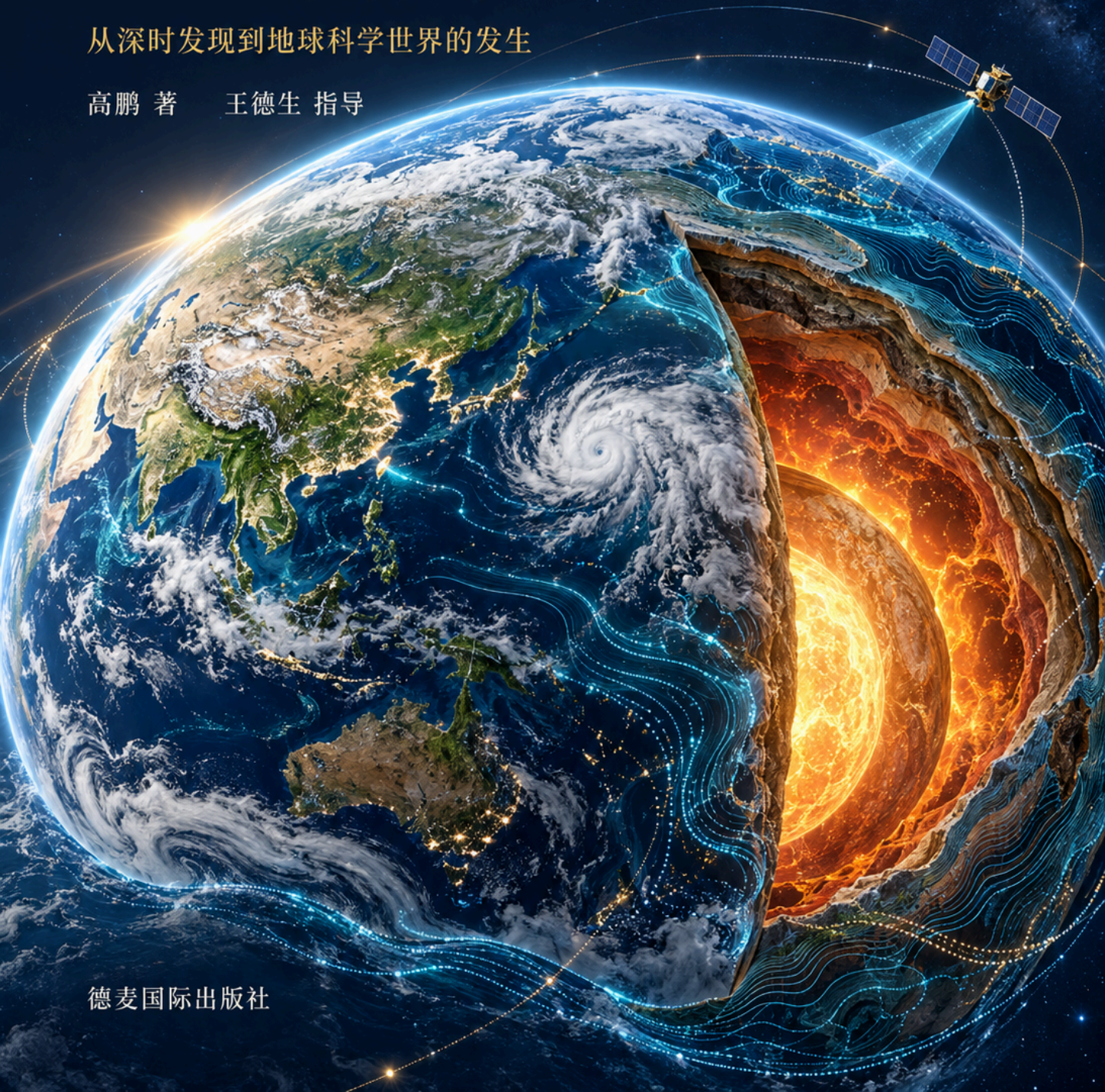
德麦国际专著 · 第375号

地球科学思想简史

从深时发现到地球科学世界的发生

高鹏 著　　王德生 指导



德麦国际出版社

出版信息

地球科学思想简史

从深时发现到地球科学世界的发生

作者　高鹏

指导老师　王德生

系列　德麦国际专著

卷号　第 375 号

ISBN　979-8-90690-586-4

定价　US$20.00

出版　德麦国际出版社 · Demai International Press

版本　统稿精修版 v1.1 · 2026 年 10 月

开本　190 × 250 毫米

本版保留完整正文、附录、参考资料与材料范围说明。

阅读说明

本书按照十二编六十章组织论述，章内保留四节结构，导论与结语分别承担问题建立和综合反思。正文中的历史材料、现代教学推演与后续研究建议分别标明，文献编号对应书末来源。

本版为统稿精修稿，保持十二编六十章的论述结构，统一术语、证据身份和章际衔接。原始资料的独立复算、部分历史档案和图件的逐项核验尚有待完成；具体限制在相关论证与书末材料记录中说明。本版以正文论证为主，原典图版的重制与逐图校核尚未全部完成。本次精修不改变上述材料范围，也不替代外部同行评议与出版终审。

正文统计为 309,511 汉字，其中导论 7,944 汉字，六十章 297,505 汉字，结语 4,062 汉字。统计排除各级标题、目录、参考资料与附录；数字、外文和标点不计入汉字。

