



波现象与智能反演成像研究组

地震波成像研究的新导向

多元信息表达及融合、宽带弹性参数估计、定量储层描述

报告人：王华忠

波现象与智能反演成像研究组 (WPI)

同济大学海洋与地球科学学院，上海

2024年04月18日

目录

- ◆一、油气勘探的新领域及油藏描述的新需求
- ◆二、地震波成像当前研究的发展困境
- ◆三、多元数据统计分析的思想与框架
- ◆四、多元信息表达与融合的合理理论框架
- ◆五、地震波成像研究的新导向
- ◆六、再论WPI当前侧重研究的问题



一、油气勘探的新领域及油藏描述的新需求

◆ 油气地震勘探最终目的：

◆ 精确地描述油气藏，进行准确的含油气性分析，做出最佳的钻井决策，获得最高的油气勘探效益。

◆ 油气勘探的新领域显然来自于各大含油气盆地。地震波成像的根本问题及今后的新导向均来自于各大盆地的实际勘探需求。

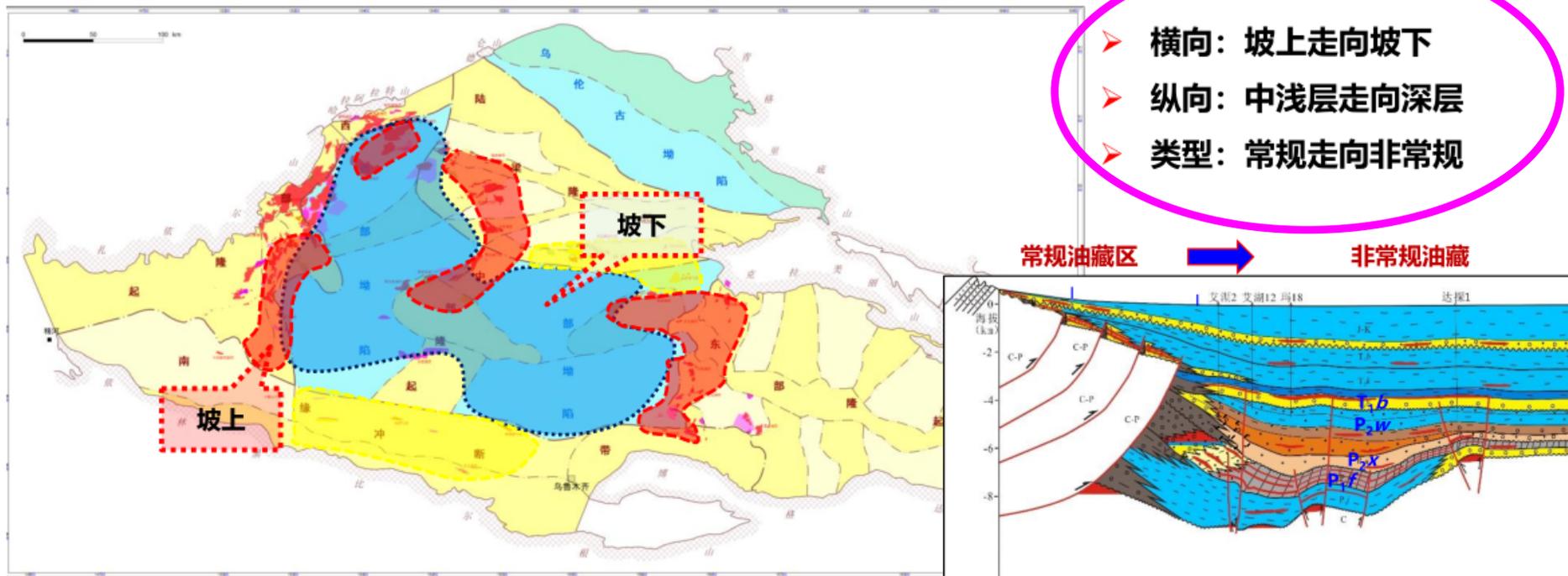
◆ 下面看看各大盆地的问题与挑战。

一、油气勘探的新领域及油藏描述的新需求

◆准噶尔盆地:

2012年以前, 准噶尔盆地周缘“中浅层(侏罗系及以上、4000m以浅、正向构造)”油气藏勘探, 有力支撑新疆油田连续10年原油稳产千万吨以上。2012以后, 勘探重点由坡上走向坡下、目的层由中浅层走向深层($\geq 4000\text{m}$ T+P)、勘探类型由常规走向非常规。

2020年度准噶尔盆地油气勘探成果图



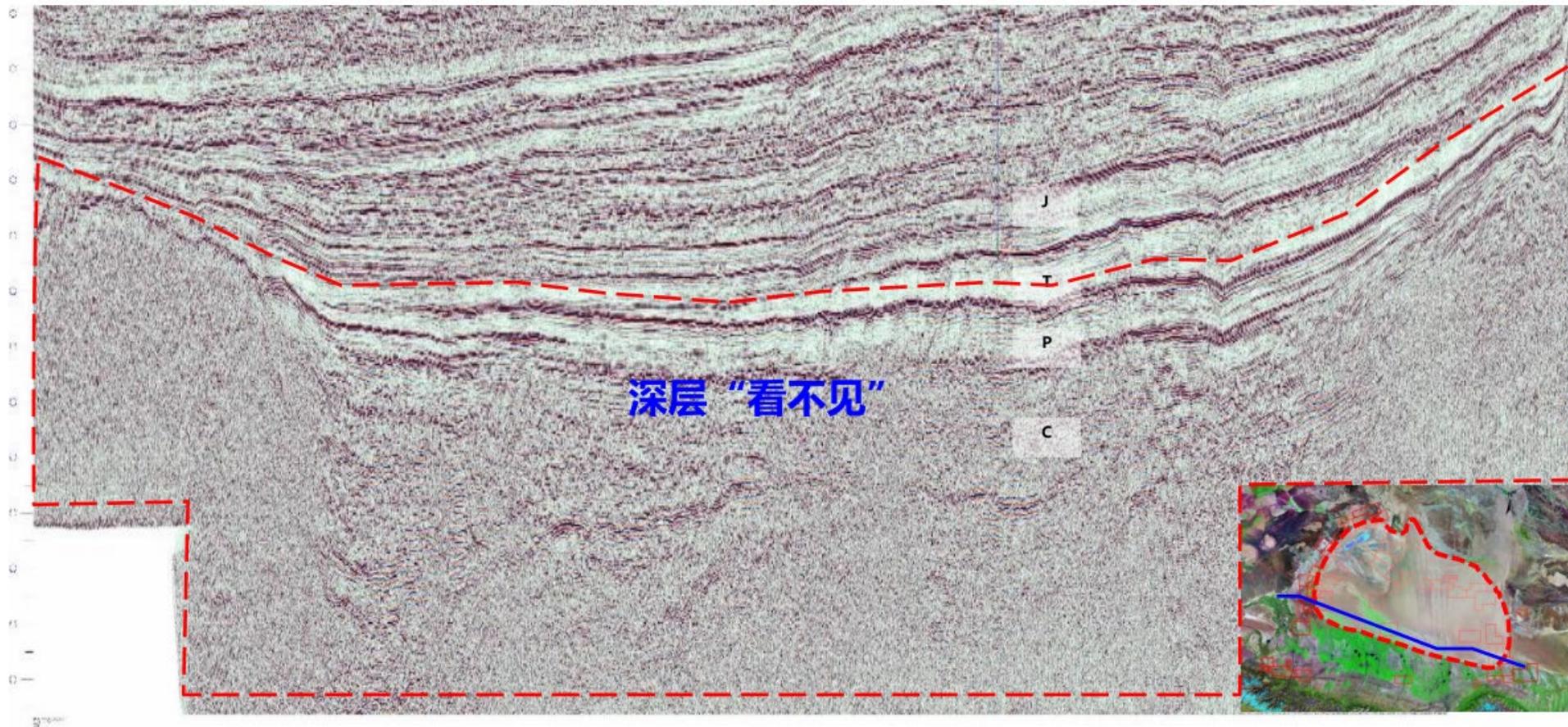
一、油气勘探的新领域及油藏描述的新需求

◆准噶尔盆地:

挑战1——深层成像“看不见”：富烃凹陷多处于厚沙漠区 ($\geq 80\%$)，以往地震资料深层普遍不成像，地层格架对比关系难以建立，制约盆地下组合整体研究与重大接替领域谋划。

EW7格架地震剖面

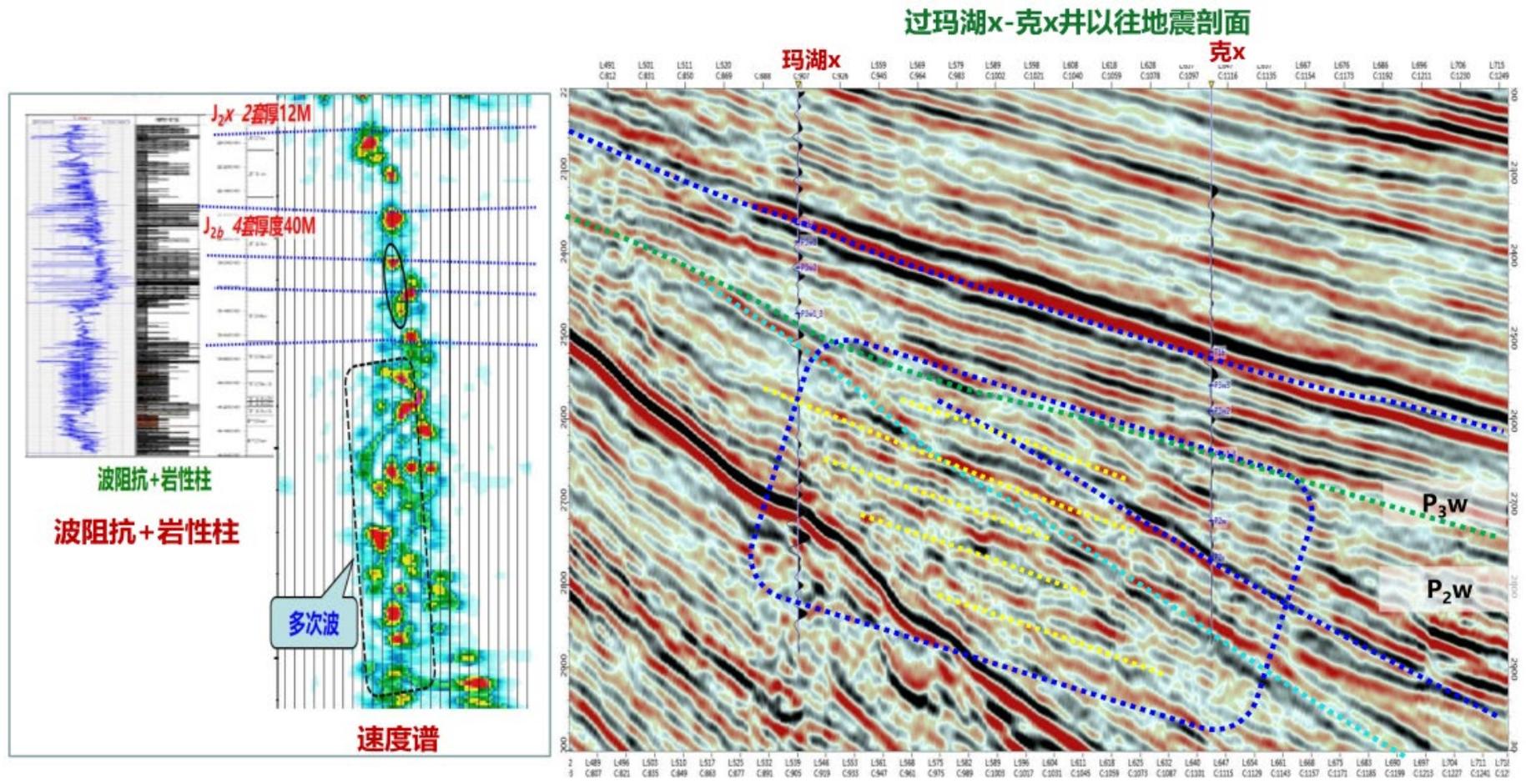
芳草1



一、油气勘探的新领域及油藏描述的新需求

◆ 准噶尔盆地:

