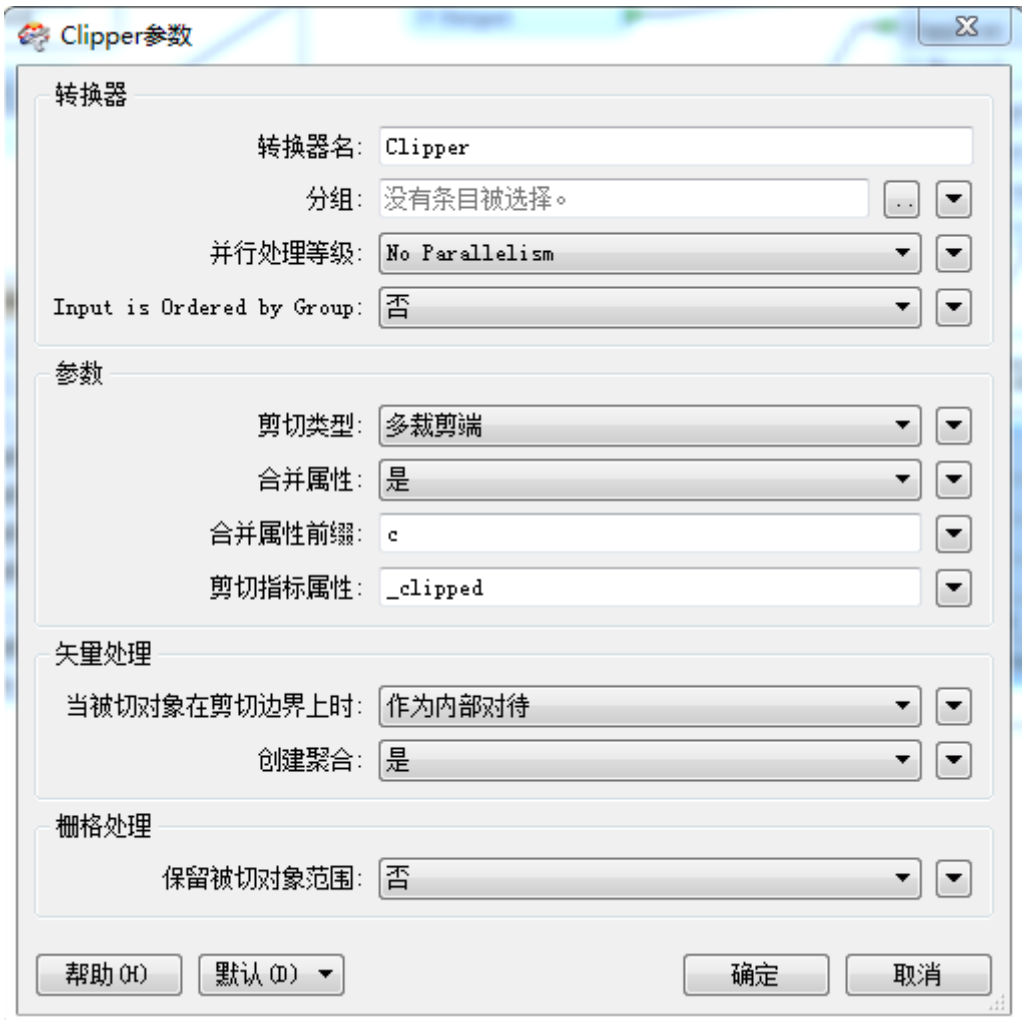
2） SubstringExtractor 转换器：根据源属性 fme_basename，按指定索引范围提取出不带后缀的 XML 文件名称 。相关参数设置界面如下:



3） FeatureMerger 转换器：根据高分二号遥感影像源数据名称，将 XML 文件中的属性合并到镶嵌面 SHP 文件属性中去。具体的参数属性设置如下：



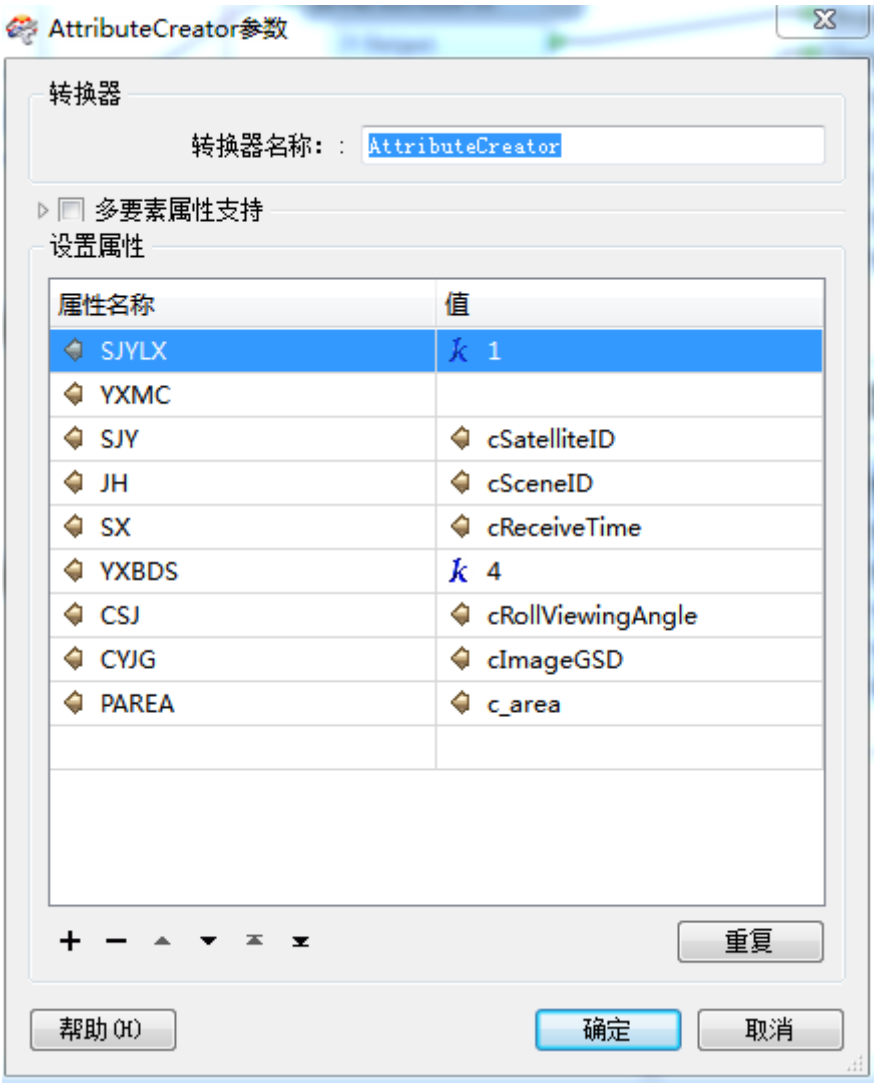
4） Clipper 转换器：以县域 SHP 矢量文件为裁剪边界，镶嵌面 SHP 矢量数据为被裁剪要素，合并属性，最终输出的 Inside 端口即为数据源情况 MBID 要素的矢量范围。具体的参数属性设置如下：



5） DateFormatter 转换器：由于高分二号元数据 XML 文件的日期格式为：年-月-日 ，根据相关技术规程要求，需将其重新格式化到指定的日期格式。具体的参数设置如下：

6） AttributeCreator 转换器和 AttributeRounder 转换器：通过 AttributeCreator 转换器重新命名裁切后所得到的属性值，与写模块中 MPID 要素中属性值一一对应，最后通过 AttributeRounder 转换器，将侧视角 CSJ、采样间隔 CYJG、面积 PAREA 等字段属性按四

舍五入法近似到指定的两位小数位数。相关参数设置如下：



3 结束语

本文基于 FME 软件对三调项目中 1:10000DOM 成果的元数据进行了研究，主要针对元数据 MDB 模板的制作、成果数据基本信息 MBII 图层和数据源情况 MPPI 图层展开了探讨，通过

一些简单的转换器，实现了不同数据类型间属性的转换和提取等操作，从以往那些冗余繁琐的重复劳动中解脱了出来，有更多时间去构思整个技术流程，大大提高了实际生产效率。

然而，由于本人初涉 FME，技术有限，对各类转换器的功能参数尚不是很了解，使用起来不够熟练，上述基于 FME 的元数据生产模板尚有诸多地方需要优化，如一级、二级检查意见问题及处理等属性，需要人机交互才能达到最理想的效果；实际 DOM 成果数据量大小的精确计算；数据生产过程信息中关于检查点个数和误差的描述信息的准确录入等等。故此，笔者认为，基于第三方开发软件对 FME 进行可视化人机交互窗口的开发，能更好的发挥 FME 在元数据生产过程中的作用，如基于 FME Plugin 的开发、基于 FME Objects 组件的开发等技术，都将是下一步需要开展的工作。以上就是笔者在利用 FME 软件生产 1:1000DOM 元数据的一些想法，希望各位专家批评指正！