目录

导论 从一块岩石进入地球科学的世界.....1

第 1 编 地球科学世界怎样成为研究对象

第 1 章 同一地球，怎样发生出不同的科学世界.....12
第 2 章 边界怎样从地层接触面成为共同标准.....16
第 3 章 深时、测量时与知识时：四条时间线.....21
第 4 章 从地方露头到全国地图：外推需要哪些许可.....26
第 5 章 怎样写一次思想转折：证据、操作与回写.....31

第 2 编 地层化石与可比较的过去

第 6 章 露头如何变成证据：吉尔伯特的调查自由.....37
第 7 章 斯特诺的层序规则：从空间接触推断时间.....42
第 8 章 化石从标本变成跨地区的时间标记.....49
第 9 章 1815 年地质图：颜色、底图与一个可流通的地层世界.....56
第 10 章 不整合：怎样把缺失读成地球历史.....65

第 3 编 深时间与地球年龄的科学生成

第 11 章 地球年龄之争：从历史年表到开尔文的冷却时钟.....73
第 12 章 渐变与灾变：把两派口号拆成可检验的问题.....80
第 13 章 从霍姆斯到煤山 矿物时钟怎样取得事件身份.....88
第 14 章 帕特森的地球年龄：为什么要离开地球岩石.....95
第 15 章 金钉子与年代标尺 地方剖面怎样承担全球责任.....103

第 4 编 地表测绘与空间尺度

第 16 章 山地不再只是风景：吉尔伯特怎样分解地形.....110
第 17 章 地图投影：一种空间为何需要多种合法表达.....115
第 18 章 尺度怎样改变可问的问题，而不改变山河本身.....121
第 19 章 地貌循环与动态平衡：两种解释怎样接受同一地形检验.....127
第 20 章 丁文江的下扬子调查：地方文献怎样成为海岸变迁证据.....133

第 5 编 大陆移动争论与海洋证据条件

第 21 章 海岸线相似：从拼图直觉到可检验的大陆组合.....139
第 22 章 魏格纳的大陆漂移：证据综合为何没有立即形成共识.....145
第 23 章 固定大陆学说：局部有效为何不能保证全球成立.....153
第 24 章 海洋观测基础设施：新证据为什么终于能够进入争论.....160
第 25 章 未被接受的思想如何存活：延续、修订与真正的失败.....167

第 6 编 海底扩张与板块世界的发生

第 26 章 海底地形：测深剖面怎样生成新的地球骨架.....175
第 27 章 海底磁条带 从异常信号到时间与空间的对应.....182
第 28 章 海底扩张：一个假说怎样承担多项独立检验.....192
第 29 章 转换断层 新运动关系怎样改变边界的含义.....200
第 30 章 板块世界的成立：运动学闭合、共同检验与新任务.....208

第 7 编 深部地球的间接显露

第 31 章 地震波与分层地球：莫霍、古登堡、莱曼的三个节点.....218
第 32 章 高压岩石实验：实验室材料如何获得地幔意义.....226
第 33 章 重力与磁场：多解性如何约束地下世界的生成.....232
第 34 章 内部热与地幔动力：能量来源怎样成为运动解释.....238
第 35 章 深部地球的公共事实：参考模型怎样稳定又怎样开放.....244

第 8 编 天气的动力学与可预报性

第 36 章 共同的天气怎样形成：测站、时刻与可比读数.....251
第 37 章 天气图、气旋与锋面：从压力分布到生命史模型.....257
第 38 章 从物理纲领到电子预报：1904、1922 与 1950.....264
第 39 章 洛伦兹 1963：确定方程为什么不保证长期轨道预测.....273
第 40 章 资料同化：观测与模型怎样共同生成分析场.....281

第 9 编 流域海洋与跨圈层交换

第 41 章 流域不是一个现成的容器：研究边界的形成.....288
第 42 章 从入渗到状态记忆：降水怎样成为径流.....294
第 43 章 从温盐水团到风驱环流：海洋对象的两种形成.....302

第 44 章 水循环从一幅图到可核算的交换系统.....310
第 45 章 任务改变，水文对象是否必须改变.....317

第 10 编 古气候生命与地球系统

第 46 章 过去的气候怎样留在树轮、冰芯与海底.....324
第 47 章 1976 年的轨道论证：多种记录怎样指向同一段过去.....332
第 48 章 1967 年的辐射—对流模型：气候反馈怎样成为问题.....340
第 49 章 氧化世界的形成：生命、岩石与大气如何共同进入地史.....348
第 50 章 1988 年的地球系统纲领：一张整体图怎样组织研究.....357

第 11 编 遥感算法与行动中的地球知识

第 51 章 同一块地表如何跨越半个世纪的卫星记录.....367
第 52 章 Destination Earth：数字孪生究竟复制了什么.....373
第 53 章 AI 标出的地震，何时成为共同体承认的事件.....380
第 54 章 雷道布特火山与航空预警：知识怎样进入行动.....388
第 55 章 版本也是证据：从 Landsat 重处理到 AI 目录更新.....394

