



中國地質大學
China University of Geosciences

艰苦朴素 求真务实
温家宝

map2loop 1.0

map2model 1.0

li



1 前言

map2model和map2loop项目旨在：改进3D地质建模部分 workflow，主要是从地质图数据到建模的转换，这是**最耗时的部分**（hours to days），属于预建模过程。转换不是唯一的，与选择使用何种地质特征参数以及联合源数据集的方法有关。

map2model和map2loop加速地质建模，特别是考虑到

- Speed
- Reproducibility
- separation of data, concepts
- interpretations.



1 前言

Jessell et al. (2014)将3D地质建模分为4种场景：

1. local-scale (mine) models
2. regional-scale sedimentary basins (HyVR)
3. regional-scale hard-rock terranes (GemPy)
4. large-scale (crustal or lithospheric) models

LoopStructural着眼于hard-rock区域建模场景，相比矿山和沉积盆地来说，**data-poor**。地下的3D地质的最佳预测因子是包含在地质图中的信息以及**测井(logged well)数据**（如果可获取的话）。不幸的是，除了沉积盆地外，钻井(drill hole)都**太浅**，不能提供在区域尺度上的约束，通常还**缺少层序信息**。



1 前言

地质图提供：

1. 主要几何数据：如层序接触位置
2. 次级几何数据：局部地质层组(formation)厚度
3. 断层和层序拓扑信息

map2model为GemPy和LoopStructural提供完整的输入文件。

输入地质建模的数据源还有卫星数据(Wellmann et al., 2019)以及从钻孔直接自动提取的信息(dh2loop, Joshi et al., 2021)，但不在本文讨论范围内。

map2loop是依赖于map2model，但独立开发的，依赖于外部开源库有：Geopandas(管理矢量格式的地理空间数据)，Rasterio(管理栅格格式的地理空间数据)，Networkx(管理网络图)和Shapely(管理2D计算几何)。

Wellmann, J. F., Schaaf, A., de la Varga, M., and von Hagke, C.: From Google Earth to 3D Geology Problem 2: Seeing Below the Surface of the Digital Earth, Developments in Structural Geology and Tectonics, 5, 189-204, 2019.



2 输入数据

input指对map2loop和map2model的输入。augmented data是map2loop的产品。 augmented data是3d地质建模的输入。所有与map2model库相关的输入和输出都封装在map2loop库中。地质图的信息分为3类几何数据：

1. 位置相关数据，如：断层、侵入体与层序接触的位置
2. 梯度数据，如：接触或断层的倾角
3. 时空拓扑数据，如：断层之间以及层序单元之间的年代关系

map2loop导入不同的数据集（手工非常费劲），输出一些标准格式文件(GML), csv, geotiff, ESRI shapefile



2 输入数据

一旦输入参数定义了，整个 workflow 是自动化的，因此需要实现决定选择哪些参数（表1为参数列表）。整个 workflow 如图2。一旦生成配置文件以及在map2loop控制脚本中定义 workflow 控制参数，所有接下来的操作都是自动化的，从访问输入数据到使用GemPy或LoopStructural构建3D模型。

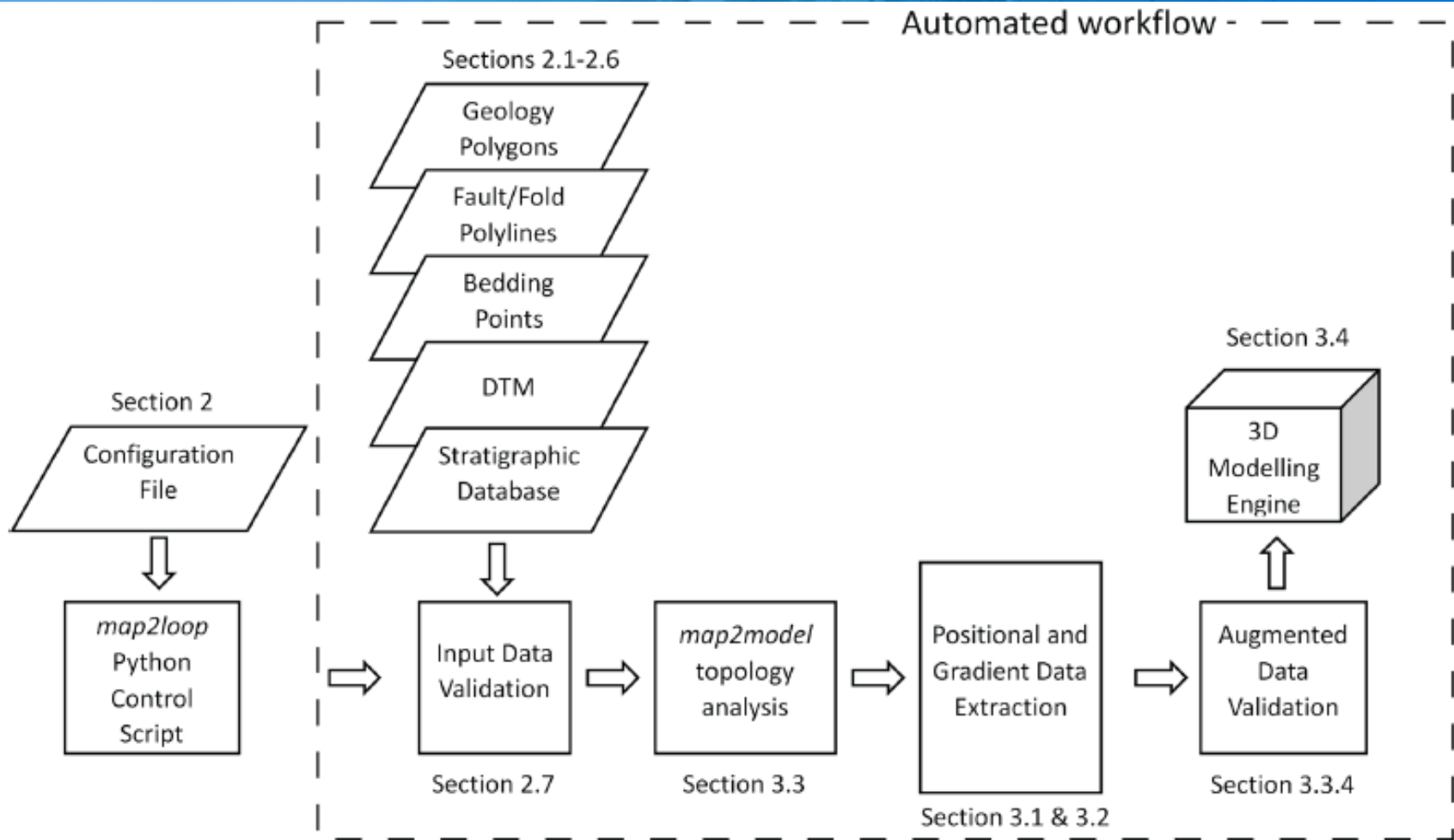


图2 自动化工作流：一旦创建了配置文件和在map2loop控制脚本中定义了工作流参数(表1)，所有虚线框内的步骤完全自动化，没有手动干扰。更细节的工作流见图4。



Table 1. Parameters that may be modified from their defaults prior to the automated workflow starting.