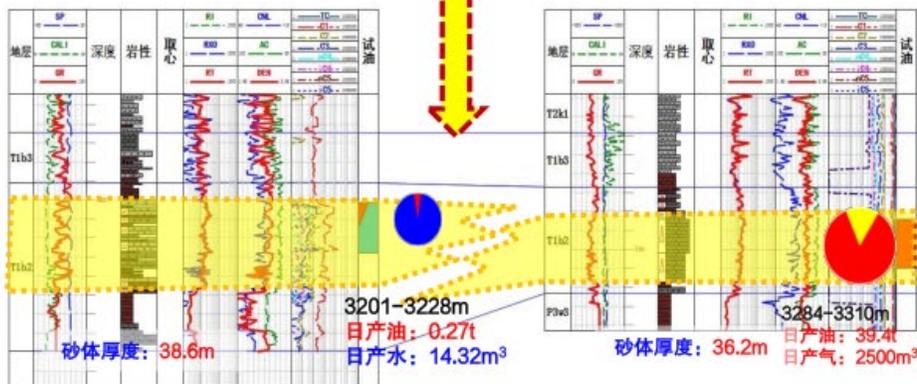
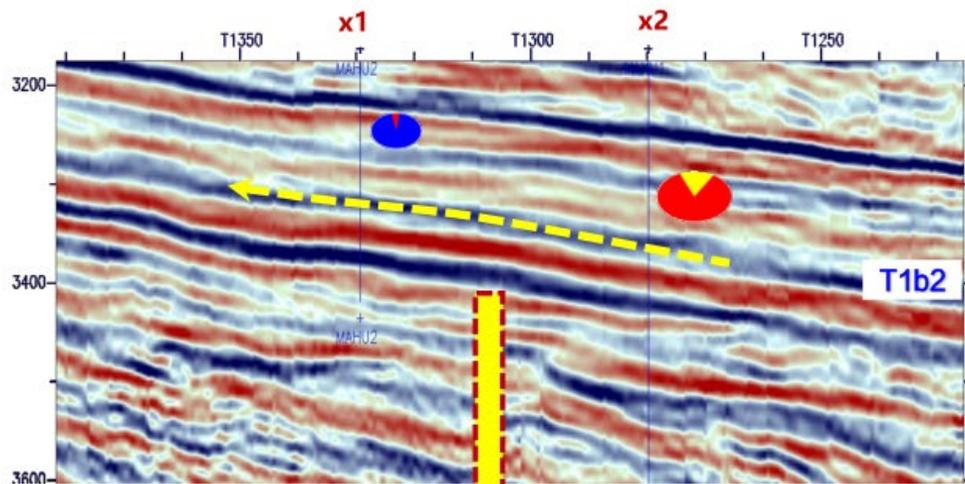
挑战2——相带“分不清”：受煤源多次波影响，二叠系地层真假难辨，严重制约地层型油气藏规模发现。



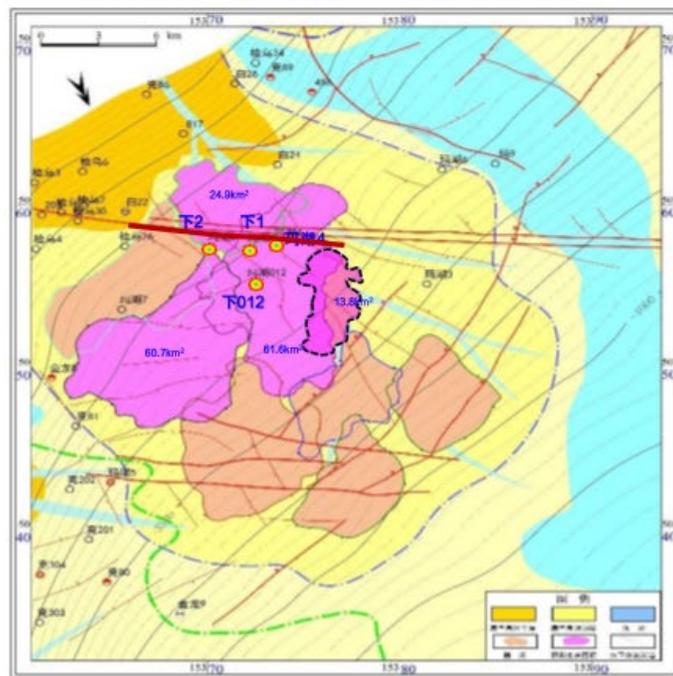
一、油气勘探的新领域及油藏描述的新需求

◆ 准噶尔盆地:

挑战3——砂体边界“分不开”：地震波吸收衰减严重，深层地震资料分辨率低，岩相变化与藏间关系难以准确描述，2012年X1风险井获得突破后，上钻X2、X3接连失利，勘探举步维艰。



克拉玛依扇百口泉组二段综合成果图



360日

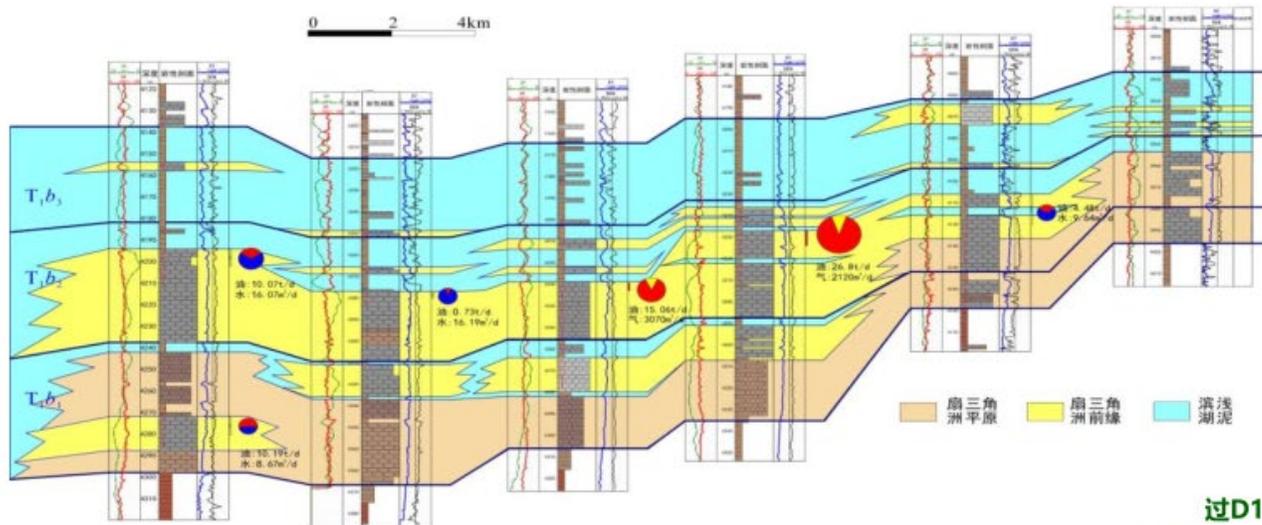
- 两井压力不同属不同油藏，但地震反射特征一致
- X2核磁共振好于X1井但产量却低于X1井

一、油气勘探的新领域及油藏描述的新需求

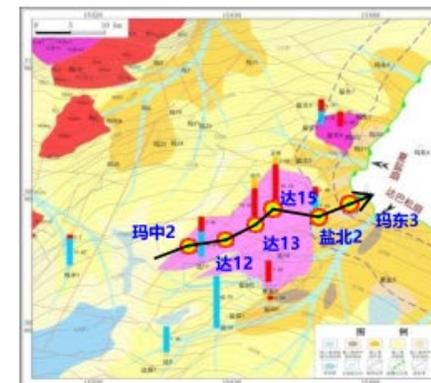
◆准噶尔盆地:

挑战4——甜点“识不准”：岩性、孔隙结构复杂，储层变化快，非均质性强，甜点体难以准确刻画。

过X2-D12-D13-D15-YB2-M3井三叠系百口泉组沉积相对比图

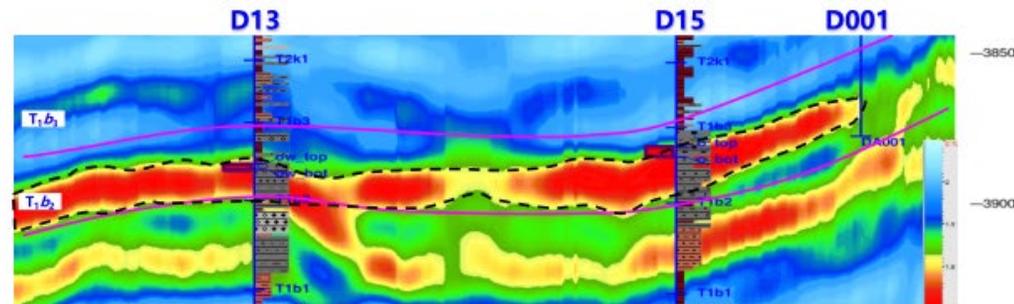


玛东斜坡区百二段勘探成果图



相同相带高部位含油水层，低部位纯油气层，油水关系复杂，更佐证了储层非均质强的特点

过D13-D15井连井纵横波速度比反演剖面



一、油气勘探的新领域及油藏描述的新需求



◆准噶尔盆地：

针对四大挑战，多年来依托集团科技重大专项、物探技术攻关，攻克技术瓶颈，形成了**深层地层岩性勘探配套技术**，支撑了**玛湖砾岩、吉木萨尔页岩油**两个**10亿吨级特大型油田**的发现。

序号	攻克4项瓶颈	研发4项配套技术	推动3大领域勘探	取得2项成果
1	深层成像“看不见”	两宽一高观测系统设计及观测技术系列	① 助力盆地上乌整体发现 ② 推动风城组的全面发现 ③ 夯实百口泉组储量落实	① 创立沙漠区深层地层岩性勘探配套技术 ② 支撑了玛湖、吉木萨尔两个10亿吨特大型油田的发现
2	煤源多次波“压不净”	层源双控多次波压制技术		
3	岩性、岩相“分不开”	全频保真高分辨处理技术		
4	相带及甜点“识不准”	相带刻画及甜点预测技术		

形成了**深层高保真地震勘探技术系列**、支撑了两个**10亿吨油田**发现

一、油气勘探的新领域及油藏描述的新需求

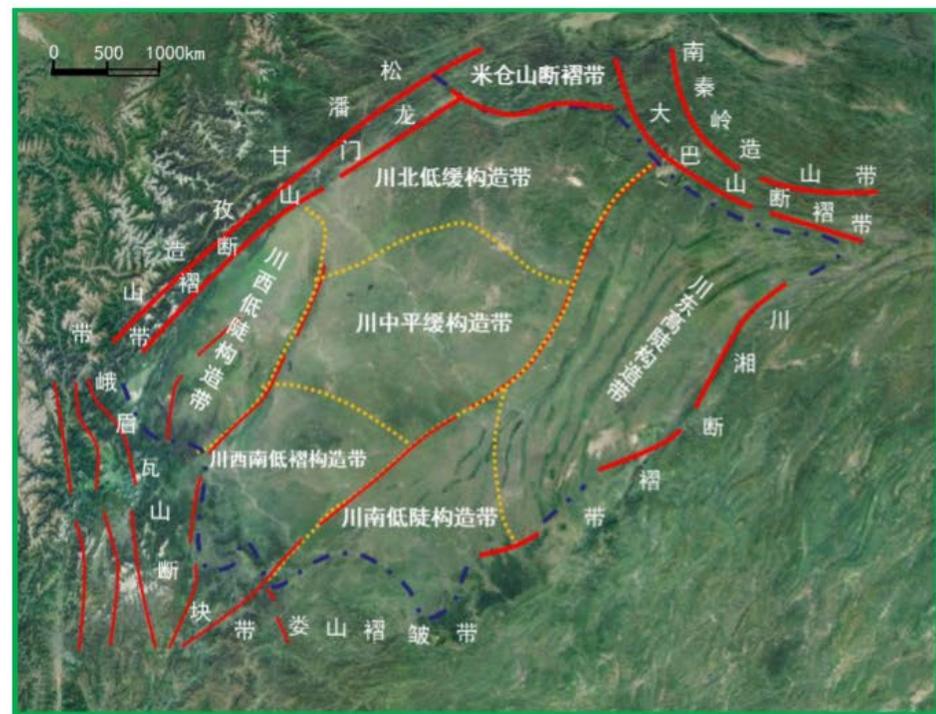
◆四川盆地:

(一) 四川盆地面积18万平方千米

- 位于处于扬子地块西缘，现今属于环青藏高原盆山体系
- 经历多期构造运动，盆内发育五大构造带



华南上扬子地区大地构造格局



四川盆地构造分区图

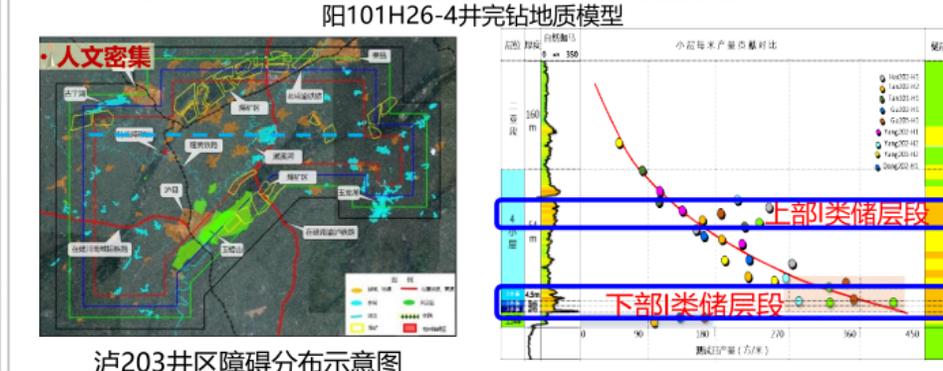
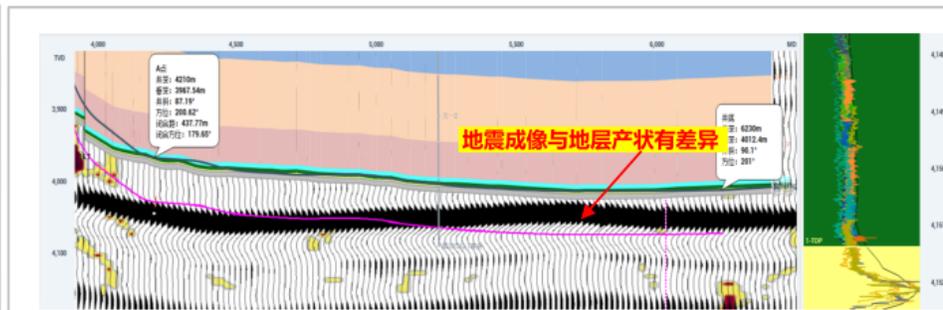
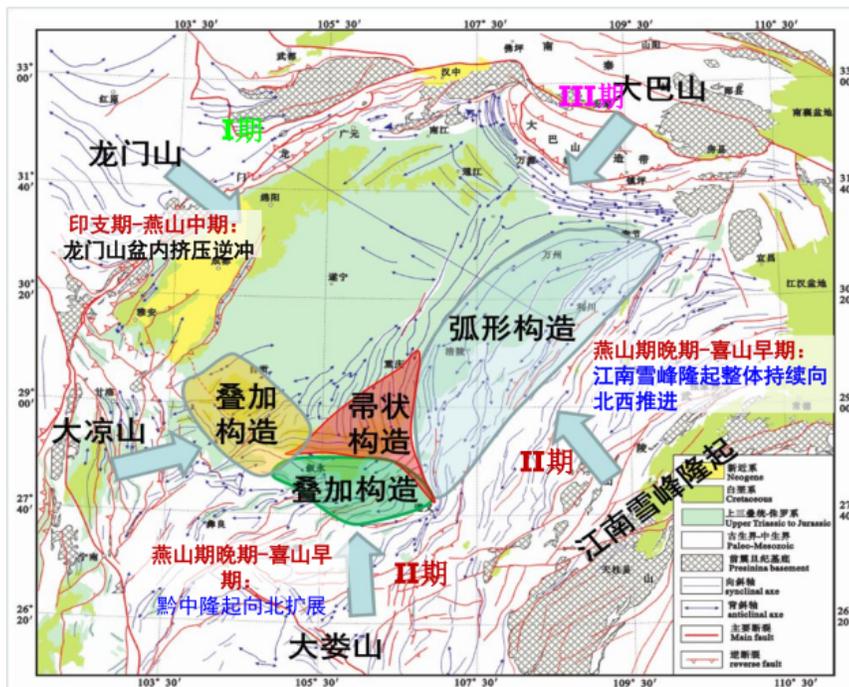
一、油气勘探的新领域及油藏描述的新需求

◆四川盆地:

三大挑战

- 印支期以来受三期构造运动挤压，地表、地腹条件复杂，加上人文密集影响，**地震精确成像难度大!**
- 构造和断裂复杂，具有多期次、多方向性，**精细预测难!**
- 页岩气开发走向I类储层连续厚度10-5m区，**薄储层反演分辨率不够!**

高产主控三要素



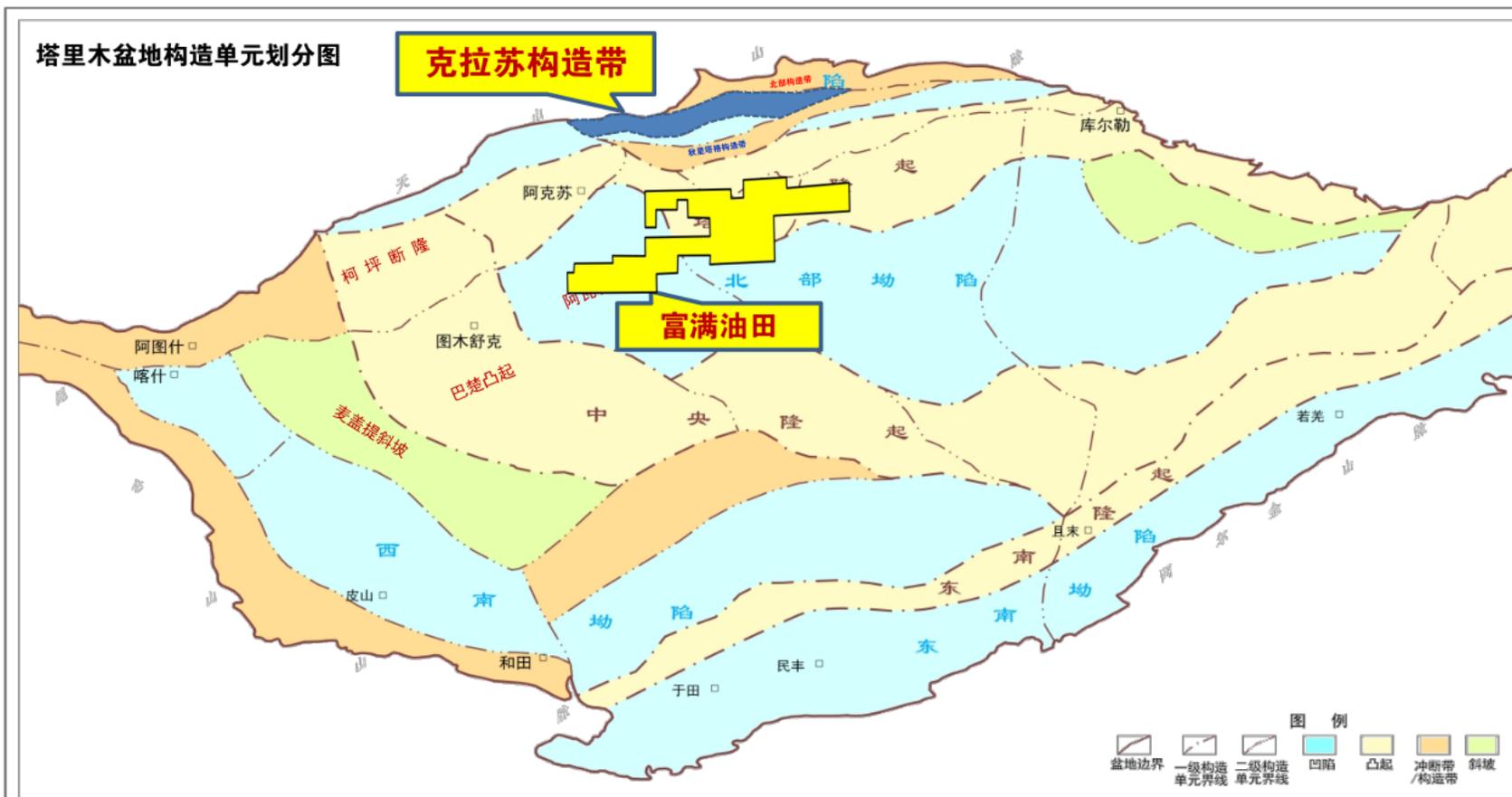
一、油气勘探的新领域及油藏描述的新需求

◆塔里木盆地克拉苏构造带：

- 克拉苏盐下逆掩叠置构造、富满碳酸盐岩断溶体储层是塔里木盆地两大典型油气勘探开发目标，同时也是塔里木油田增储上产的“两大粮仓”，2022年年产量达2100万吨（占油田整体约64%）

➤ 复杂山地之克拉苏构造带：
 位于库车山前盐下逆掩叠置构造区，勘探面积5500平方千米，主力产气层白垩系巴什基奇克组，2022年产气226亿方

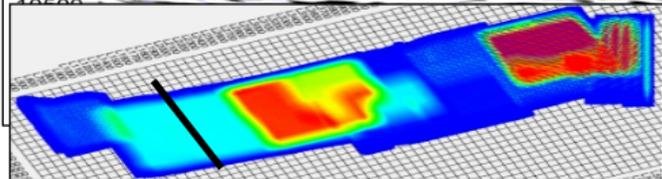
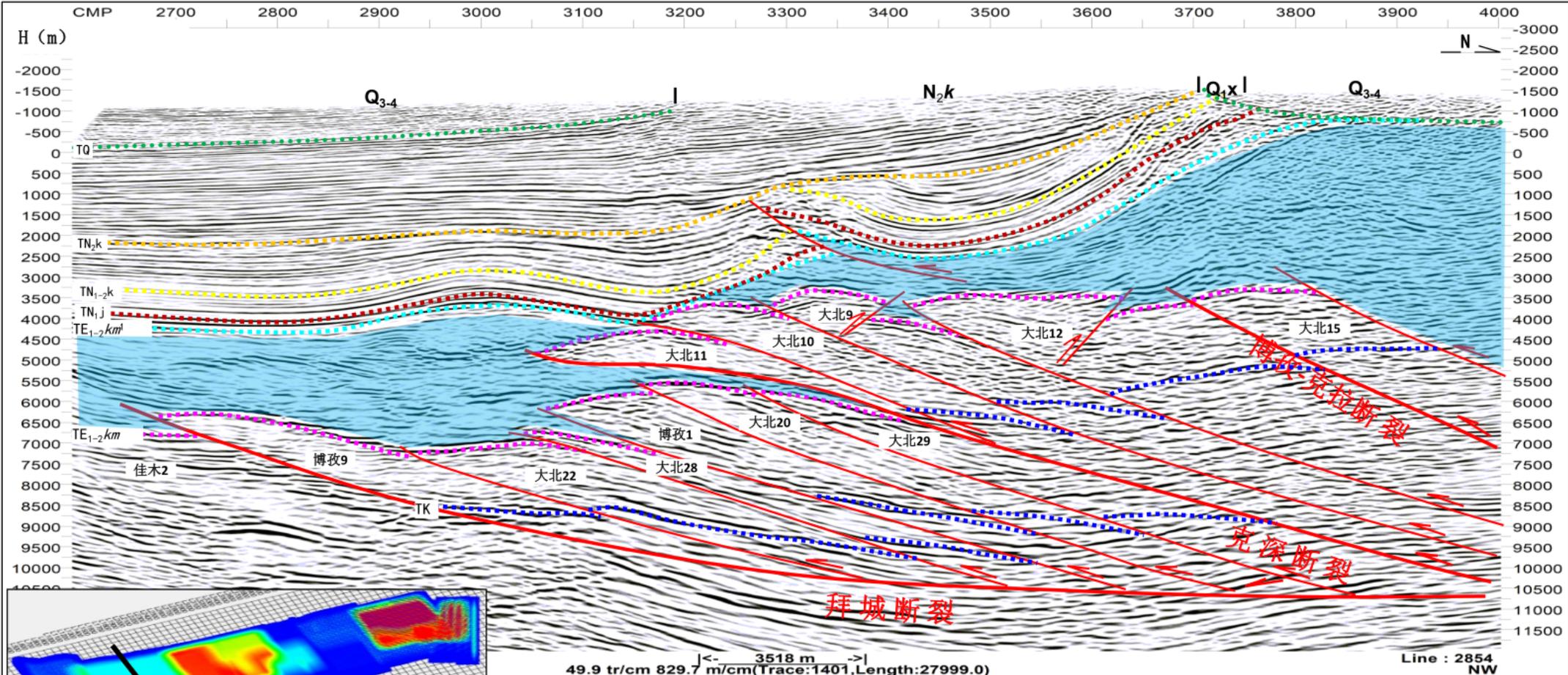
➤ 大沙漠区之富满油田：
 位于塔北隆起大沙漠区，勘探面积1.7万平方千米，主要目的层奥陶系一间房-鹰山组，2022年产油300万吨