第 12 编 发生学综合与研究边界

第 56 章 不能重演的过去，怎样接受可以重做的检验.....401
第 57 章 三方程的增量在哪里：从提示框架到可败解释.....406
第 58 章 六路径与三原理：比较转折，不替历史排队.....412
第 59 章 两项承重论证：样品年龄与板块对象的形成.....418
第 60 章 开放的地球：下一轮问题从哪里发生.....424
结语 在稳定的知识中保留继续发生的能力.....429
附录一至三.....434
附录四 洛伦兹系统的教学计算.....437
参考资料与来源索引.....439
补充来源与材料范围.....447

导论 从一块岩石进入地球科学的世界

◆ 导论一 岩石提出的不止一个年龄问题

把一块岩石放在桌上，人们很容易问它有多老。这个问题仿佛只缺一个数字，等到技术足够精密，岩石便会交出答案。然而，我们稍稍追问，就会发现问题的对象还没有完全确定。我们问的是矿物结晶的时间，岩石堆积的时间，岩体冷却的时间，还是其中生物死亡的时间？如果岩石后来经历加热、变形和流体作用，这些事件是否留下另一套时间记录？所谓岩石的年龄，有时是一个有用的简称，有时却把几种不同过程藏进同一个名词。

设想一层含有碎屑颗粒的沉积岩。颗粒可能来自更老的岩体，先经历风化，再被搬运到此处。沉积以后，新的胶结物把颗粒结合起来；再后来，裂隙中又生成其他矿物。我们从中取出某颗矿物，得到一个看似精确的年龄，并没有因此完成整块岩石的历史。那个数值必须获得事件身份：它约束哪一个过程，又凭什么能够约束这个过程？这段设想是对地质推理的简化说明，不对应本书新采集的样品。它的作用是让读者看见，一个数字从仪器结果进入地史叙述，还需要跨越哪些连接。

连接不能靠语言的便利来完成。矿物与围岩是否同时形成，要由产状、结构与成分等材料约束；测量是否记录形成过程，要考虑体系的保存与改变；目标事件与取样层位之间的关系，也需要独立交代。如果这些问题没有解决，增加小数位数可能只会让一个归属不清的结果显得更加权威。反过来，一个不很精确、但对象明确的年龄区间，可能已经足以排除某种因果顺序。

本书的一个主要案例来自煤山剖面。这里的时代界线、化石出现、火山灰层定年和主要灭绝区间，经常出现在同一幅剖面图中，却不能互相替代。国际地层委员会登记的印杜阶底界位于煤山 D 剖面 27c 层底，主要生物标志为 Hindeodus parvus 的首次出现，批准年份为 2001 年。这个批准年份属于共同标准的历史，绝不是该岩层的形成时间。[K01] 图上的一个点、一段层位和一组测年数值，各自回答不同的问题。正是这些答案能够受约束地连接起来，煤山才成为超出单一地点的研究节点。

我们不在导论提前展开全部定年技术。这里先保留一个要求：读者见到年龄时，应当能够顺着论证返回样品、返回操作、返回事件定义。若这条返回路线断裂，数字即使来自可靠仪器，也还没有获得它在论述中被赋予的全部意义。思想史不仅记录更大的年数怎样取代较小的年数，也记录人们如何学会区分不同的计时对象，如何设计能够支持这些区别的操作。

同一块岩石还可以进入另外的世界。在地质图中，它被放入某一地层单位，成为连接相邻露头的依据；在工程调查中，它可能成为具有特定变形和渗流性质的材料；在古环境研究中，它的颗粒、化石和化学组成被用于约束过去的条件。这里并不是任意选择一种说法。不同问题改变取样尺度、允许的比较、需要排除的误差，也改变研究者必须具有的辨认能力。

假如一位初学者只能辨别颜色深浅，他和一位能够识别层序倒转、次生脉体以及化石保存状态的地质学家，面对的物理位置可以相同，实际开展的观察却不同。后者不是在一份现成而完整的信息表中多读取了几行。他能够提出的区分，需要学习、工具、操作和可被他人纠正的规则共同形成。这种辨识能力也不是主体内部孤立增长的能力。离开标本、剖面、比较材料和共同验证，它便失去发挥作用的条件。

在 SDE 框架下，本书把这些对象属性、辨识能力和操作关系的共同发生，称为地球科学世界的发生。它不是把自然界改写成人的任意产物，也不满足于说人们更换了看待世界的方式。新的研究复合体能够稳定地区分原来混在一起的事件，连接原来难以比较的记录，发出原来不能接受检验的判断。这些新增能力构成我们必须解释的实质。

有时，变化确实只发生在精度上。同一种事件已经能够识别，新的仪器让误差范围收窄，这是一项重要改进，却不必被命名为新世界的诞生。有时，新的精度使过去不可区分的两个过程得以分开，原先的一个年龄对象由此被重构为有先后的事件关系。这时，仪器进步与对象发生之间存在联系，但联系仍须逐步论证。不能因为仪器更新就宣布本体更新，也不能因为对象名称相同就断定它的规定性从未改变。

这一分辨将贯穿全书。地层、板块、气团、水团、流域以及数字地球，并不是同一类对象。它们的历史也不能被压成一个不断放大的发现故事。我们寻找的是不同案例之间可以比较的发生环节，而非一条足以吞并所有差异的总公式。岩石是入口，因为它最容易提醒我们：看得见的材料与可以据以论证的对象之间，存在需要认真解释的工作。