Parameter name	Meaning	Default value	Data type
aus	Indicates if the area is in Australia for using ASUD	TRUE	bool
close_dip	Shows the dip to assign to limbs of folds, -999 means use interpolated dip as local dip estimator, otherwise apply fixed dip assuming normal younging.	-999	int
contact_decimate	Decimates contact data by only saving every n th value.	5	int
contact_dip	Shows contact dip information, -999 means use interpolated dip as local dip estimator, otherwise apply fixed dip assuming normal younging.	-999	int
contact_orientation_decimate	Decimates orientation data by only saving every n th value.	5	int
deposits	Mineral deposit commodities for focused topology extraction. Not discussed in this paper.	"Fe,Cu,Au,NONE"	str
dist_buffer	Buffer for processing plutons to ensure faults that stop at plutons are correctly analysed by <i>map2model</i> .	10	int
dtb	Shows the path to depth-to-basement grid.	10	str
fat_step	Shows how much to step out normal to the fold axial trace for limb orientation to be added.	750	int
fault_decimate	Decimates fault data by only saving every n th vertex.	5	int
fault_dip	Shows default fault dip, -999 means add randomly assigned value between $\pm 60^\circ$.	90	int
fold_decimate	Decimates fold data by only saving every n th vertex.	5	int

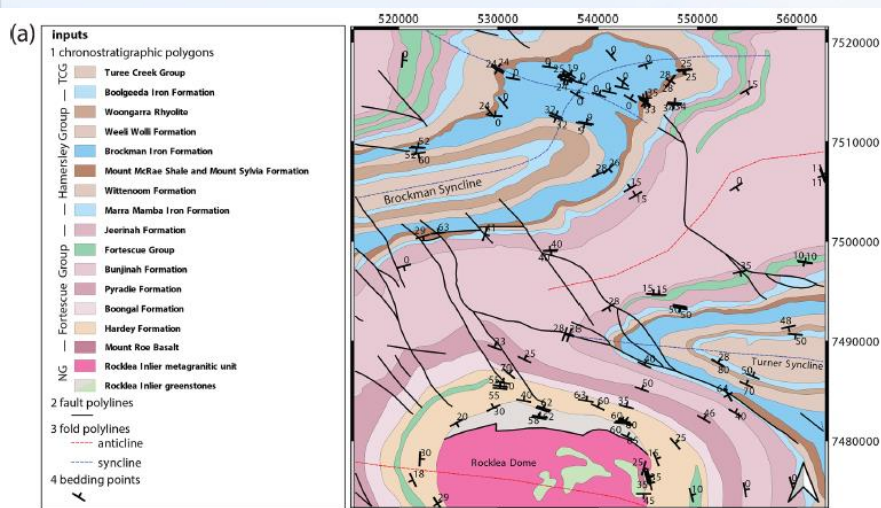


interpolation_scheme	Shows which interpolation method to use for <i>scipy_rbf</i> radial basis or <i>scipy_idw</i> inverse distance weighting.	"scipy_rbf"	str
interpolation_spacing	Shows interpolation grid spacing in metres.	500	int
intrusion_mode	Takes a value of 0 to only exclude sills or 1 to exclude all intrusions from basal contacts.	0	int
max_thickness_allowed	When estimating local formation thickness, this provides an upper limit to valid thicknesses to avoid unlikely thickness values.	10 000	int
min_fault_length	Gives the minimum fault length (tip-to-tip straight line distance) to be used.	5000	int
misorientation	Gives the maximum misorientation of pole to great circle of bedding between stratigraphic groups such that they still are considered to be part of same supergroup.	30	int
null_scheme	Value of null values (i.e. surface outcrop) in the depth to basement grid.	"null"	str
orientation_decimate	Decimates orientation data by only saving every <i>n</i> th value	0	int
pluton_dip	Shows the default pluton contact dip.	45	int
pluton_form	Gives possible forms from "domes", "saucers" "pendants", "batholiths".	"domes"	str
thickness_buffer	Shows how far away to look for the next highest unit when calculating formation thickness.	5000	int
use_fat	Uses fold axial trace info to add near-axis bedding info.	TRUE	bool
use_interpolations	Uses all interpolated dips for modelling.	TRUE	bool



2 输入数据

使用2016 1:50万解译的西澳大利亚的基岩地质图以及WAROX露头数据库，构建Hamersley Region of Western Australia地区的first-pass模型（图1）。



组成有：Archean褶皱和覆盖在Archean基础之上的Proterozoic(元古代)地层，被50° NW-SE方向的，形成Nanjilgardy断层系统的一部分断层切割。

(b) Stratigraphic (overlies ->) Relationships

Turee Creek Group -> Boolgeeda Iron Formation
Wittenoom Formation -> Marra Mamba Iron Formation
Marra Mamba Iron Formation -> Jeerinah Formation
Jeerinah Formation -> Bunjinah Formation
Bunjinah Formation -> Pyradie Formation
Pyradie Formation -> Boongal Formation
Weeli Wolli Formation -> Brockman Iron Formation

图1 输入map2loop的6类输入: (a) 基岩地质图(1:50万); (b) 前7类的层序关系; (c) SRTM DTM



(c) Digital Terrain Model

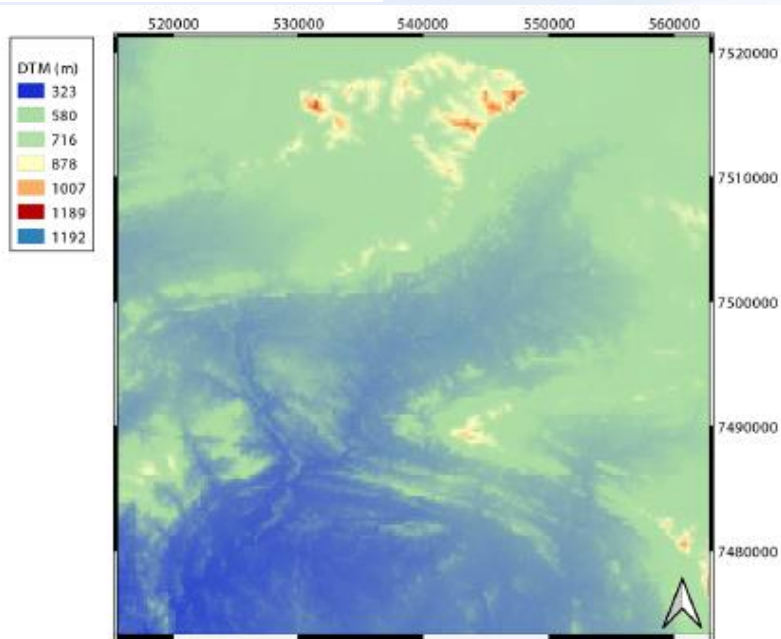


图1 输入map2loop的6类输入: (a) 基岩地质图(1:50万); (b) 前7类的层序关系;
(c) SRTM DTM



2 输入数据

map2loop使用Geopandas库从几种格式（shapefiles, MapInfo tab, JSON）文件加载数据，地理空间数据可以是任意标准的参考坐标系统（EPSG）。使用这些库加载和转换输入的地质几何和属性（表2）。

Table 2. Geometric features imported and saved by *map2loop* and *map2model*. The geometric objects refer to specific *Geopandas* data objects.

Geometric object	Input geological feature	Augmented-output geological feature
Point	Bedding	Bedding, contacts, faults, fold axial traces
Polyline	Faults, fold axial traces	None
MultiPolyline	Faults, fold axial traces	None
Polygon	Stratigraphic domains	None
MultiPolygon	Stratigraphic domains	None
Raster	DTM	DTM



2 输入数据

map2loop与map2model使用的**6种输入数据**（图1），这2个库使用一个配置文件，用户指定GIS层的哪个场。Jupyter Notebook帮助用户从输入层创建**HJSON**格式的配置文件（**Utility 1 – Configfilegenerator.ipynb**）。map2loop需要的最低输入数据见附录A。



2.1 Chronostratigraphic Polygon（必须）和 MultiPolygon layer

The layer may contain a mixture of single Polygons, MultiPolygons (sets of Polygons with the same non-spatial attributes), and or Polygons with holes (also stored as MultiPolygons, 图3)。这些对象作为Geopandas数据对象。

每个Polygon需要包含如下元素：

- （1）定义节点的有序封闭循环x y坐标的列表；
 - （2）层序编号或更低层级（如formation, member）的名称，称之为"units"
 - （3）一个或更多高层级的层序定义（如group, supergroup, spuersuite, province），称之为"groups"
 - （4）一个或更多岩性描述，有助于确定unit是否是火山、基石或其他类型的侵入体或其他类型的沉积岩？
 - （5）可选地（但重要的）：细尺度层序单元的最大和最小估计年代。
- 基岩地质图包含各formation的最大和最小年代，但是由于缺少绝对的地质年代约束，一个group内共享相同的范围。