一、油气勘探的新领域及油藏描述的新需求

◆塔里木盆地克拉苏构造带：

□ 地下复合盐层塑性复杂变形、构造复杂，最厚达4000m，盐下逆掩推覆多层叠置，最多超过10层

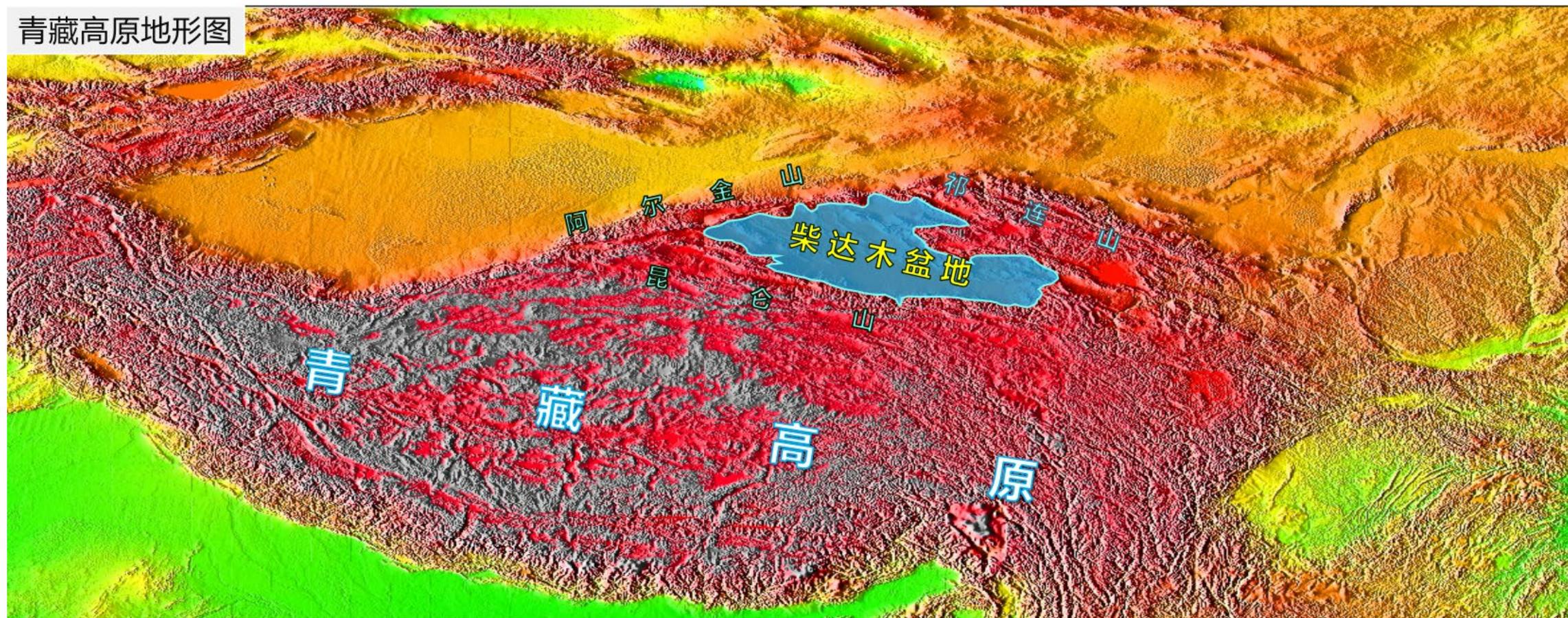


克拉苏博孜段PSDM剖面

一、油气勘探的新领域及油藏描述的新需求

◆柴达木盆地:

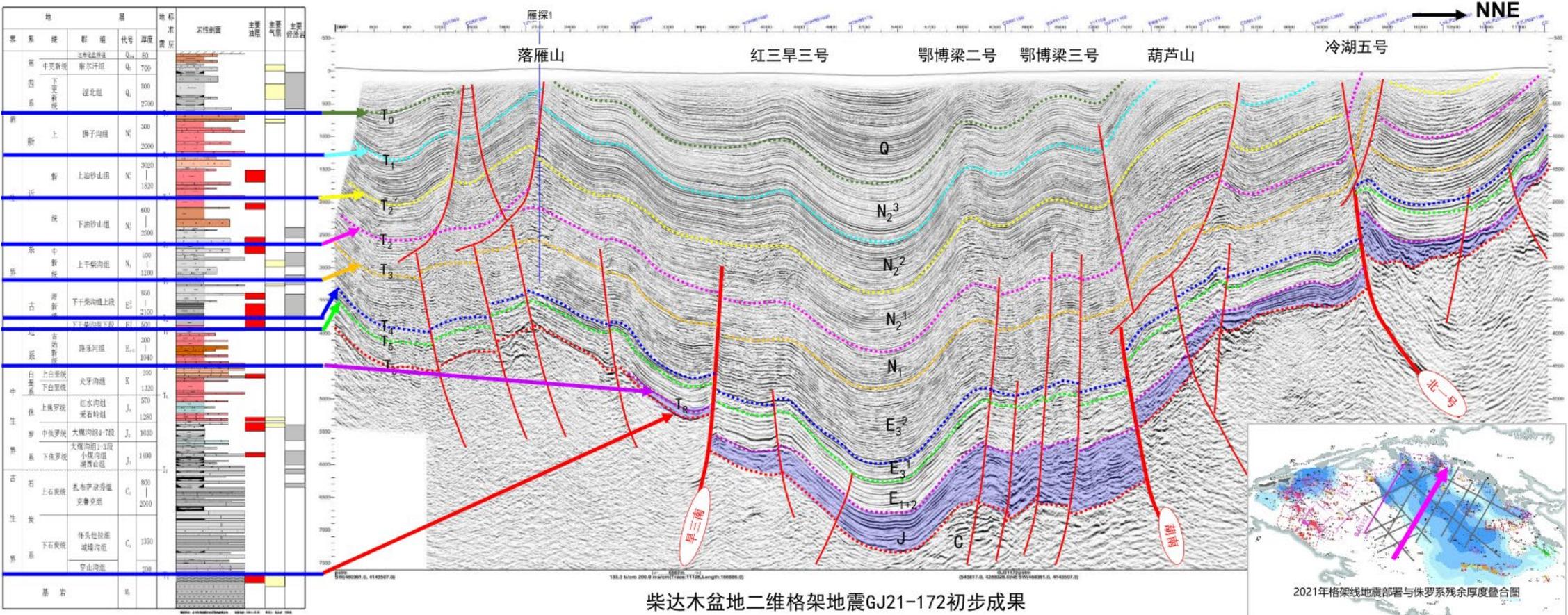
- **盆地位置:** 柴达木盆地位于青藏高原北部，为昆仑山、祁连山、阿尔金山所环抱的菱形山间高原盆地，地势西高东低、平均海拔3000m。高寒缺氧、地理环境特殊、地质条件复杂造就了其独特的石油地质特征。



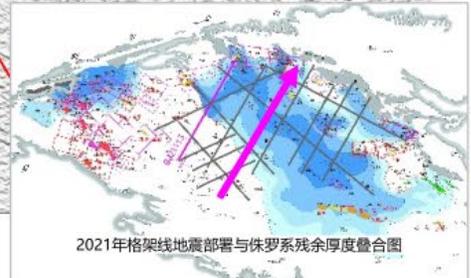
一、油气勘探的新领域及油藏描述的新需求

◆柴达木盆地:

◆ **沉积地层:** 沉积岩面积 $12.1 \times 10^4 \text{km}^2$, 最大厚度 $1.7 \times 10^4 \text{m}$, 纵向发育石炭系、侏罗系、白垩系、古近系、新近系及第四系六套地层。



柴达木盆地二维格架地震GJ21-172初步成果

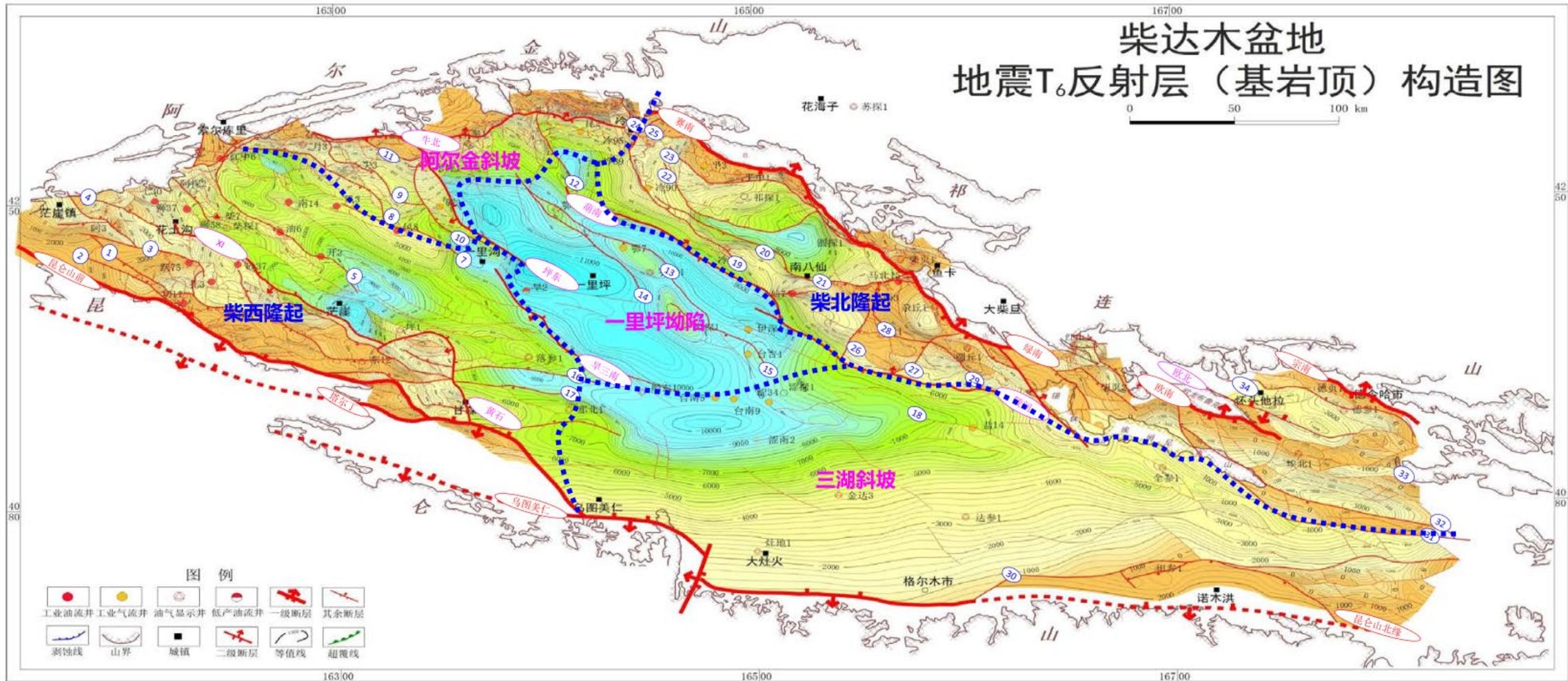


2021年格架线地震部署与侏罗系残余厚度叠合图

一、油气勘探的新领域及油藏描述的新需求

◆柴达木盆地:

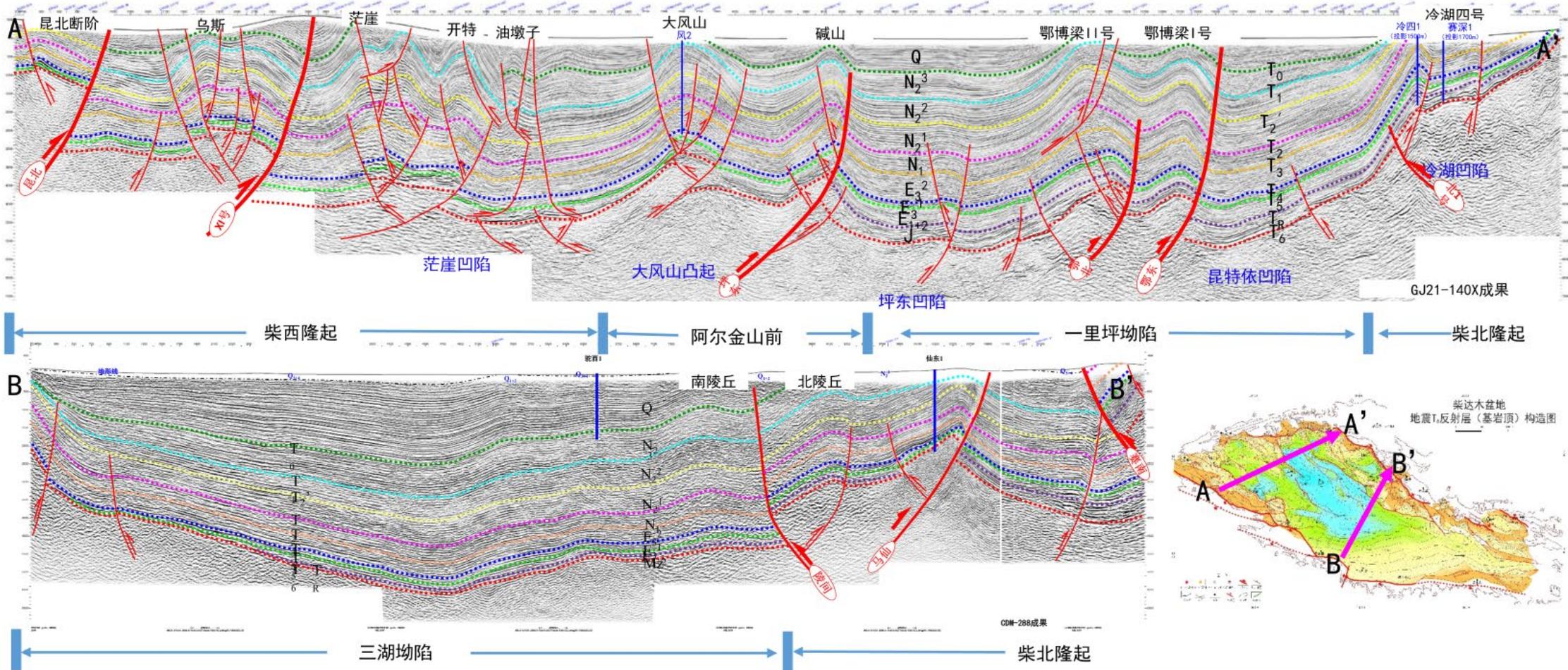
◆ **构造格局:** 从基岩顶面现今构造看, 基本呈现两大隆起、一大坳陷、两大斜坡的单元分区。



一、油气勘探的新领域及油藏描述的新需求

◆柴达木盆地:

◆ **地质结构:** 受两侧昆仑山、祁连山逆冲控制, 具有“南北对冲, 中央拗陷”特征。

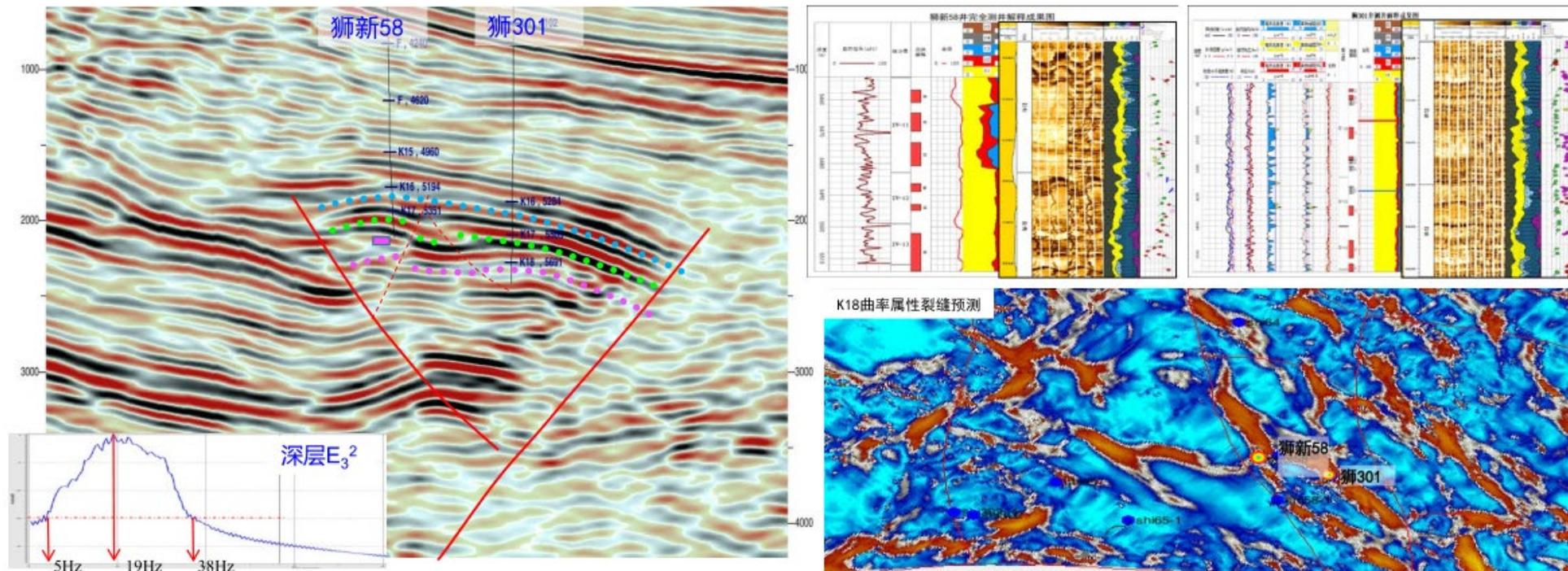


一、油气勘探的新领域及油藏描述的新需求

◆柴达木盆地:

挑战二：巨厚盐下湖相碳酸盐岩缝洞体精准刻画

◆ **问题描述：** 英雄岭英中狮新58井E₃²IV-13油层5502.0-5514.0m已累产油14.63 × 10⁴t，累产气6188.4 × 10⁴m³，基质储层物性、裂缝发育程度明显优于临近的狮301井（与狮新58井距离1km，试油日产油仅1.3m³）。

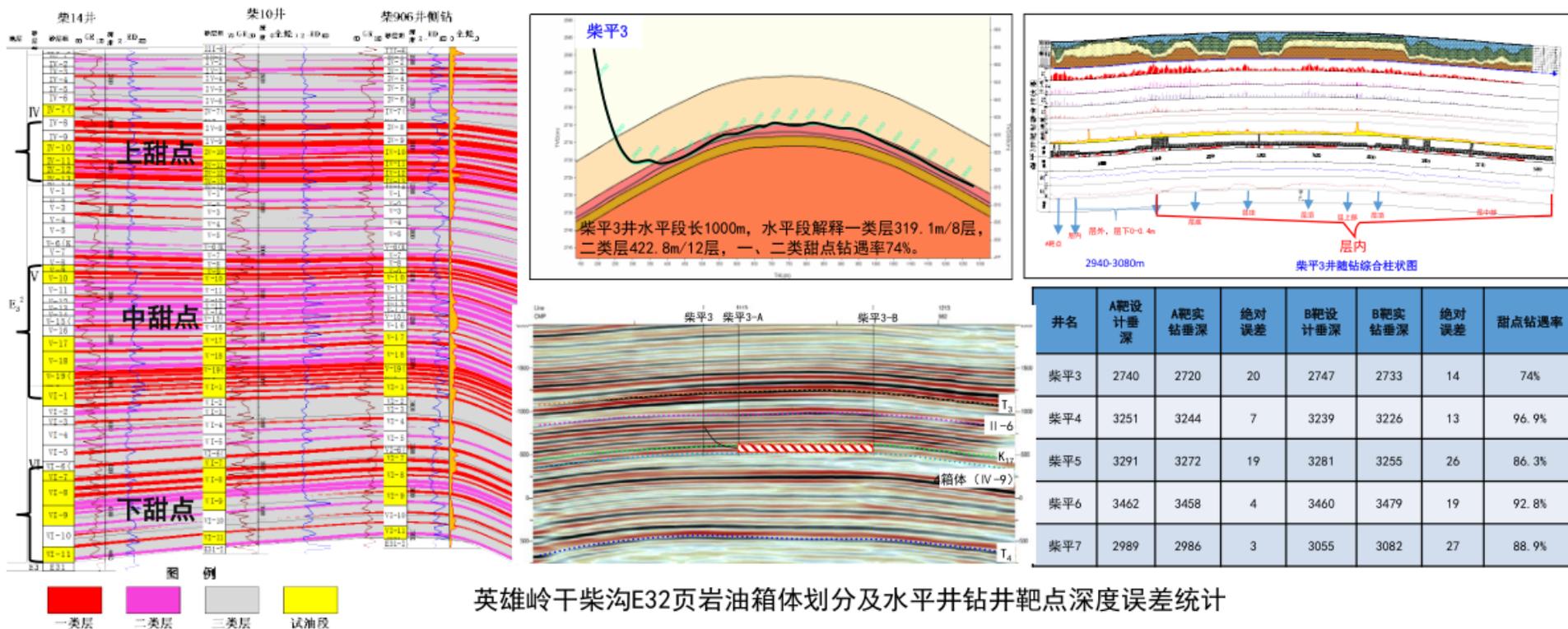


一、油气勘探的新领域及油藏描述的新需求

◆柴达木盆地:

挑战三：页岩油气4-5米黄金靶区空间形态准确表征及水平钻井实时导向

◆ **问题描述**：英雄岭干柴沟、黄瓜峁极低信噪比地区，叠前深度偏移资料不能支撑精细刻画5米级靶区低幅度构造形态，水平井钻进易脱靶，水平段甜点钻遇率较低，难以满足提高甜点钻遇率和单井产量的要求。



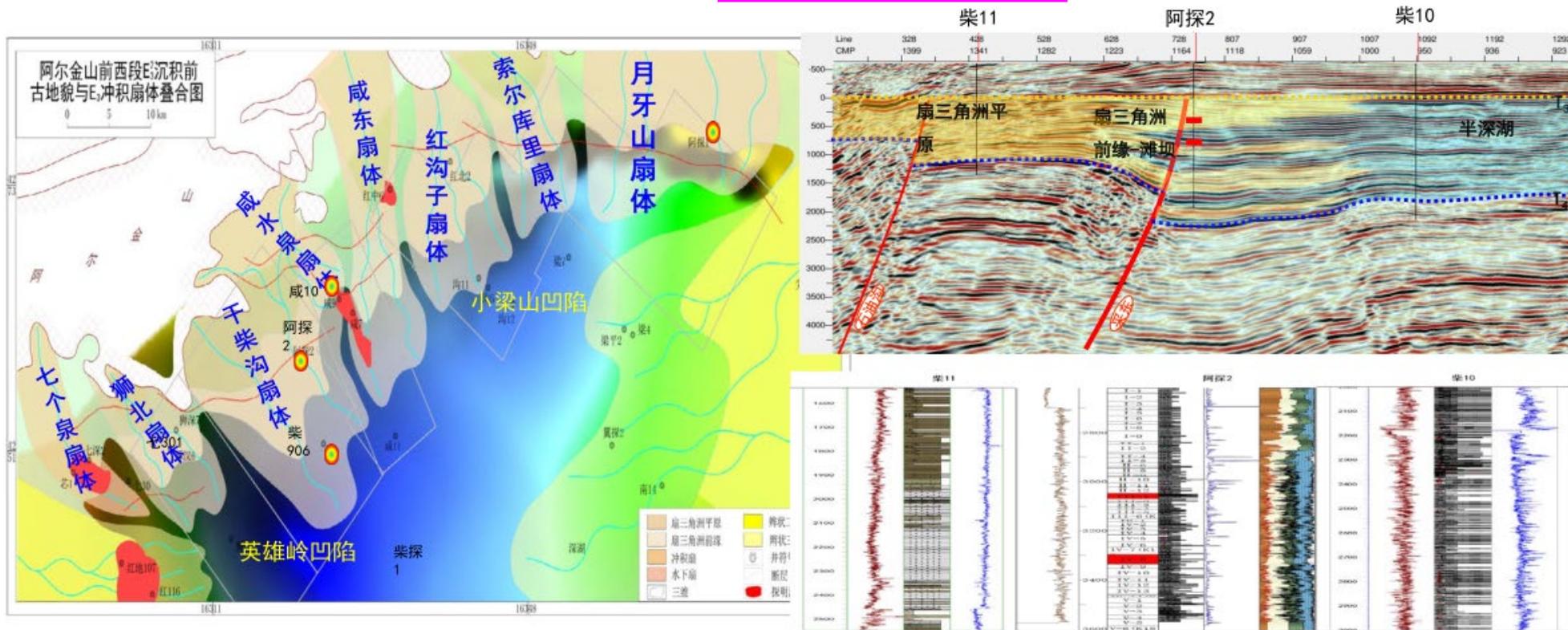
英雄岭干柴沟E32页岩油箱体划分及水平井钻井靶点深度误差统计

一、油气勘探的新领域及油藏描述的新需求

◆柴达木盆地:

挑战四：山前带近物源砂砾岩扇体储层预测及岩性圈闭刻画

◆ **问题描述：**阿尔金山前带古近系发育多个扇体，阿探1、阿探2、七301等探井见到了良好苗头，有望成为新的勘探领域。受限于三维资料品质，扇体有利相带预测、岩性圈闭刻画等与实钻符合率较低。





一、油气勘探的新领域及油藏描述的新需求

◆大致总结一下，我国油气勘探的新领域包括：

- ◆1、复杂地表+复杂地下结构的山前带及前山带勘探领域。
- ◆2、深层超深层（湖/海相碳酸盐岩）油气藏的勘探。
- ◆3、页岩气和页岩油的非常规油气勘探。



一、油气勘探的新领域及油藏描述的新需求

◆我国各大含油气盆地地震勘探遇到的成像问题，大致可总结为：

◆复杂地表+复杂地下结构情形下，构造成像问题也还没有得到很好地解决。

◆是当前PSDM技术不行或/和速度层析建模方法技术不行？

◆我觉得都不是（地震数据信噪比有保障的情况下，PSDM和TOMO效果应该是可以的）！主要是地震数据品质不行！构造极为复杂多变的探区，及时有高信噪比数据，速度建模和成像依然存在问题。叠前数据本身信息不足以解决此类问题。

◆深层超深层的构造成像与岩性成像都没有得到很好地解决。

◆我认为除了层间多次波压制与成像是个没有很好解决的问题外，其他的成像问题还是归结为地震数据问题，满足不了高精度成像的要求。叠前数据本身信息不足以解决此类问题。

◆成像分辨率问题是一直存在的至今没有得到很好解决的问题。

◆分辨率问题的本质是：用于反演成像的信息不足，反演问题存在多解性，目前的反演成像方法只能得到光滑的近似解。



一、油气勘探的新领域及油藏描述的新需求

- ◆ **我国各大含油气盆地地震勘探遇到的成像问题，大致可总结为：**
 - ◆ **反射系数为目标的地震波成像还没有进展到定量的、宽带的弹性参数的成像，满足不了隐蔽油气藏、非常规油气藏和岩性油气藏的（半）定量的精确描述的需求。**



一、油气勘探的新领域及油藏描述的新需求

◆油藏描述的新需求，也可以大致总结为：

◆描述油气藏，无非是三个主要方面：

- ◆对油气藏进行进行精确的三维空间中的定位；
- ◆对油气藏的几何（外形）形态（包括复杂油气藏的空间关系）进行精确的刻画；
- ◆对油气藏的岩石物性，尤其是孔隙度、含流体情况等，进行准确的、定量的预测。

◆无论是旧领域还是新领域油气勘探中的油藏描述，都是如此。

◆所谓的油藏描述的新需求，我认为是在“两宽一高”地震数据观测+FWI/LS_RTM新技术下，针对非常规油气藏、海相碳酸岩盐油气藏、地层岩性油气藏等描述所提出的。新需求希望反演成像给出**定量的、宽带的弹性参数成像结果（宽带纵波阻抗+宽带横波阻抗）**。在新的成像结果基础上，依然是解决油藏描述上述三方面的问题。

◆技术发展基本上就是一个螺旋上升的过程。

目录

- ◆一、油气勘探的新领域及油藏描述的新需求
- ◆二、地震波成像当前研究的发展困境
- ◆三、多元数据统计分析的思想与框架
- ◆四、多元信息表达与融合的合理理论框架
- ◆五、地震波成像研究的新导向
- ◆六、再论WPI当前侧重研究的问题

二、地震波成像当前研究的发展困境

- ◆从学术发展的角度看，地震波成像的科学问题是十分清楚的，即：已知观测的叠前地震数据（包括一些先验信息）和选定的波动方程，在Bayes参数估计理论体系下，进行弹性参数的估计或称为地震波反演成像。
 - ◆具体的Bayes参数估计理论下的FWI方法，以及它的问题、它的各种退化和变种，就不要再重述了！WPI的学生应该非常清楚了！
- ◆但是，这样的理论和方法技术的研究路线似乎走到了尽头！
- ◆看起来理论很完美，但实际问题得不到有效的解决。

二、地震波成像当前研究的发展困境

- ◆从岩石物理角度，更新广义胡克定律，导出新的波动方程，似乎可以从学术上继续推进地震波成像研究的不断深入。
- ◆李幼铭老师一伙人弄的广义波动方程考虑了应力应变关系的二阶项，得到了所谓的广义波动方程。但是，（勘探地震这种）弱应力情形下，应力应变关系应该可以做线性化处理的，目前的波动方程对实测波场还是有很好的预测结果的。广义波动方程的预测结果有物理实证的支持吗？目前没有找到！

二、地震波成像当前研究的发展困境

◆ 岩石物理机制首先是微观的，然后是中观和宏观的，不同尺度下，岩石物理关系是不同的。岩性参数与岩石模量之间到底是关系？不同的岩石物理关系对应不同的波动方程，不同频带的震源激发出的波场是非常不同的。地震波成像都退化到去研究这些基础的内容吗？即便研究清楚了，对油气勘探有多大的指导性？

◆ 英国统计学家George E. P. Box说：“All models are wrong, but some are useful.” 所有模型都是错的，但其中有些是有用的。"Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful."

二、地震波成像当前研究的发展困境

- ◆ 因此，从岩石物理角度，更新广义胡克定律，导出新的波动方程是正确的学术方向吗？能有颠覆性的创新吗？能颠覆就意味着目前线弹性理论的波动方程是存在重大缺陷的。
 - ◆ 这里面的重点是新的岩石物理关系。
- ◆ 另一方面，更复杂、更接近实际的波动方程适合用于参数反演吗？
 - ◆ 这个问题早就被CWI反演的逻辑给否定了。当然，我们的观点也是来源于总结提炼抽象了现有的反演理论下形成的。
- ◆ 这就产生了地震波成像方法技术的发展困境。学术上，到底应该研究什么？

目录

- ◆一、油气勘探的新领域及油藏描述的新需求
- ◆二、地震波成像当前研究的发展困境
- ◆三、多元数据统计分析的思想与框架
- ◆四、多元信息表达与融合的合理理论框架
- ◆五、地震波成像研究的新导向
- ◆六、再论WPI当前侧重研究的问题

三、多元数据统计分析的思想与框架

- ◆ 基于多元信息融合的、定量的宽带弹性参数反演成像是我认为的地震波成像的发展方向。
- ◆ 仅仅靠一个优化迭代算法得到由低波数、到中波数，再到高波数成分的宽带弹性参数估计结果是不现实的。
- ◆ 多元信息约束下生成尽可能高精度的先验约束模型是一个核心研究议题。
- ◆ 因此，有必要理清楚多元数据统计分析的基本思想。事实上，多元数据统计分析奠定了当前大数据时代所有领域数据分析的逻辑基础。

三、多元数据统计分析的思想与框架

◆多元数据的本质含义是什么？

- ◆多元数据是对一个或多个统计总体的抽样。这是统计学语言下的定义。
- ◆多元数据的合理定义：调查或统计某一现象全部数据的集合。某一现象就是我们关注的研究对象（对应统计总体）。
- ◆多元数据的具体形式：研究对象受若干 p 个（随机）因素的影响，每个（随机）因素随机采样 n 个结果，就形成一个或多个统计总体的随机事件具体实现结果的数字记录或采样。抽样结果形成一个或多个数据矩阵。
- ◆每个统计总体对应一个这样的数据矩阵。
- ◆多元统计分析就是对一个或多个多元数据矩阵进行分析，抽取其中蕴含的信息，然后进行决策。

三、多元数据统计分析的思想与框架

◆多元数据的本质含义是什么？

多元统计分析是以 p 个变量的 n 次观测数据所组成的数据矩阵

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \cdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \begin{matrix} p \text{ 个变量或影响因素} \\ n \text{ 个样品} \end{matrix}$$

$$X = [X_1, X_2, \cdots, X_p]$$

三、多元数据统计分析的思想与框架

◆多元数据统计分析的本质是什么？

- ◆ 我认为：利用对统计总体的抽样（样本）估计其（联合）概率分布函数和联合概率密度函数的具体形式及其中包含的参数，并基于此进行推断和判决是多元数据统计分析的最抽象的本质。
- ◆ 我认为：分析、寻找统计总体的影响变量之间的相互依赖关系是最本质的目的。最好是建立一个变量和另一个变量间的（回归）关系；一个变量和若干变量间的（回归）关系；多个变量和多个变量间的（回归）关系。利用这种（回归）关系进行推断与判决。
- ◆ 这种关系一般是线性（回归）关系，还是仅仅用二阶统计量建立的。
- ◆ 当前，多元数据统计分析基本上仅仅用到了这些随机的影响变量之间的一阶和二阶统计量。

三、多元数据统计分析的思想与框架

◆多元数据分析应该包含的内容：

- ◆1、依据统计总体的抽样（样本）估计其（联合）概率分布函数和（联合）概率密度函数的具体形式；
- ◆2、假设（联合）概率分布函数和（联合）概率密度函数的具体形式已知，估计其中包含的参数。
- ◆3、不同影响变量之间线性（回归）关系式的建立。最好是建立不同影响变量之间非线性（回归）关系式。
- ◆4、基于二阶统计量的影响变量之间相关关系的研究及应用。包括：分类分析（判别分析）、聚类分析、主成分分析、因子分析、对应分析、典型相关分析、偏最小二乘回归分析。都仅仅用到了统计总体的均值与方差。

三、多元数据统计分析的思想与框架

◆多元数据分析的具体内容:

- ◆**分类分析**: 已知若干 (n 个) 总体的一阶和二阶统计量, 对新加入的总体 (已知一阶和二阶统计量), 用一阶和二阶统计量的距离, 判决新加入的总体与已知若干 (n 个) 总体中的哪个距离最近, 从而实现分类判决。
- ◆**聚类分析**: 已知若干 (n 个) 总体的一阶和二阶统计量, 基于类内方差最小, 类间方差最大的原则, 把若干 (n 个) 总体分成几个子集, 从而实现聚类。
- ◆**主成分分析**: 把 p 个 (相关的) 影响因素通过某种线性组合实现解相关, 组合因素之间相互不相关。组合因素方差最大化是主成分分析的原则要求。
- ◆**因子分析**: 对 p 个 (相关的) 影响因素进行解相关, 并分析清楚影响因子与 p 个 (相关的) 影响因素之间的关联关系, 赋予影响因子以明确的实际意义。

三、多元数据统计分析的思想与框架

◆多元数据分析的内容：

- ◆**对应分析**：把 p 个影响因素和 n 个样品对等起来看，提取各自的主成分（数据标准化处理后，实际上二者的主成分是相同的），**一般地**按第一主成分和第二主成分展示在二维平面上，凸显不同的影响因素和不同的样品之间的关联关系。
- ◆**典型相关分析**：把单个总体的主成分分析推广到两个（A和B）总体上，在交叉组合因素方差最大化的条件下，进行各自的主成分分析。A的主成分分析结果受B的制约，B的主成分分析结果受A的制约。

三、多元数据统计分析的思想与框架

◆多元数据分析的内容：

◆线性回归分析：

◆包括单因变量多自变量线性回归关系式 $y = f(X)$ 和多因变量多自变量线性回归关系式 $Y = f(X)$ 两种情形。

◆核心是寻找“最佳”线性回归方程。

◆偏最小二乘线性回归分析：

◆在对多自变量进行主成分分析的基础上，建立单因变量多自变量线性回归关系式

$y = f(X)$ 和多因变量多自变量线性回归关系式 $Y = f(X)$ 。

三、多元数据统计分析的思想与框架

◆ 目前多元数据统计分析的局限性：

- ◆ 根据上述讨论可以清楚地看出目前多元数据统计分析存在明显的局限。我们可以肯定地知道：总体的 p 个影响因素之间可能存在复杂的**依赖关系**。而目前的分析基本上是在假设总体满足（或近似满足）Gauss分布，仅仅用一阶和二阶统计量就可以完整地刻画其统计特征。

相关关系？ 因果关系？

- ◆ 基于高阶统计量的、非线性的**依赖关系**的挖掘，然后用于推断，是多元数据统计分析的必然发展方向。事实上，最近几十年大数据、ML和AI的发展，也正是奠基在多元数据统计分析的这种发展上的。譬如非线性回归分析、非线性SVM，L1/L0范数下参数估计、非线性聚类（SOM）等等。