由此，思想史的解释需要同时交代材料的范围与推断的步骤：哪些关系已经得到支持，哪些连接仍需补证，哪些属于帮助理解的现代推演。

◆ 导论二 四条时间线与思想史的责任

地球科学研究过去，但研究本身也有过去。沉积物形成于某个地质时期，今天从剖面采得样品，几年后出现新的解释，更晚才有共同标准。这几个时间可以写在同一页，却不能被当成一条没有层次的时间线。若把标准确立日期当作发现日期，就会遗漏先前的争论；若把理论发表日期当作其全部证据已经齐备的日期，就会将后来的验证偷偷交给早期研究者。

第一条时间线属于自然过程。它包括沉积、结晶、侵蚀、海底运动、大气演变与生命活动。第二条属于记录和操作：何时取样，何时测量，使用什么位置、仪器和处理程序。第三条属于公开论证：什么观点在何时以什么材料被提出，哪些反对意见进入了讨论。第四条属于共同使用和标准：某种界线、模型、数据产品或操作规范在什么条件下获得持续采用。

四条线并不总能得到同样精确的日期。发表时间可能准确到日，某次早期采样的时间却只知道季节；自然过程可能只能落在一个区间，标准采用又可能经历多年过渡。思想史的诚实，不在于把所有空白填满，而在于不同精度的信息仍能支持哪些判断。如果一条关键因果链需要先后顺序，而现有资料无法确定这个顺序，解释就不能靠叙事流畅来补足。

板块研究提供了鲜明例子。1963 年的海底磁异常解释、1965 年的转换断层概念与 1967 年的震源机制检验，属于相互联系而不能合并的证据时点。[K04—K06] 后来的地震结果可以增强早期模型的可信度，却不能被写成早期作者当时提出模型的现成理由。历史分析要同时做到两件事：理解模型后来何以成立，也保存模型初次提出时实际承担的不确定性。

这样处理，并不会削弱科学家的创造。恰恰相反，只有保留当时尚缺少什么，才能理解提出新关系需要跨越什么困难。若把后来积累的结果全部摆到早期人物面前，任何转折都会显得显而易见。英雄叙事表面上赞美个人，实际上可能抹掉他的判断空间，因为它把创造写成了对后世答案的提前抄录。

反方向的误读同样常见。一项早期理论中存在错误的机制，不代表它提出的全部比较都没有价值；后来的理论取得成功，也不代表早期异议一概出于保守。关于大陆移动的争论涉及几何拟合、古生物与地层联系、运动机制以及观测可及性。不同问题拥有不同的证据质量。思想史必须把这些理由分别安放，才可能理解某种主张为什么在一处富有启发，在另一处难以接受。

所谓当时的可知条件，也不是把时代想象成一个内部完全一致的盒子。不同研究者拥有不同航线、标本、仪器、语言和发表渠道。同一时期可以存在不对称的信息分布。某项材料已经被某人记录，不等于它已经成为另一个研究群体可调用的公共证据。记录存在、公开可见、能够比较与能够承担判断，是不同层次的状态。

地质图特别适合展示这些差别。野外露头是局部的，图面却常常连续。图把位置、颜色和层序连接起来，使分散观察进入同一个比较空间。它的公开传播可能让后来的人沿着色带继续调查，也可能暴露先前不明显的冲突。地图的出版因而既是成果出现的时点，也是新的研究条件形成的时点。但是，我们要证明某幅图实际改变了调查路线，还需要后续使用记录，不能只由地图具有这种可能性推出历史已经如此发生。

本书称这种后续作用为回写。回写并非把过去重新造成一遍，而是已经形成的知识改变接下来的操作条件。一个被接受的地层对应关系可能影响采样目标；新的测量又可能要求修改该关系。若只有教材措辞改变而调查与判断方式不变，那是一种传播变化，未必构成对象规定性的重组。为了不把不同变化混在一起，本书将分别讨论理论回写、操作回写和公共规范的维护。

思想史的编排还涉及另一种时间错觉。按章节先后阅读，很容易以为所有学科都按同一顺序进步。本书从地层进入深时，再转向空间、板块、深部、天气和地球系统，主要依据问题之间的联系。这不是一份宣称各领域同步转折的编年表。海洋测量、气象观测和地质调查经常并行发展，所依赖的技术与制度也彼此交叉。章节顺序帮助读者理解，不能替代历史顺序。

我们因而需要两种阅读眼光。一种顺着本书的问题链，从局部记录走向多尺度系统；另一种随时返回案例内部，检查事件、操作、论证和共同使用是否被正确区分。前者提供全书的连续性，后者阻止连续性成为制造历史必然性的工具。二者之间的张力，正是写作一部思想史应当保留的困难。

◆ 导论三 怎样使发生学分析承担证据责任

本书用 S、D、E 三个符号组织上述问题。S 指正在形成或维持的显露，D 指具有承接关系的差异序列，E 指相关特征彼此规定其成立条件的纠缠。三方程分别写作 S＝F(D,E)、D＝G(S,E)、E＝H(S,D)。它们在这里安排三种相连的提问：某种可辨认对象如何成立；它如何影响下一步操作；操作与结果又怎样改变后续成立条件。[M01]

这是一套研究框架，不是衰变定律、流体方程或地震波传播方程的替代品。矿物年龄要由适当的物理化学模型估计，气象预报要受动力学与资料条件约束。发生学分析关心这些计算中的对象、变量、比较和使用条件怎样形成，不能凭一个一般关系式算出实际年龄或天气。把元层次的问题与领域模型混同，会同时削弱二者。