2.1 Chronostratigraphic Polygon 和 MultiPolygon layer

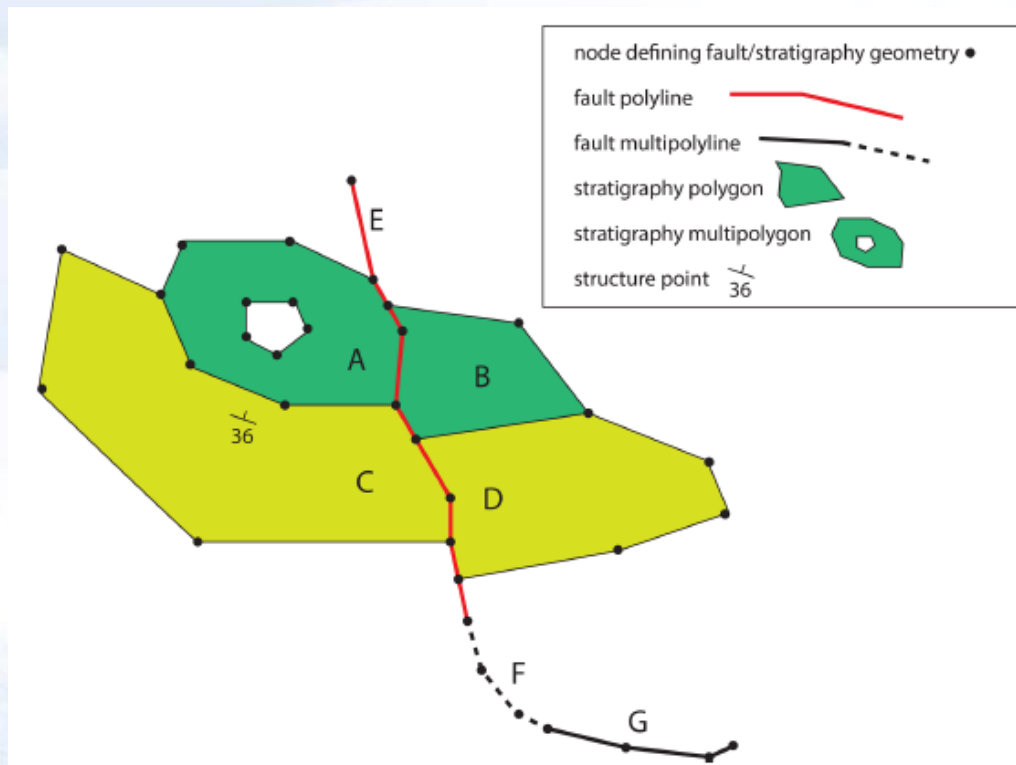


图3 地质图中使用的几何单元：B, C, D是层序Polygon，由节点的xy坐标定义。A是一个MultiPolygon，因为其包含一个洞。E是一个断层Polyline。F, G是断层MultiPolyline，描述相同断层的片段（断层E也是）。构造观测（层理测量）是Point类型。所有几何单元可能拥有多个属性，通过增加栅格的DTM信息，被转换为3D几何单元。



2.2 Fault Polyline （必须）和 MultiPolyline layer

该矢量层描述在**surface**处的断层或窄的剪切带的位置、方向和位移信息。

该层由**MultiPolylines**（具有相同非空间属性的**Polylines**组）混合组成。**map2loop**将**MultiPolylines**分解为分离的**Polylines**，允许正确实施断层长度和方向的分析。比用户定义长度还短的断层给过滤掉，降低模型的复杂度。

各**Polyline**需要包含如下元素：

- （1）定义节点的顺序开放循环**xy**坐标列表；
- （2）唯一的识别号，以某种方式标记断层；
- （3）在中点处存储断层的倾角和倾角方向（或走向）。

使用1:50万的地质图，由**map2loop**过滤提取断层。



2.3 Fold axial trace Polyline layer (可选)

map2loop过滤地质图，提取褶皱轴向轨迹。



2.4 Bedding orientation point layer

该矢量层描述层理的局部方向，一般不包含在地质图内，但可以从单独的数据库或地质野薄找到。还可以通过航拍图像解译或3点分析得到。

各点应包含如下元素：

- (1) 定义Point的单个xy坐标
- (2) 倾角信息；
- (3) 倾角方向或走向信息，这里称为"azimuth"以避免混淆；
- (4) 层理的Polarity (upright or overturned)。

使用2016 WAROX露头数据集。



2.5 Reference stratigraphy

Australian Stratigraphic Units Database, ASUD, 可以下载组(formation)的层级及上层的层序细节信息, 如2.1节提到的最大-最小年代的Polygon。层序信息一般没有空间信息, 但假设映射的层序年代Polygon与ASUD的编码相同, 如此得到层序关系(如A overlies B), map2model执行拓扑分析, 这样有助于定义地图区域上的局部层序。map2loop目前从ASUD数据集(定义相邻层序关系对, A overlies B)密集提取, 来定义局部层序(图1b)。

(b) Stratigraphic (overlies ->) Relationships

Turee Creek Group -> Boolgeeda Iron Formation
Wittenoom Formation -> Marra Mamba Iron Formation
Marra Mamba Iron Formation -> Jeerinah Formation
Jeerinah Formation -> Bunjinah Formation
Bunjinah Formation -> Pyradie Formation
Pyradie Formation -> Boongal Formation
Weeli Wolli Formation -> Brockman Iron Formation

图1(b)