三、多元数据统计分析的思想与框架

◆ 勘探地震中多元数据分析方法的应用

- ◆ 广义地，勘探地震学就是由“数据”中提取出用于油藏描述的信息，对油藏含油气性进行推断，并做出钻井与否决策的学科。
- ◆ 以前，没有把勘探地震学放在多元数据分析理论框架下统一地审视。而是围绕“弹性参数估计” + “弹性参数映射成储层岩石物性参数”这两个核心问题展开的。
- ◆ 叠前地震数据采集 + 信号处理理论下数据预处理 + 偏移速度估计 Tomo + PSDM => 弹性参数估计
 - ◆ 这是 Bayes 参数估计，而不是 FWI 迭代算法
- ◆ AVA/FWI 弹性参数 + 岩石物理关系 + 储层岩石物性参数 => 油藏描述与推断
 - ◆ 这是 Bayes 统计推断，而不是深度神经网络算法

三、多元数据统计分析的思想与框架

◆ 勘探地震中多元数据分析方法的应用

- ◆ 主成分分析，用在信号建模基础上去噪、数据规则化等中；用在图像处理中。
- ◆ (广义) 线性回归分析，广泛地应用在各种信号与图像的建模过程中。
- ◆ PSDM、TOMO、FWI就是MLE (最大似然估计) 或MAP (后验概率最大化) 参数估计 (点估计方法) 。
 - ◆ 但是，这些点估计算法逐渐不合时宜。遗憾的是，目前区间估计还没有在勘探中得到应用。连Bootstrap这种数据驱动的点估计结果后评估方法也没有做。
 - ◆ 随机反演生成不同于MLE和MAP点估计解的做法也很少应用。
 - ◆ 这是值得我们关注的问题！

三、多元数据统计分析的思想与框架

◆ 勘探地震中多元数据分析方法的应用

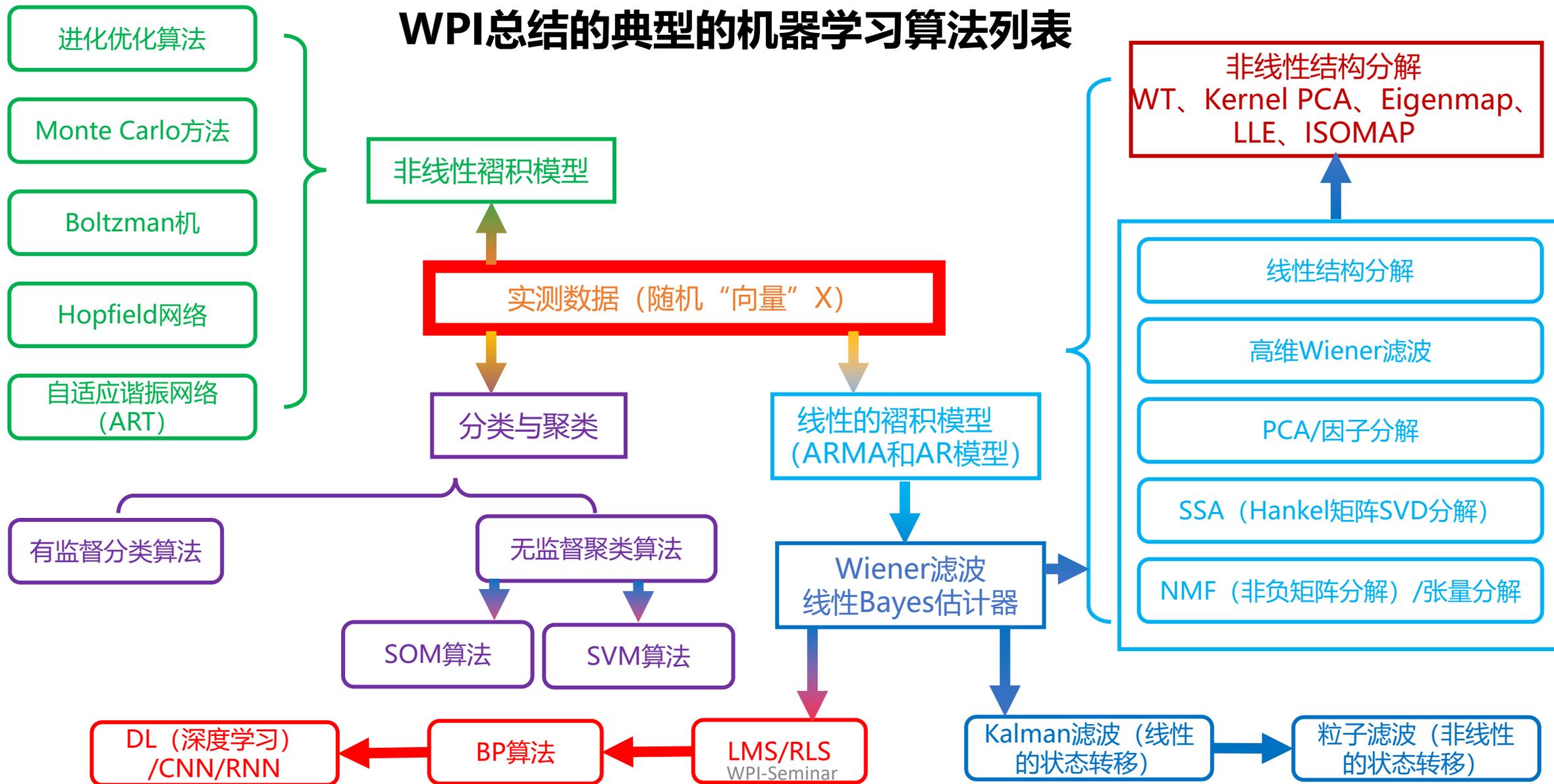
- ◆ 最近几年，由于先验信息的引入在高精度成像中的需求越来越大，分类（有监督）、聚类（无监督）方法，事实上包括各种机器学习算法，应用得越来越多。
- ◆ 但是，多因变量和多自变量之间的依赖关系分析，非线性的回归关系式构建等还是没有在岩石物理、地质知识的表达中得到深入应用。
- ◆ 目前的各种机器学习算法可以说都是在多元数据分析的理论基石上发展出来的。
- ◆ 井中获取的岩心测量数据、测井数据与岩石物性参数之间关系的挖掘，本质上这就是数据驱动地建立局部成立的岩石物理关系，以及扩展到三维空间中，提高宽带弹性参数估计的精度以及储层描述的精度，进而提升钻井决策的成功率是今后我们要特别关注的问题。

◆典型的机器学习算法

机器学习算法是在多元数据统计分析理论下发展出来的!



WPI总结的典型的机器学习算法列表



三、多元数据统计分析的思想与框架

◆ 高维数据分析与多元数据分析的对比---数据含义的对比

- ◆ 地震信号处理、图像处理、地震波成像，主要面对高维数据体。
- ◆ 各种井中数据、岩石物理测量数据，岩性物性参数测量数据，归属于多元数据。

◆ 统计学角度，多元数据是统计总体的抽样样本。具体形式：

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

- ◆ p 个影响因素（随机变量），每个抽 n 个具体采样值。

三、多元数据统计分析的思想与框架

◆高维数据分析与多元数据分析的对比---数据含义的对比

◆统计学角度，如何理解高维数据？

◆我认为不同的数据体有不同的含义解释。

◆譬如地震剖面，每个地震道对应一个影响因素（即一个震动随机事件/随机变量）。

它们之间存在（空间时间）相关性，据此蕴含有信息（反射信号）。

◆各种信号处理方法，据此产生。

◆对于空间数据，每个空间点对应一个影响因素（随机变量）。空间点随机变量之

间存在相关性。因此，蕴含有参数变化的空间信息。

◆各种空间变化规律的预测方法，由此产生。

◆对数据的认识，是发展合理的分析处理算法的基础。否则，后续的就不用谈了！

三、多元数据统计分析的思想与框架

- ◆ **高维数据分析与多元数据分析的对比---数据分析方式的对比**
 - ◆ **高维数据分析方法是多元数据分析方法的应用特例。只要把二者的数据含义能统一起来，多元数据分析方法的思想，甚至具体做法可以直接用于高维数据分析。**
 - ◆ **理解数据是数据分析最基础的事情！**
 - ◆ **再说一次，多元数据统计分析是大数据、ML和AI时代所有领域数据分析的共同基础。**

目录

- ◆一、油气勘探的新领域及油藏描述的新需求
- ◆二、地震波成像当前研究的发展困境
- ◆三、多元数据统计分析的思想与框架
- ◆四、多元信息表达与融合的合理理论框架
- ◆五、地震波成像研究的新导向
- ◆六、再论WPI当前侧重研究的问题

四、多元信息表达与融合的合理理论框架

- ◆无约束（或弱约束）优化框架下的地震波成像，即：TOMO、PSDM/RTM、FWI/LS_RT_M，无论各向同性介质或是各向异性介质，甚至粘弹性介质下的成像问题，从方法理论上讲，都已经研究成熟了（基本上很成熟了）。
- ◆但是，距离实际应用，还相差甚远。具体原因大家都应该是明瞭的。
- ◆今后，地震波成像的发展重点是定量的宽带弹性参数成像，更是要实用化的宽带弹性参数成像，能有效地促进油藏描述精度提升的成像。
- ◆当然，也可以把地震波成像延伸到储层岩性参数的估计阶段。
- ◆多信息表达与融合，产生对反演解的有效约束，使得反演结果有实用价值，能解决问题，就成了今后地震波成像技术研究的核心问题了。
- ◆勘探地震中的多信息，主要是来自井中的数据、岩石物理知识和地质知识。

四、多元信息表达与融合的合理理论框架

- ◆勘探地震中，岩石弹性参数和岩石物性参数都是空间变化的量。空间点上的多元数据包含了地下岩石介质的、各方面的信息。这是一个空间多元数据体。
- ◆这与一般多元数据是不同的。一般多元信息并不一定有空间属性（即空间相关性）。
- ◆一般地，物理多元数据是空间时间相关的。这是与非物理多元数据的显著不同点。
- ◆但是，无论是物理多元数据或是非物理多元数据，统计分析的本质还是找其中蕴含的相互依赖关系，最好能找到一个回归关系表达式来描述这种相互依赖关系。
- ◆一般地，其中蕴含的相互依赖关系的描述，就是用总体的数字特征，进而用总体样本的统计特征描述的。说白了，所谓统计特征无非就是一阶、二阶统计量。



四、多元信息表达与融合的合理理论框架

◆多元信息表达与融合的目的：

- ◆我认为：目的是依据现有的多元先验信息，生成一个能对反演解进行有效约束的初始先验模型。就是从多元先验信息中抽取出能对反演解进行有效约束的模型。

四、多元信息表达与融合的合理理论框架

◆多元信息表达与融合的基本思想和方法

- ◆最高层次的多元信息表达与融合的基本思想应该是Bayes推断，即把多元信息记为 $C = [C_1, C_2, \dots, C_p]$ ，把期望生成的模型记为 m 。假定已知后验概率密度函数为 $p(m|C)$ ，对期望生成模型的推断，就是在MAP准则下生成该期望模型，

$$m = \arg \max_{m \in \mathcal{M}} p(m|C)$$

- ◆这个抽象的理论框架是没有问题的，但是如何具体地进行勘探地震中的多元信息表达与融合？
- ◆我认为应分为：静态的（不随时间变的）和动态的（随时间变的）两种情形。

四、多元信息表达与融合的合理理论框架

◆ 静态的多元信息表达与融合-算法与要解决的问题

◆ 空间数据统计建模方法- Kriging插值方法

◆ Kriging插值方法不是或不完全是奠基在空间多元数据统计分析基础上发展出来的方法。

◆ 空间点之间的多元数据是统计相关的，所谓的空间数据统计建模应该是：充分利用空间点之间多元数据统计相关性，发展出更好的建模方法。目前，Kriging插值方法仅仅利用两点之间的变差函数进行预测模型构建，显然存在明显的理论缺陷。

◆ 径向基函数插值更是没有考虑空间多元数据本身内蕴的统计特征。用Laplace方程靠散乱数据支撑起合理的空间函数更是不合理的。

◆ 合理的选择是：在地质知识约束下（地震相约束下），考虑空间多元数据本身内蕴的统计特征，构造新的空间数据统计建模方法。地震相约束下的空间数据建模。

四、多元信息表达与融合的合理理论框架

◆ 静态的多元信息表达与融合-算法与要解决的问题

◆ 1、井中多元数据的统计分析与建模

- ◆ 因子分析+回归分析

- ◆ 我认为这是很值得发展的一个技术方向。在多井地区，是很有实用价值的。

◆ 2、基于成像剖面的地震相模型构建

- ◆ 层位识别---ML算法

- ◆ 地质知识表达---知识表达算法

- ◆ 地震相模型构建---??