为了让这种区别具体可见，我们构造一个地层对比的教学情境。两个地点都有浅色岩层，研究者首先怀疑它们属于同一地层单位。颜色相近只支持一个很弱的起点，因为不同时间也可能形成相似岩性。随后，研究者比较上下层序与化石组合，发现某些排列关系相符。新的对应关系使他把第三个地点选为检验位置，而那里的一组材料又迫使他区分原先合并的两个单位。这个情境并非某位历史人物的实际日志，而是用于拆解推断责任的例子。

在情境的第一阶段，S 可以是尚待检验的跨地点地层对应，D 是露头之间、上下层位之间的比较序列。E 不能只写成地质环境、仪器和人员，因为这样的清单没有解释任何关键连接。应当具体写出：化石组合的身份限制层位对比，层位关系又约束某个化石出现能够表示什么；保存状态改变缺失的意义，而预期层序又帮助判断需要重查什么位置。只有这些相互限定被说清楚，E 才承担了分析责任。

第二阶段，已形成的对应关系选择下一步比较。如果第三地点能够在两个竞争判断之间给出不同预期，去那里调查就有明确的判别价值。此时 D＝G(S,E)关注的不是未来结果神秘地改变过去，而是当下的目标与已有条件如何筛选可行行动。我们可以根据模型提出这种选择建议；要把它写成历史事实，则必须找到行动者确实选择该地点的记录。方法上的合理路线与实际走过的路线不能互换。

第三阶段，新地点改变原先对应关系的成立条件。一个曾被当作同层证据的相似特征，现在需要分成不同类型；旧标本可能因此被重新归类，旧图例也可能需要调整。这是

E＝H(S,D)要求追踪的内容。但如果新材料只是落在原来的预测区间中，分类和规则无需改变，那么这次结果主要加强已有关系，而不必宣布条件发生全面重构。

这个教学例子也给出了失败条件。如果所有地点都按预期归类，而预期本身又决定哪些样品被保存，所谓成功就可能来自筛选。如果任何不一致都能通过临时修改分类而被吸收，模型便失去承担风险的能力。如果档案只记录最后的地图，没有中间操作，我们也不能从漂亮的终态图倒推出全部回写已经发生。发生学解释必须允许材料阻止它完成闭环。

本书将六路径用于比较研究入口：S→D→E、S→E→D、D→S→E、D→E→S、E→S→D、E→D→S。这些排列不是六种科学家的固定性格，也不是每个发现必须经过的心理步骤。一次研究可能从异常读数开始，另一项可能从已有模型要求检验开始，还有一些从分类条件彼此不相容开始。路径帮助我们检查从哪里进入问题，以及转换到另一维时有什么证据；它不准许我们凭叙述顺序重建人物意识。

在历史材料不足时，最稳妥的说法是本书采用某条复审路线，而不是历史以某条路线发生。复审路线可以反向进行：先拿到一个成熟结论，再追问它依赖哪些比较，最后检查那些比较的共同条件。历史发生的过程却可能多线并行，包含失败、偶然相遇和未公开的中间尝试。分析的清楚与历史的复杂不必相互排斥，只要不把前者伪装成后者。

三原理进一步提出相位问题。本书以 P1、P2、P3 提示两维之间的张力如何推动另一维改变，并追踪这种改变的回写。这些分析代号的内容，要由案例给出：方向预测相反、样品归属冲突、尺度不匹配，以及任务要求超出现有分辨率，分别要求不同的处理。相位分析只有进入这些具体联系，才可能说明变化的机制；它本身不构成额外的地学规律。

更不能把每次困难都解释为创造的必然前奏。异常可以被校准修正，目标可以被合理降低，旧分类也可能经受住新材料。失败的仪器、错误的样品标签与真正挑战对象身份的结果，具有不同意义。发生学研究若只挑选成功革命，就会失去判断什么时候需要改变对象、什么时候只需要修复操作的能力。

本书因此在强常规解释旁边安放 SDE 解释。例如，地层学本来就知道样品年龄不等于全部地层事件的年龄，误差分析本来就处理共同校准的问题，历史科学的方法论也早已重视竞争假说与判别证据。不能先把这些解释压缩成机械的资料积累，再宣布发生学发现了它们从未考虑的事情。比较必须使用同一组材料，并给予对方完整表达其能力的机会。

SDE 分析可能提供的增量，需要落在可检查的地方：是否定位了先前被混合的对象身份；是否说明某个结果怎样改变下一轮操作；是否发现表面独立的证据共享同一筛选条件；是否在新的材料面前提出了不同而可执行的复核任务。若这些事情已经由既有分析充分完成，本书应当承认自己提供的是统一组织和教学上的价值，而不夸大成独立历史机制的发现。

这种承认不会使研究失去意义。一部跨学科思想史本来就需要能够迁移的问题语言。关键在于迁移不能消除领域差别。矿物的封闭性、海磁记录的几何与气象资料的时空代表性，不是一个词能够替换的同一种条件。统一框架的工作，是让读者看见它们在推断中承担的相似责任，同时保留这些责任在技术上怎样实现的不同。