2.6 Digital terrain model

SRTM 90m

中國地質大學



2.7 输入数据的校验

首先，将数据剪切到建模区域；

然后，检查新的层，确保有足够的层理数据，因为使用的算法需要至少3个方向插值，得到完整的层理方向场；

然后，检查是否地质Polygon文件有数据。没有数据，则出错。



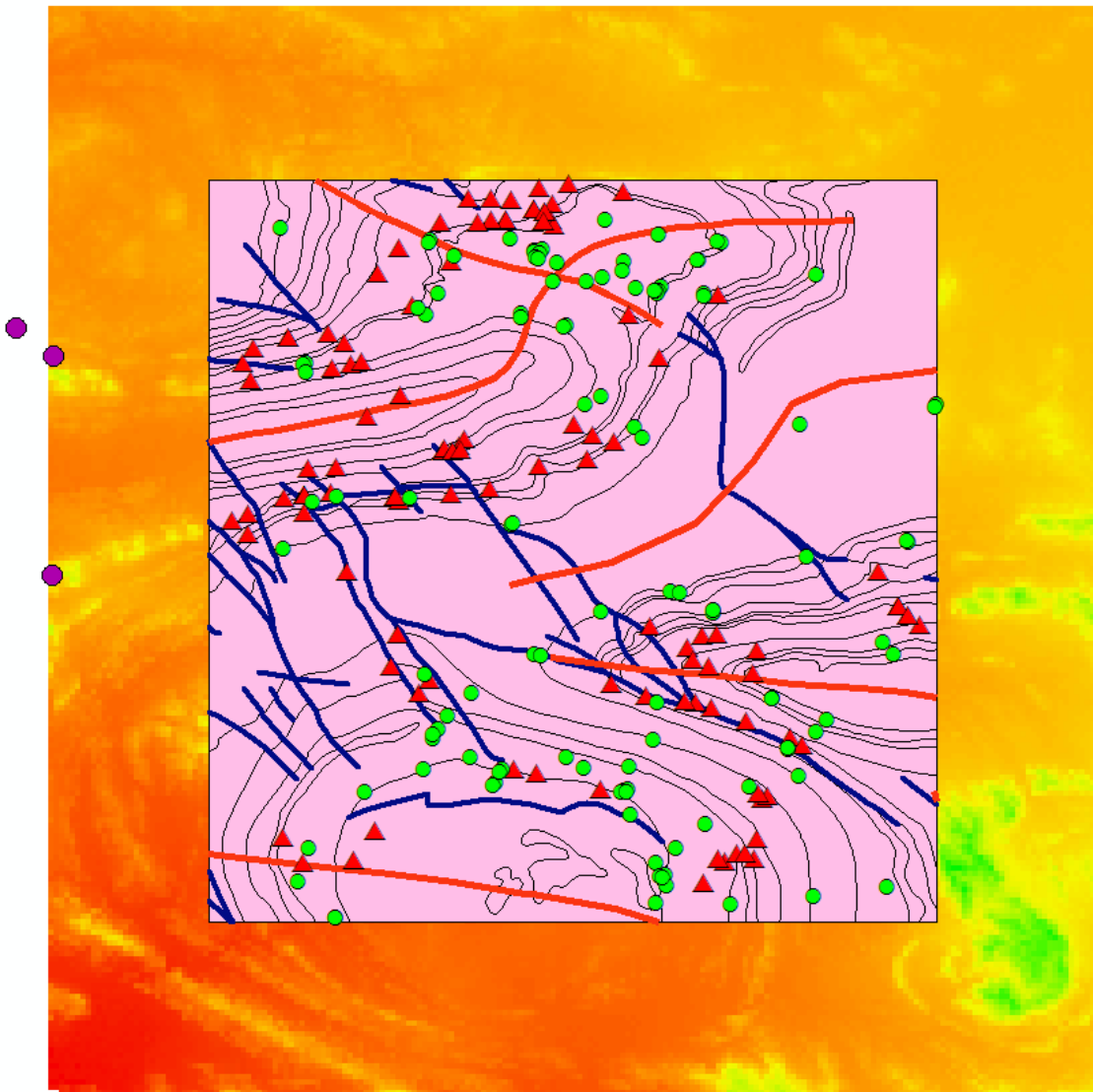
Table Of Contents

✕



Layers

- [-] E:\Reservoir_modelling\GeoModelling\LOOP3D\
- ☒ structure_clip
- ☒ null_mindeps
- ☒ mindeps_clip
- ☒ folds_clip
- ☒ faults_clip
- ☒ geol_clip
- [-] E:\Reservoir_modelling\GeoModelling\LOOP3D\
- ☒ dtm_rp.tif
- Value
- High : 1175
- Low : 297





3 方法

map2loop与map2model融合节2的输入数据，生成一系列csv，geotiff和gml 格式的输出，可直接用于3D地质建模或2D研究的分析数据源。

map2model 实施地质图的空间和时间拓扑分析：断层-断层交互、断层地层交互和局部地层分析。**map2loop进一步细化分析结果，融入非地图数据源信息，如地层数据集，作用是封装map2model，并实施其他计算。**

关键计算的伪代码见附录B。

3维建模引擎用到的具体输出见表3。



3 方法

Table 3. Comparison between model engine inputs.

Modelling engine	Digital terrain model	Stratigraphy	Orientation data	Stratigraphic units	Faults	Fold axial traces
<i>LoopStructural</i>	Used	two-level	Bedding, Cleavages	Position, thickness of units	Position, age relationships w.r.t. units and each other, displacement, ellipsoid for limited extent faults	Not used directly
<i>GemPy</i>	Used	two-level	Bedding	Position	Position, age relationships w.r.t. to units and each other, displacement, ellipsoid for limited extent faults	Not used directly

表3 GemPy与LoopStructural建模引擎输入数据的比较



3 方法

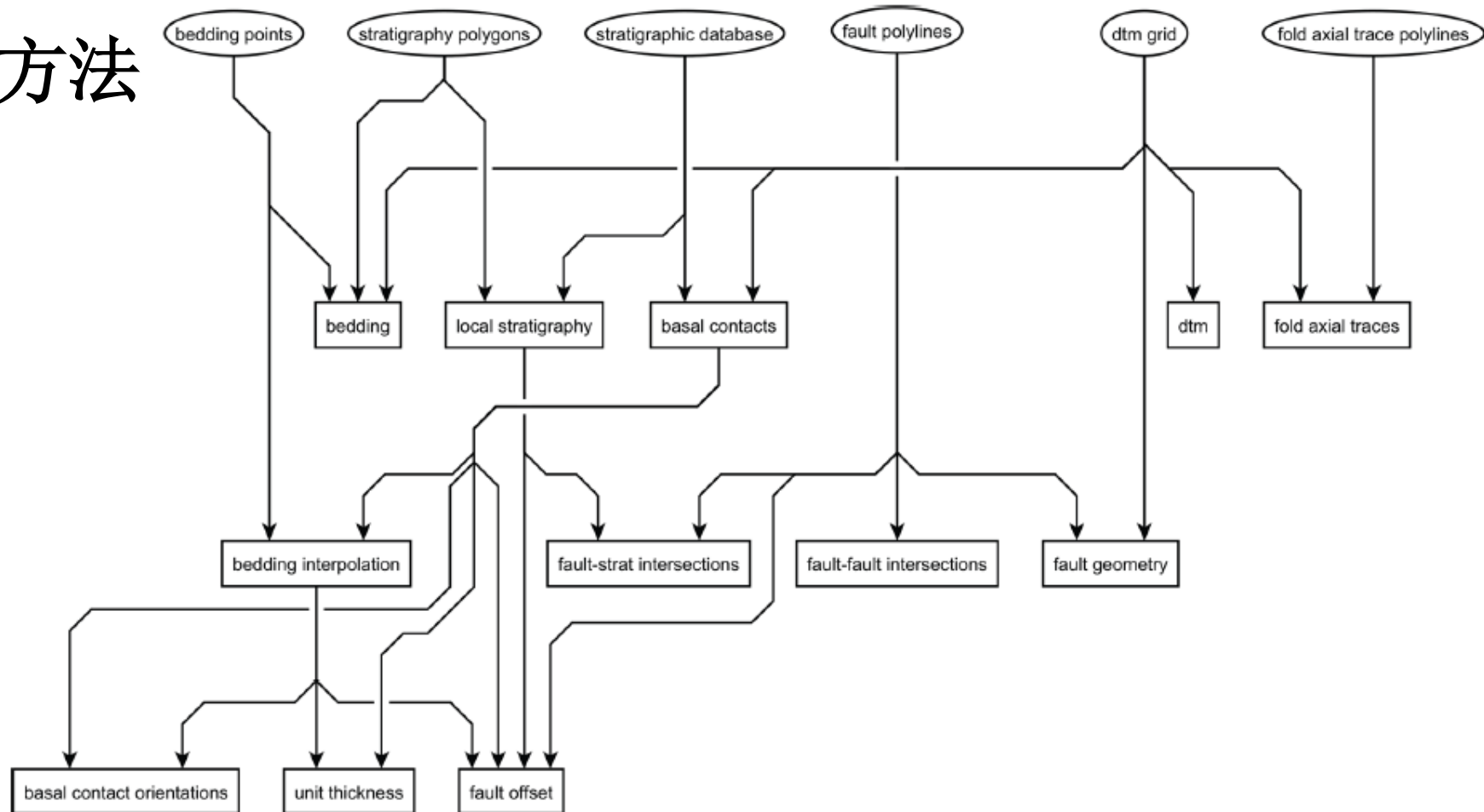


图4 GIS地图层, Web服务器和地层数据库提供的数据流输入 (椭圆框), 处理后的数据 (矩形) 是整合输入数据和map2loop工作流得到的。map2model处理拓扑关系分析: 断层-断层相交、断层-地层相交和局部的地层分析, 所有其他计算都是由map2loop管理的。