◆ 3、两种信息融合下的地震岩相模型构建

- ◆ 地震相约束下的空间散乱数据插值---生成地震岩性模型

- ◆ 粗糙的背景速度和背景密度建模也属于此类

地质统计学建模的核心目的就应该是生成地震岩相模型。

四、多元信息表达与融合的合理理论框架

◆ 动态的多元信息表达与融合

- ◆ 这是对对应4D地震等随时间不断补充新的空间数据的情形的。在这样的情形下，如何依据新信息不断地调整已有的空间初始模型，得到最佳的目标模型，是动态的多元信息表达与融合的基本模式。
- ◆ 隐式Markov随机场的最佳演化是这种情形下问题的本质。寻找最佳的演化结果作为不断随时间调整的空间初始模型。
- ◆ 粒子滤波算法是代表性的方法。
- ◆ 动态的多元信息表达与融合算法应该以静态的多元信息表达与融合算法为基础。

目录

- ◆一、油气勘探的新领域及油藏描述的新需求
- ◆二、地震波成像当前研究的发展困境
- ◆三、多元数据统计分析的思想与框架
- ◆四、多元信息表达与融合的合理理论框架
- ◆五、地震波成像研究的新导向
- ◆六、再论WPI当前侧重研究的问题

五、地震波成像研究的新导向

◆地震波成像的现状:

◆速度层析成像

◆带限反射系数成像

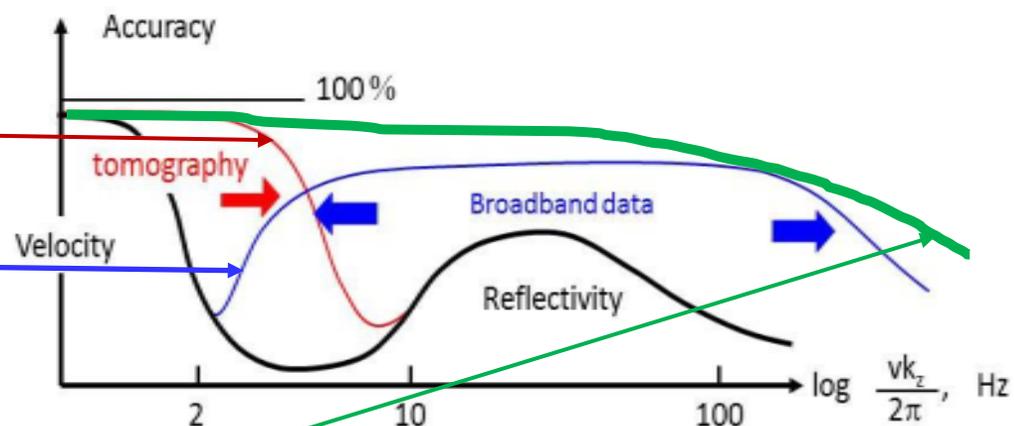


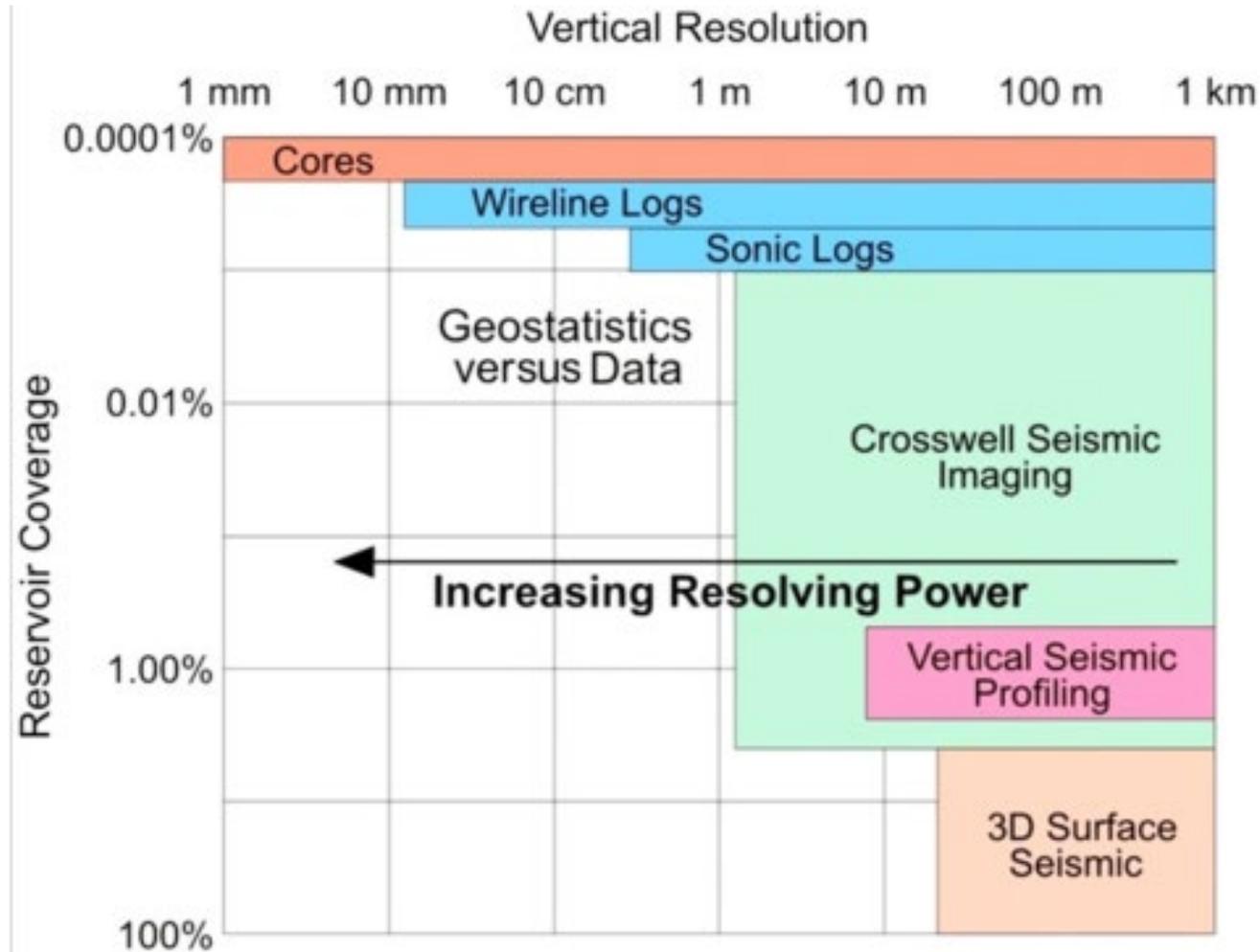
Figure 3: In black, the classic sketch by Claerbout (1985) illustrating the resolution we can expect from velocity analysis (“Velocity”) and imaging (“Reflectivity”). The gap between 2 Hz and 10 Hz is now filled by the improved resolution provided by high-resolution tomography and FWI (red) and the impact of broadband acquisition on imaging (blue). From Lambare et al. (2014).

◆地震波成像的今后导向:

◆定量的、宽带的弹性参数成像

◆期望是：宽带纵波阻抗+宽带横波阻抗。

五、地震波成像研究的新导向



Stratigraphic Reservoir Characterization for Petroleum Geologists Geophysicists, and Engineers.
(2th Edition) R. M. Slatt

五、地震波成像研究的新导向

◆破局的导向是什么？

- ◆ 1、Bayes参数估计的理论框架已经顶到天花板了，不可能再发展了。
- ◆ 2、Bayes参数估计的理论在勘探地震中的应用技术---FWI，尤其是FWI实用化，远远没有达到期望。
- ◆ 3、目前的重点不在于更新现有的线弹性理论下的波动方程。
- ◆ 4、破局的首要关键在于获得更满足FWI要求的叠前地震数据体。
 - ◆ 地震数据观测技术是油气地震勘探获得突破的核心。
- ◆ 5、破局的第二个关键点是：宽带弹性参数成像需要的先验信息的获取、表达和正则化应用。
 - ◆ 岩心测量数据、测井数据、岩石物理知识、地质知识是主要的先验信息源。
 - ◆ 多元信息约束下生成尽可能高精度的先验约束模型是重要议题。期望是一个地震岩相模型。

五、地震波成像研究的新导向

◆破局的导向是什么？

◆6、FWI实用化的具体技术路线和关键技术也是值得考究的。

- ◆满足“同因之果”关系的时差测量在背景速度估计中是核心问题。
- ◆RFWI的梯度分解及梯度正则化是核心问题。
- ◆保真定量无旁瓣反射系数的获取是核心问题。

◆7、把震源和检波器的物理机制放在波动方程正演模拟中是正确的导向，而不是可以追求更复杂的、预测结果更好些的波动方程。

- ◆震源和检波器的物理机制对实测波场的影响远超波动本身。
- ◆这个问题的研究目前很少见到。用地表一致性处理统一后就认为与实际情况一致了，真是这样吗？

五、地震波成像研究的新导向

◆破局的导向是什么？

◆8、把宽带弹性参数映射成储层岩石物性参数（孔隙度、流体饱和度等）是更为非线性性的问题。

◆描述储层物性参数的尺度是关键因素。

◆岩石物理关系是这个映射的正问题。尺度又是重要影响量。不同尺度下，有不同的岩石物理关系。

$$\text{弹性模量} = f^{\text{尺度}}(\text{岩石物性参数}) \quad \text{弹性参数} = f^{\text{尺度}}(\text{岩石物性参数})$$

◆DL算法进行数据驱动的这种映射，必须奠基在大量样本和岩石物理关系约束下。

◆把地震数据（无论叠前叠后）用DL算法映射成岩石物性参数的做法，太不符合逻辑了。人为地提一个更强的非线性反问题来求解。

◆我认为WPI的路线是：收集大量井数据，用多元统计分析方法，建立更合理的、当地化的岩石物理映射关系。基于此，把宽带弹性参数映射成储层岩性参数。一定把研究奠基在物理上。不迷信数学算法，算法就是工具。

目录

- ◆一、油气勘探的新领域及油藏描述的新需求
- ◆二、地震波成像当前研究的发展困境
- ◆三、多元数据统计分析的思想与框架
- ◆四、多元信息表达与融合的合理理论框架
- ◆五、地震波成像研究的新导向
- ◆六、再论WPI当前侧重研究的问题

六、再论WPI要侧重研究的问题

◆方向性的重点问题：

- ◆1、FWI的实用化问题，包括LS_RTM的实用化问题；
- ◆2、宽带弹性参数重构问题；
- ◆3、特征多次反射波预测、压制及反演成像问题。
- ◆4、井、震、岩石物理和地质知识的有效结合生成三维空间中的、尽可能准确的、有效的初始模型。
 - ◆在近地表建模中，数字地质露头模型生成，也是类似的问题。
- ◆5、高维数据的线性与非线性建模及在地震数据分析中的应用
 - ◆强噪音弱信号情形下的信号建模问题

四、再论WPI要侧重研究的问题

◆方向性的思想方法：

- ◆ 1、 Bayes推断的思想与方法；
- ◆ 2、 Bayes参数估计的思想与方法；
- ◆ 3、 高维数据分析与多元统计分析的思想与方法；
- ◆ 4、 空间数据分析的思想与方法；
- ◆ 5、 动态空间数据分析的思想与方法。

到了以地震信号分析为主转到以多元地震数据分析为核心的时候了。地震数据分析涉及的范围要宽泛很多，更针对油藏描述的需求。

四、再论WPI要侧重研究的问题

◆ 更具体的研究问题：

- ◆ 1、数字地质露头模型生成；
- ◆ 2、水体模型及含海底附近地层的水体模型生成；
- ◆ 3、复杂水体相关噪音的压制；
- ◆ 4、复杂地表相关噪声的压制；
- ◆ 5、高维数据规则化；
- ◆ 6、特征反射层相关多次波预测、压制及反演成像；
- ◆ 7、陆上透射波FWI+复杂条件下的初至波识别；
- ◆ 8、陆上反射波FWI实用化；
- ◆ 9、非线性时差检测。

四、再论WPI要侧重研究的问题

◆ 更具体的研究问题：

- ◆ 10、成像道集最佳照明优选；
- ◆ 11、LS_RTM的实用化问题；
- ◆ 12、井中数据的多元统计分析建立数据驱动的、本地化的岩石物理关系。
- ◆ 12、井、震、岩石物理和地质知识的有效结合生成三维空间初始模型；
- ◆ 13、宽带弹性参数（波阻抗）重构；
- ◆ 14、宽带弹性参数（波阻抗）映射成岩石物性参数。



谢谢
欢迎批评指正