主体也不能在这一框架中消失。地球科学世界的形成包含研究者如何获得新的辨识与判断能力，团队如何分担不可能由一人完成的测量和校核，以及后来的使用者如何学会信任某种封装结果。公共模型的使用可以减少每次从头检查的负担，但这种便利依赖校准、比较和纠错机制持续有效。知识被封装后，历史没有结束；它进入了一种新的维护与再生过程。

据此，本书在讨论新对象的发生时，追踪四项联系：新增的区分、支持区分的操作、可能使它失效的材料，以及它对后续研究的影响。各领域可以有不同的论述方式，但这些联系应当可被读者检查。即使暂不接受 SDE 的全部哲学主张，读者也能据此判断案例的解释是否充分。

一个容易忽略的问题，是研究对象的稳定性究竟由谁承担。初学者可以借教材认识某种地层单位，工程人员可以调用已经发布的数据，另一个学科的研究者也可以把标准年龄当作输入。这些使用并不要求每个人重新完成全部原始研究。可靠性的责任通过文献、标本、校准体系和专业分工分布开来。因而，科学世界的公共性不是所有人具有同样经验，而是不同位置的人能够在需要时找到相互检查的接口。

接口失效会产生特有的错误。一个使用者可能保留了数值，却丢掉它所对应的版本；另一个人继承了分类名称，却换掉分类所依据的尺度；还有人把对一个地区有效的校准直接搬到另一个地区。每一环单看都像在合理使用知识，连接起来却可能改变原结论的意义。思想史追踪这些连接，能够说明成熟成果如何在传播中扩大能力，又如何在传播中失去条件。

这也说明本书为何同时关心知识的生成、传播与失效。生成不是从无材料处凭空创造，传播不是把完全不变的内容从一个头脑搬到另一个头脑，失效也不总意味着原结论当初错误。一项关系可能在明确尺度上长期成立，后来被用于超出该尺度的任务而失效。我们需要判断是模型本身遭遇反证、应用条件被改变，还是中间的翻译和封装发生了问题。三种情形要求不同的修复，不应都用范式革命来命名。

这种阅读训练最终应转化为独立辨认连接的能力。拿到一幅地震剖面，可以问它的颜色表示什么量；拿到一段气候曲线，可以问时间轴怎样建立；拿到一张预警图，可以问自然状态、模型判断和行动阈值分别在何处进入。问题越具体，理论越有机会接受限制，也越可能产生新的研究方向。

全书的和谐性因此不依靠消除差异来取得。地层研究需要耐心辨认先后，数值预报需要在有限时间内处理状态，灾害行动还必须在不确定条件下作出决定。它们各有节奏，不能用同一种成功标准衡量。共同的发生学语言应当使这些差异得到更清楚的表达，并让一种领域积累的检查能力有条件地进入另一领域，而不是把所有方法改写成同一套仪式。

◆ 导论四 进入全书的路线

这部思想史没有试图收集地球科学的一切人物与概念。它选择那些能清楚呈现对象形成问题的转折，沿着从地层和时间到多圈层系统的路线展开。取舍意味着许多重要领域无法获得完整学科史待遇。读者不应把没有专章讨论的主题理解为不重要，也不应把本书的章节篇幅直接当成历史贡献的排名。

第一编建立分析工具：什么是可共同辨认的对象，为什么边界、时间和外推需要各自的证据。第二编进入地层与化石，使读者看见相对先后如何从空间关系中获得约束。第三编转向数值年龄，讨论物理时钟怎样与地质事件连接。相对年代和数值年代不是简单的落后与先进关系；数值解释仍需要层序和事件归属，二者的结合扩大了可检验问题的范围。

第四编讨论空间表达与尺度。从露头到地形、从地方调查到地图，观察如何被转成可比较的空间结构？比例尺与投影选择不会重新制造山河，却会改变哪些特征被保留、哪些关系能够被检查。这里的限制使我们理解，科学图像的连续性不能自动保证证据的连续性，也不能把图面整齐误认成世界已经没有空白。

第五、六编承担板块案例。先交代大陆移动争论与海洋观测条件，再进入海底地形、磁异常、扩张、转换断层与运动学闭合。这样的编排避免把板块学说讲成一个孤立灵感得到全世界立即承认的故事。几何、时间、运动方向和活动位置分属不同检验，只有说明它们怎样连接，又保留什么独立性，证据收敛才具有说服力。

第七编进入不可直接抵达的深部地球。波的记录、材料实验与重磁反演使某些结构成为可讨论对象，同时也暴露间接推断的多解性。第八编转入天气，观测同步、动力模型、计算能力、敏感依赖和资料同化共同改变了可预报性的含义。两编可以对照阅读：一个主要由外部记录约束内部结构，一个持续更新状态并面向未来，但二者都不能把模型输出等同于无条件的直接观察。

第九编讨论流域、入渗、海洋与水循环。系统边界并非在地图上画一道线就可以完成，地下交换、时间尺度与任务选择会改变核算责任。第十编将古气候、生命和大气放入更长的联系中。过去的气候需要代理记录，代理又需要校准与年代框架；地球系统的整体表达需要说明组成之间怎样交换，而不是把多个圈层名称放在同一张图上便宣布整体性成立。

第十一编面对遥感、数字孪生、人工智能与行动。数字产品使地球知识更容易更新、传播和调用，也使版本、训练资料和处理规则进入证据链。能够显示一幅高分辨率图像，不等于已经复制自然地球；能够标出疑似地震，不等于已经完成事件关联、人工复核和机制解释。相关章节将具体讨论这些阶段，而不以对新技术的赞美或警惕代替分析。