3.1 位置输出

map2loop算法输出的第一类建模约束提供位置性数据，即定义一种地质特征的x,y,z位置，包括断层、侵入体及层序接触的位置。

3.1.1 DTM

3.1.2 基岩接触

3.1.3 断层位置及维度

3.1.4 褶皱轴向迹的位置及维度

3.1.5 局部单元厚度

3.1.6 局部断层位移



3.1 位置输出

3.1.6 局部断层位移

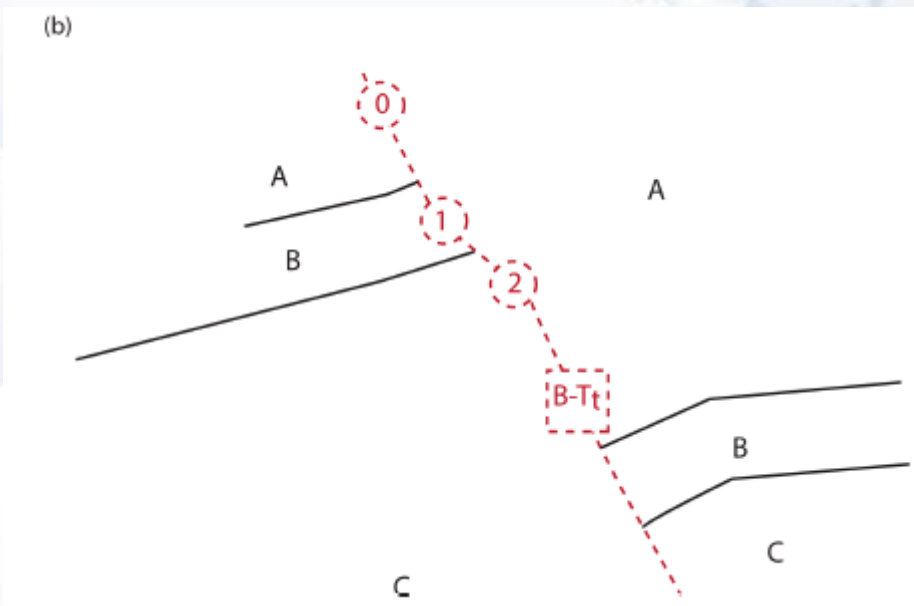
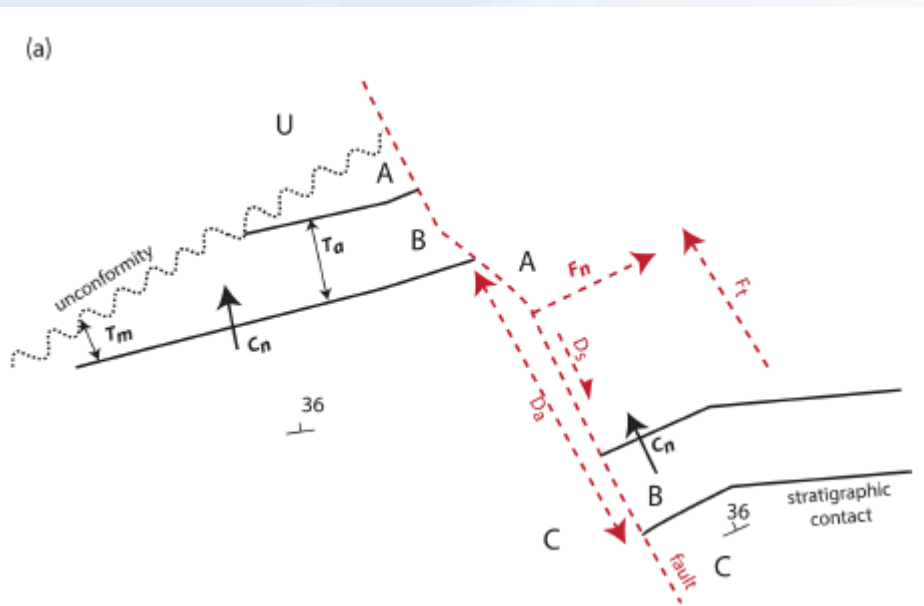