第十二编返回全书主张。它不把前面案例重新缩写一遍，而是比较哪些共同问题得到支持，哪些只是方法提示，哪些仍需新史料或数据复核。尤其要分开两种成功：一种是某个地学模型相对另一个地学模型的成功，另一种是 SDE 历史解释相对其他历史解释的增量。前者再有说服力，也不能自动充当后者的证据。

读者可以顺序阅读，也可以从两项承重论证开始。第 13、15 章建立矿物年龄、地层事件与标准界线之间的区别；第 27、29 章展示磁条带与转换边界怎样获得新的运动意义。读完这些章节，再返回导论和第一编，抽象方法会更容易接受具体检验。对气象和系统科学更熟悉的读者，也可以从第 38 至 40 章或第 46 至 50 章进入，然后比较同样的提问在其他对象上是否成立。

书中使用的材料分成三种角色。第一种用于重建历史上的实际主张和检验，必须尊重发表时点与原文范围。第二种用现代知识澄清早期问题，例如为了说明为什么不能由一个读数直接推出事件身份而给出的简化模型。第三种是本书提出的进一步研究方案。三者可以放在同一章，但必须让读者知道何处发生了角色转换。现代说明不能伪装成早期作者的算法，研究建议也不能写成已经完成的复算。

引文采用可追踪的来源编号。来源不是权威装饰：题录只能确认题名、作者和出版信息，摘要只能支持其确实表达的内容，全文相关段落也未必等于全部补充资料已经核验。对原始图表的重新计算与对论文结论的解释性讨论，是不同工作。正文保留影响结论成立的限定，书末来源记录集中交代核读范围与补证任务，供读者沿线检查。

我们还要警惕一种阅读习惯：以为科学史的终点就是当下教材，而当下已经拥有无须再发生的最终世界。实际上，成熟知识的可用性往往依靠不断维护的观测、模型与标准。新的材料可能只改善估计，也可能打开原先没有的区分。地球科学世界的开放性不意味着一切随时崩溃，而意味着稳定性具有条件，其延续需要工作。

本书从一块岩石出发，最终希望训练的也是一种具体能力：面对一个已经被封装为事实的结论，能够适度打开其形成过程，辨认它解决了什么问题、依赖什么联系、在哪个层次仍然可靠。不是每次使用都必须从头复算，但当对象改变、条件改变或证据发生冲突时，我们需要知道应当打开哪一层封装。这样的历史阅读，使过去的思想转折成为今天判断和创造能力的一部分。

第 1 编

地球科学世界怎样成为研究对象

第 1 章 同一地球，怎样发生出不同的科学世界

第 2 章 边界怎样从地层接触面成为共同标准

第 3 章 深时、测量时与知识时：四条时间线

第 4 章 从地方露头到全国地图：外推需要哪些许可

第 5 章 怎样写一次思想转折：证据、操作与回写

第 1 章 同一地球，怎样发生出不同的科学世界

1.1 从一块岩石进入

在一张地质图上，一条色带穿过山谷，绕开城市，又沿另一侧丘陵伸向远方。熟悉这种地图的读者会把它看作某一地层的地表分布。可是，纸面没有告诉我们这条连续色带经过了多少次不连续的观察：某处是采石场新鲜的断面，某处只是农田里翻出的碎石，还有一些地方被土壤和植被覆盖。颜色的一致，显然不等于观察条件的一致。

把色带拆开，我们至少得到四种需要分别承担责任的判断。地点判断回答材料从哪里来；层位判断回答它与上下岩层有什么关系；标本判断回答化石或矿物具有哪些可辨认特征；制图判断则回答这些局部材料为什么能够连接。前面三项即便各自正确，也未必足以保证最后一项正确。两处相似的砂岩可以形成于不同时期；同一时期的沉积也可以在海岸附近成为砂岩，在较远处成为细粒岩层。

史密斯 1815 年的地图使这种连接获得了能够公开读用的形式。伦敦地质学会保存的地图及其馆藏说明，将化石序列、底图和设色放在同一作品中呈现。[A01]沿着图上的色带追问，发现便展开为一连串比较：某处露头怎样被辨认，相邻地点怎样取得共同层位，跨过覆盖区的连接又凭什么成立？成图的历史意义，要在这些连接中寻找。

这正是本书所谓科学世界的基本尺度。它不是另造一颗物质地球，而是使某类差异能够稳定显露，使研究者能够据此辨认、判断和行动的整体系统。一块石头成为某地层的证据，既涉及岩石保存的特征，也涉及采集位置、比较操作与辨识能力。把其中任何一项单独当作全部原因，都难以解释为什么同一标本离开层位标签以后会失去一部分科学价值。

因而，地层世界不是岩石名称的仓库。它包含名称之间可检验的先后和连接，包含什么能够反驳一条连接，也包含发现矛盾以后应到哪里继续调查。在这个意义上，科学对象的成立同时发生在对象属性、互动操作和研究者能力之中。

1.2 四层对象的边界

古代沉积物落到盆地底部，是自然过程。研究者记录一处砂岩的厚度、方向和化石，是记录操作。把若干露头判为同一地层，是对象判断。让不同调查者按照同一图例表达并复核这种判断，则涉及公共标准。四者联系紧密，却不能互相替代。