图5



3.2 梯度输出

3.2.3 断层方向

(b) faults

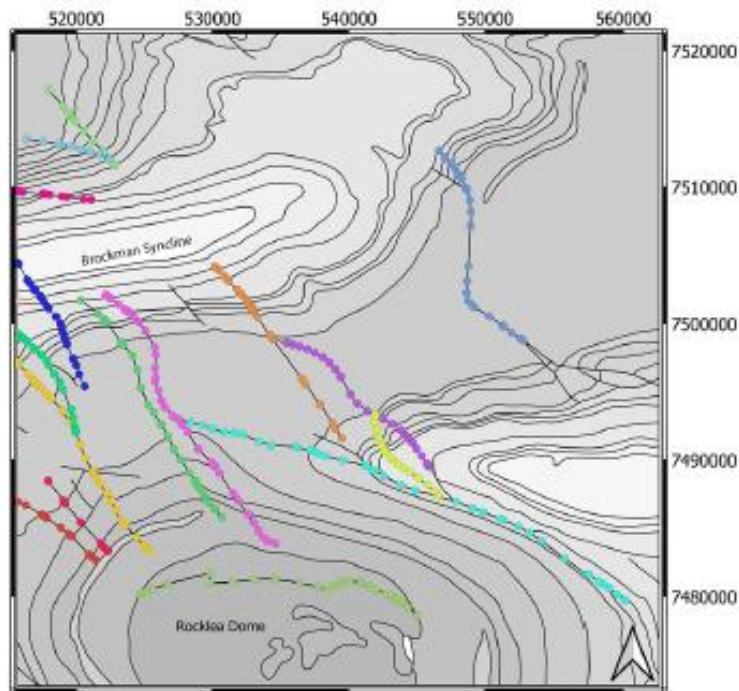


图7



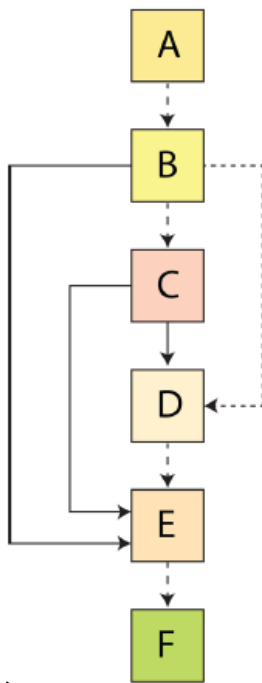
3.3 拓扑输出

map2loop使用map2model C++库，从地质图提取当地层序、构造和侵入关系。

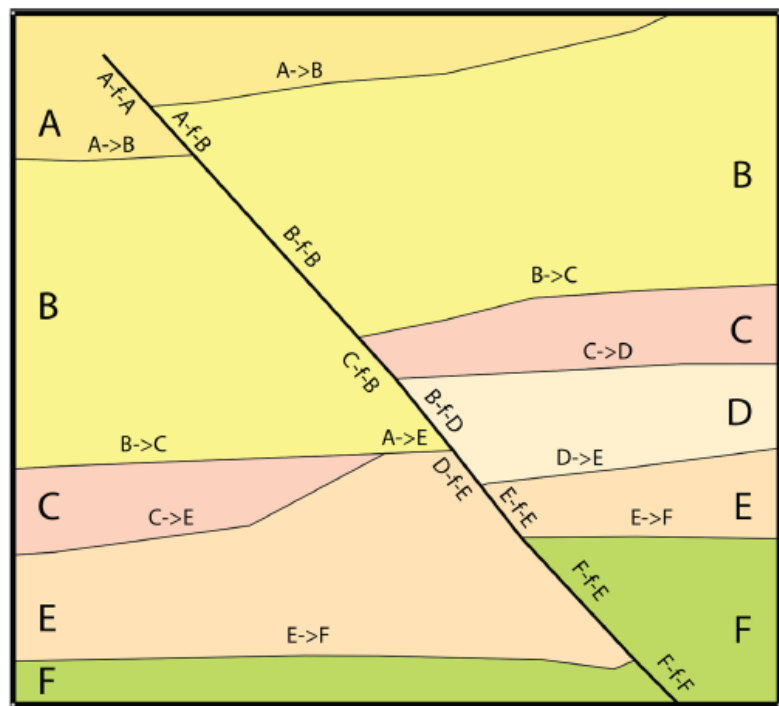
map2model使用map2loop提供的2个图层：

- 层序年代Polygon层（2.1节）
- 断层Polyline层（2.2节）

map2loop的输出是层序表(csv格式)，定义Unit与Group的顺序，以及一个表格（哪个groups形成待插值的super-group）。



3.3.1 局部层序



A->B : A stratigraphically overlies B

A-f-B : A is in faulted contact with B

——> stratigraphically overlies

- - - -> stratigraphically overlies and is also in faulted contact with

- - - - -> is in faulted contact with

图6 map2model提取的拓扑关系（层序关系）



3.4 处理(Augumented)数据的验证

自动完成最终所有处理后数据的检验，确保没有orphan数据。

3.5 使用map2lopp和map2model处理数据3D建模

2个开源库处理的数据比一定是3D建模有关的，但对地质背景区域分析提供有用的数据。

map2loop与map2model的输出提供所有构建3D地质模型(GemPy与LoopStructural)需要的信息。

数分钟完成建模，降低地质学家的通过地质建模做地质分析的入门障碍。

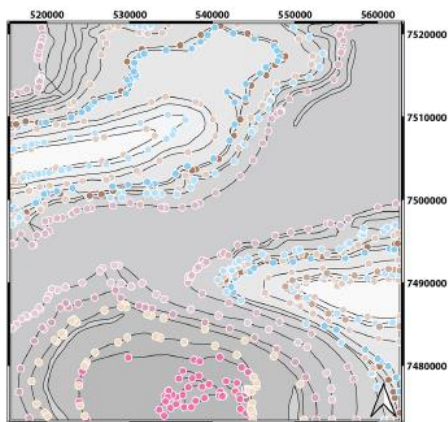


4 结果

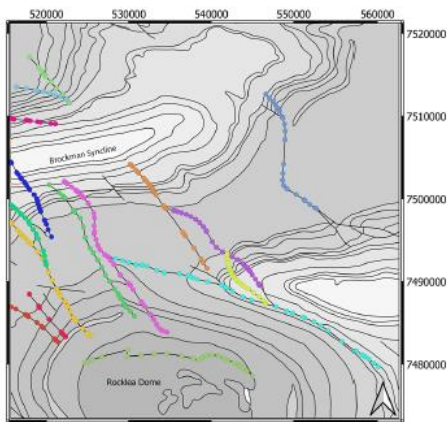
图7

4.1 位置计算的结果

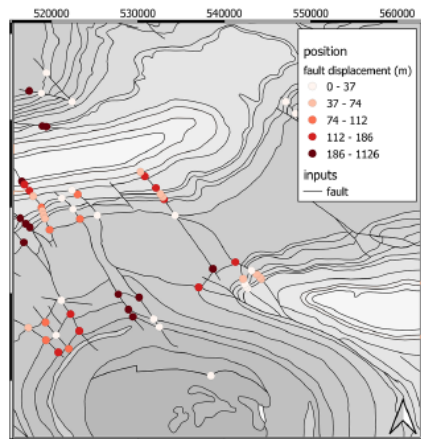
(a) contacts



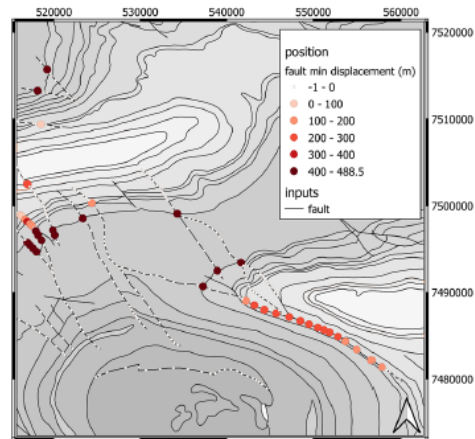
(b) faults



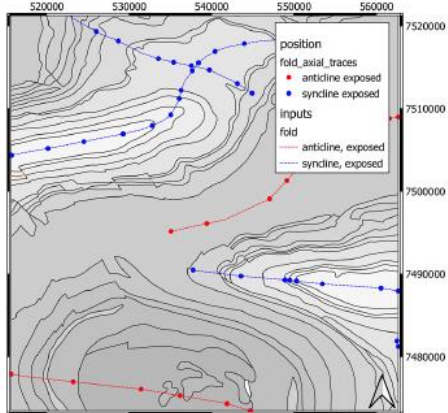
(e) fault offset



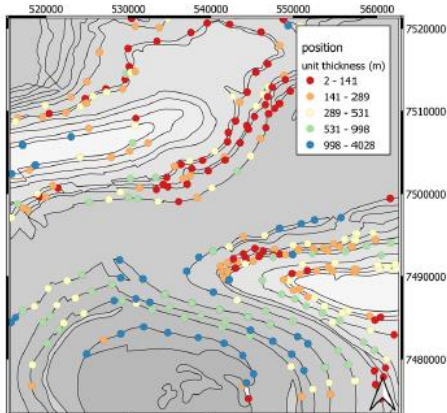
(f) fault min offset



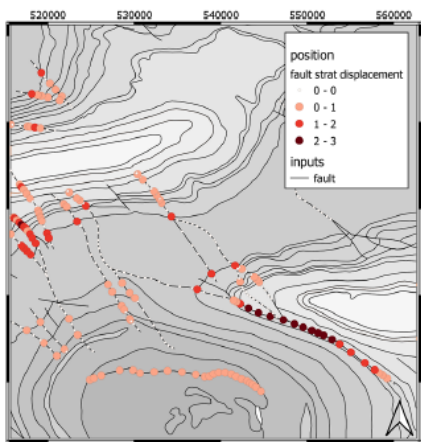
(c) folds



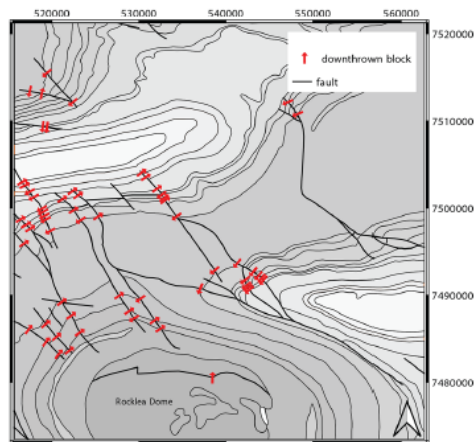
(d) unit thickness



(g) fault strat offset



(h) downthrow

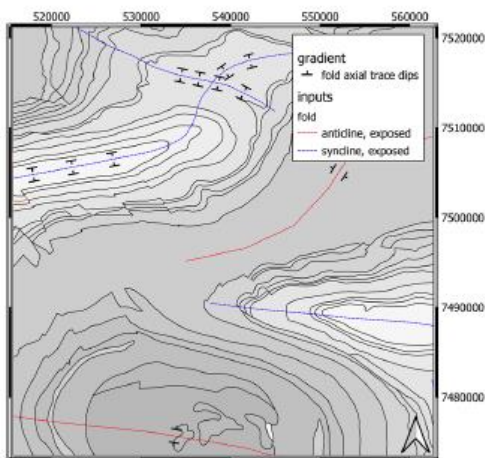




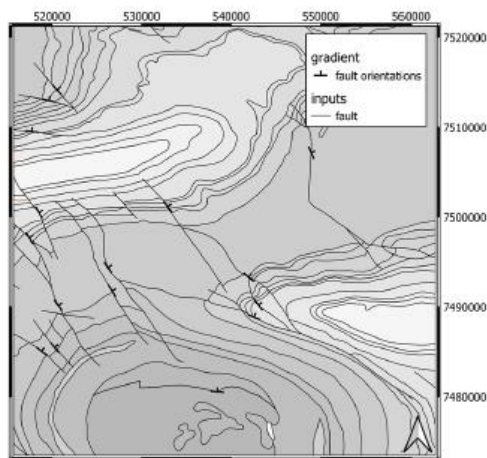
4 结果

4.2 梯度计算的结果

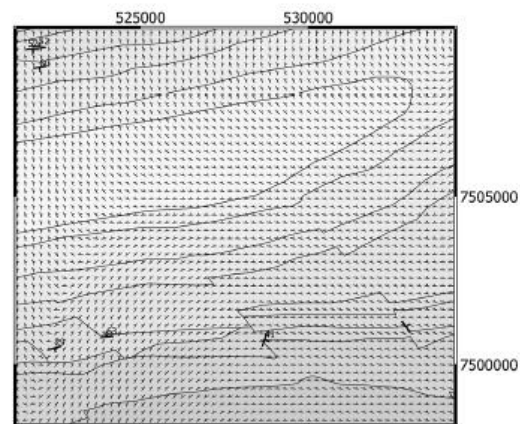
(a) fold orientations



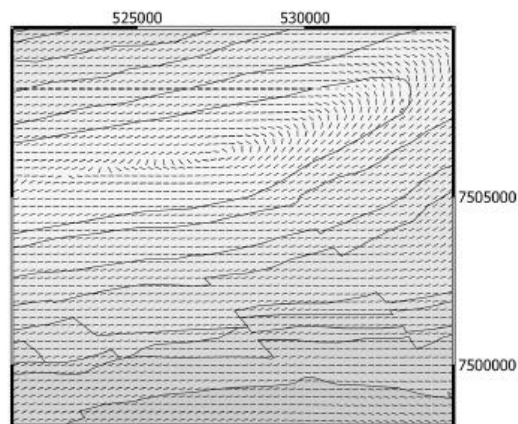
(b) fault orientations



(c) orientation field



(d) contact field



(e) combined field

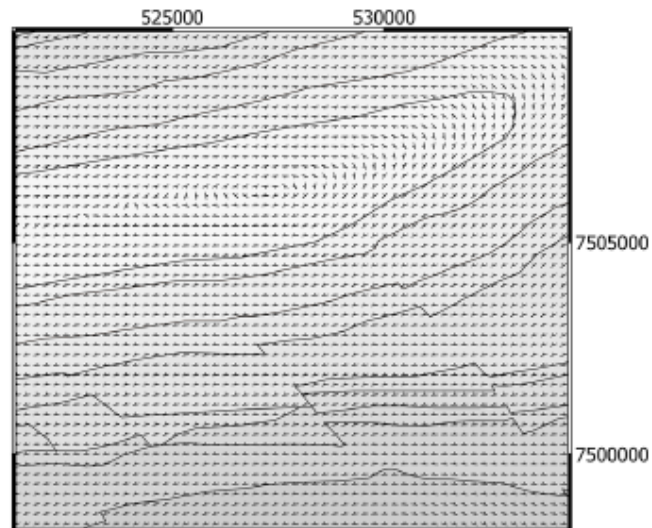


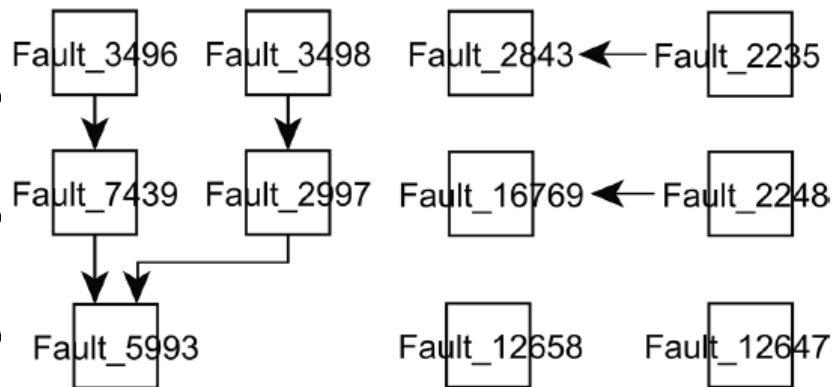
图8 从地质图求得的梯度信息，放大到Brockman向斜(syncline)



艰苦朴素 求真务实

4 结果

4.3 拓扑计算的结果



(c) fault-unit relationships

unit/fault	Fault_7436	Fault_7439	Fault_2235
Boolgeeda_Iron_Formation	0	0	0
Boongal_Formation	0	0	1
Brockman_Iron_Formation	0	1	0
Bunjinah_Formation	0	1	1
Fortescue_Group	0	0	1
Hardey_Formation	1	0	1
Jeerinah_Formation	0	1	1
Marra_Mamba_Iron_Formation	0	1	1
Mount_McRae_Shale_and_Mount_Sylvia_Formation	0	1	0
Pyradie_Formation	0	0	1
Rocklea_Inlier_greenstones	1	0	0
Rocklea_Inlier_metagranitic_unit	1	0	0
Turee_Creek_Group	0	0	0
Weeli_Wolli_Formation	0	0	0
Wittenoom_Formation	0	1	0
Woongarra_Rhyolite	0	0	0

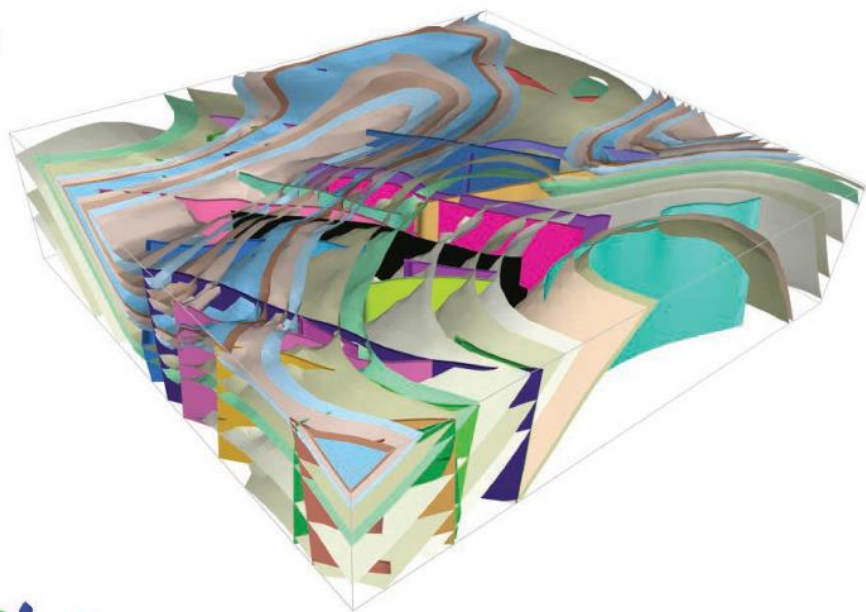
图9 从地质图求得的拓扑信息: (a) 层序年代关系



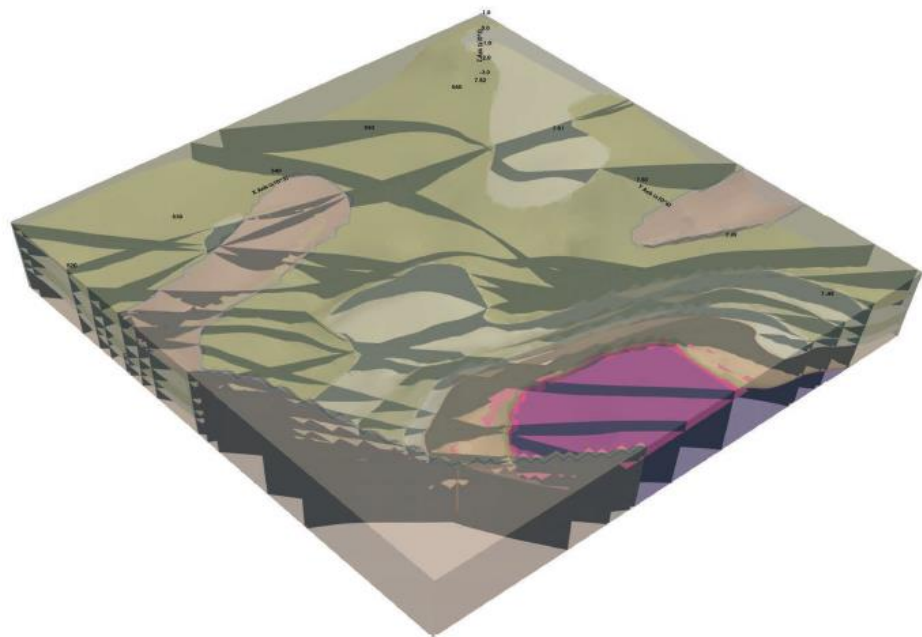
4 结果

4.4 三维模型计算的结果

(a)



(b)



LoopStructural

GemPy

图10 LoopStructural和GemPy最终构建的3D地质模型



中国地质大学
China University of Geosciences

艰苦朴素 求真务实

温家宝

5 讨论

5.1 对计算的改进

艰苦朴素
求真务实

中国地质大学



5 讨论

5.2 三维建模的局限

5.2.1 数据不充足

1、**基岩地质数据的获取**：（数字化）地质图仅能提供潜在的地表信息，表层地质图不能为**3D**模型提供足够的信息。很多地区，地表覆盖阻碍基岩地质的观测，此时没有地质图可以解构。随着很多地区的地球物理数据集变的可获取，产出了基岩顶部的解译地质图，以及覆盖-基岩界面几何的计算（如澳大利亚的**GSWA**）。**map2loop**的示例代码展示了，如何合并这些数据，来代替地表地质，作为地质建模的输入，但这对**Hamersley**案例不需要。**Loop**项有将地球物理数据整合入**LOOP**的工具(**Tomofast-x v1.0**, Giraud et al., 2021)，但不在本文讨论范围内。



5 讨论

5.2 三维建模的局限

5.2.1 数据不充足

2、地球物理数据的约束及地质学家的经验融入：还需要解译的断面数据来约束3D地质，但是即使有诸如地球物理数据、或者地震数据解译、或者重磁校验等的约束，仍然相比地表数据的约束要差。如果地震数据可获取，解译经验是偏差的显著来源。钻孔数据目前尚未考虑在本工作流中（[dh2loop](#), Joshi et al., 2021的工具可集成）。

地质学家绘制的没有地球物理约束的断面图依赖两种信息源：

- (1) 地质图，将来map2loop会提供当量信息；
- (2) 地质学家的经验，这难以程序化，是一个很大的挑战。

map2loop目前没有利用一定置信度的接触和断层的数据（传统地质图上的虚线）。



5 讨论

5.2 三维建模的局限

5.2.1 数据不充足

- 3、不是所有的地图都符合层序年代逻辑，例如图11，
- 4、制图早期，接触位置是很难定义的，因此一种方法要回避一块使用多个接触，SURFE软件允许不预定义接触位置的3D建模。
- 5、对于埋藏在盆地或风化层下面的地质，基于地图的方法可能就不合适了。地质学家很擅长构建这种缺乏数据的区域地质图，尽管3D地质模型的验证常受限于稀少的钻孔数据，此时模型可能是错误的，但很难说为什么。
- 6、以数码形式的地质数据的获取：各国家存储的地质调查数据及露头数据库的格式都不同，还可能根本就没有数据提供，除了地质图。对于露头数据库，有些国家就不开放。因此，建立对各类算法

- geology polygons with stratigraphic code and rock type info (required)

- fault polylines (required)

- bed dips as points in dip, dip direction (required)

- mineral deposit layer (optional)

- fold axial trace layer (optional)

- (1) 带层序编号和岩石类型信息的地质多边形（必须的）

- (2) 断层多段线（必须的）

- (3) 层理dip（dip和dip方向的点）（必须的）

- (4) 矿物沉积层（可选）

- (5) 褶皱轴向迹层（可选）

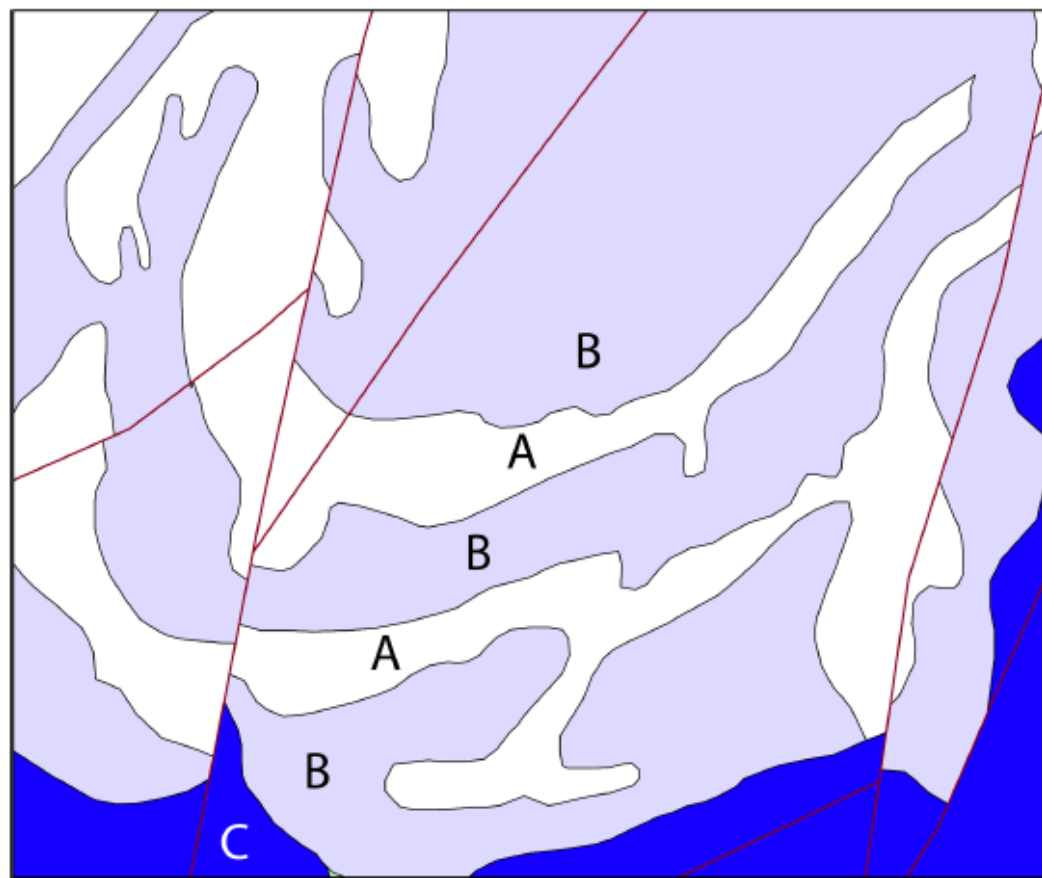


Figure 11. Example of lithological map descriptions that need re-coding in order to work in a chronostratigraphic modelling workflow. Assuming that the repetition of units is not structurally controlled, the lithostratigraphic sequence C-B-A-B-A-B-A needs to be recoded as C-B1-A1-B2-A2-B3-A3.





5 讨论

5.2 三维建模的局限

5.2.2 数据质量差

地质图信息错误导致如图8b的错误的封闭断层。

如果输入地图中层序信息是错误的，可以做些地图逻辑分析，如3.3.2节描述的。

如果使用解构数据构建3D模型失败了，可能是输入数据的不一致导致的，但模型不能告诉是哪个数据造成的错误，因此以后还需要一个健壮的地图校验程序，提供正确建模的指导。目前，重合Polygons和Polylines之间的节点的小的不匹配是可接受的。



5 讨论

5.2 三维建模的局限

5.2.3 数据的错误解构(deconstruction)

map2loop在解构地图数据过程中做了很多简化。可以自动检验断层位移和单元厚度的计算，保证沿着一个断面或接触的一致性。

5.2.4 不完备的3D建模算法

map2loop不能生成满意的3D地质模型的最后一个原因是：模拟系统本身不能很好地考虑所有类型的地质场景。隐式建模的模拟引擎在well-defined和变形不大的层序的区域，工作很好，尽管LoopStructural也可以处理poly-deformed的阶地。一旦overprinting构造很重要，隐式建模需要越来越多的信息（通常是解译后的数据，不是原始数据），来浮现地质学家设想的地质模型。地质学家头脑中的概念模型可称为“先验的概念”，这很难编成程序。



5 讨论

5.3 将来的工作

- (1) 随机模型系综生成器: https://github.com/Loop3D/ensemble_generator。
- (2) 批量的地质建模（1000个建模与1个建模的时间是差不多的）。
- (3) GemPy有自己的先进的框架分析不确定性。下一步，封装数据提取、3D地质建模、地球物理正反演工作流到Bayesian分析框架，因此所有建模、不确定度定量化和联合地质-地球物理反演决策的累积影响可以统一的方式进行分析。
- (4) 集成自动提取钻孔数据信息(Joshi et al., 2021)和考虑sill-like侵入接触(Alvarado-Neves et al., 2020)。
- (5) map2loop及相关代码，需要管理更多种类的数据数据集，比如钻孔和剖面数据（已经在进行了）。
- (6) 最大的挑战还是融入地质学家的先验概念到解构工作流(Jessell, 2021)。



中國地質大學
China University of Geosciences

艰苦朴素 求真务实

温家宝

艰苦朴素
求真务实
温家宝

中国地质大学