地图出版不会使已经发生的沉积再发生一次，也不会改变当时水流搬运颗粒的方向。它改变的是哪些分散材料可以在今天共同进入比较，哪些问题具有可操作的形式。地图允许读者提出一个以前难以系统提出的问题：如果这条地层带确实延伸到某地，为什么当地看到的化石序列与图示不合？这个问题可以组织新的观察，而不只是引起语言上的争论。

也不能把记录理解为自然过程的透明副本。记录者必须决定在哪里测、沿什么方向描述、哪些细节值得留下。决定并非意味着任意制造事实，而是意味着一次记录有选择性。被选中的特征能够支持哪些结论，必须在后来的论证中重新说明。地质笔记里的一条线可能表示接触面，也可能表示地形轮廓；若没有符号说明，几何上同样的笔画会承担完全不同的证据任务。

对象判断又比记录多走了一步。两块化石被划为同一种，不只因为眼下看起来相像，还因为某些差异被认为来自保存、磨损或个体变化，另一些差异则被认为具有分类意义。相似究竟在哪个尺度上成立，决定它能否参与地层相关。标准的作用是把这种取舍变得可以交流和修正；标准本身并不保证每一次使用都正确。

从 SDE 关系看，显露 S 在这里是可共同辨认的地层对象，D 是有次序的露头、标本和剖面对比，E 则体现为化石身份与层位关系的相互规定。化石帮助判断地层，层位又帮助识别哪些外观差异可能具有年代意义。这种互证要成立，两条证据链必须保有可分别检查的部分：化石有形态判据，地层有接触或层序依据。若全部层位判断只来自化石，而化石的意义又完全取自假定层位，比较虽能自洽，却没有增加独立约束。

一项地层判断还会在不同任务中具有不同精度。用于说明某地区总体层序时，一个较宽的单位可能已经足够；用于判定两个相邻化石出现的先后时，同一单位又过于粗糙。因而，对象稳定不等于所有尺度固定。研究需要能够说明何时可以把差异暂时合并，何时必须重新拆开。若把全国图的分类直接搬到局部精细剖面，就可能把表达上的同一误认为过程上的同时。

这种尺度意识也使共同标准具有弹性。稳定的标准可以保存基本参照，同时允许更精细的局部描述。真正需要检查的是，新细分能否回到原来的参照中，以及合并时丢失了什么。科学世界的发生因此既包含新区别，也包含有责任地组织区别的能力。

◆ 1.3 第二种世界的进入

将视线从十九世纪地图移到二十世纪的铅同位素研究，问题会发生变化。地图首先让分散地点获得可比较的层序身份；定年研究则追问某种材料记录的过程距离现在有多远。后者不是给前者的色带简单添加一个数字，而是引入另一组材料筛选和推断条件。

帕特森、蒂尔顿和英格拉姆 1955 年发表于《科学》的《地球年龄》，其作者机构库摘要介绍了比较地表铅与铁陨石铅同位素组成的方法。[A10]这项研究入口已足以显示问题的转换：为什么来自地球之外的材料能够参与地球年龄的判断？回答必须涉及材料之间可比较的初始关系，以及放射性演化怎样改变这种关系。不能因为仪器给出了精确读数，就省略这些前提。

这里存在三个容易混淆的层次。第一是读数能否稳定复现；第二是读数是否准确代表所测样品；第三是样品读数能否支持被询问的地史事件。仪器精度提高主要直接作用于第一、第二层，不能自动跨越第三层。假如样品后来交换过元素，即使测量极其精密，所得关系仍可能记录后期过程，而不是最初形成。

因此，年龄对象的更新不能只用小数位增多来衡量。若研究把某个数字从笼统的“岩石年龄”区分为结晶、冷却或后期改造的时间，测量对象的身份就发生了实质变化。这个变化会反过来改变选样：需要检查什么矿物、什么部位、哪些保存状态，取决于研究究竟询问哪个事件。对象不是测量完成后贴上的名字，它还参与组织测量。

样品保存史同时约束事件名称的使用。不一致结果可能提示污染，也可能揭示原先合并的事件需要拆开；如何取舍，应有能够独立实施的排除规则。例如，先依据保存状态区分样品，再检查所得年龄是否相合，与看见年龄不合后才把样品宣布为污染，论证方向不同。事件身份决定什么材料值得比较，保存条件决定比较可以回答什么问题，两者在这一操作中相互限定。

史密斯的图与铅同位素研究因此可以比较，但不可合并为同一种技术。前者重在空间分布与相对次序的连接，后者重在演化关系与数值时间的连接。共同问题是：什么使一个读数或一块标本具有超出自身的证明能力？不同科学世界的发生，正体现在这些证明能力的差异之中。

◆ 1.4 全书主张的责任

“新世界”是一个有吸引力、也容易被滥用的说法。本书为它设置三项最低要求：必须指出一项新判别，一项由此能够实施的操作，以及至少一条可以追查的后续改变。新判别可以是区分同岩性异年代的地层；新操作可以是据图选择补测地点；后续改变可以是修改图例、样品档案或比较规则。三项不必在同一天发生，但不能只凭作者的解释把它们补齐。

如果新术语与旧术语逐项对应，研究者仍按同一规则选择样品、处理矛盾，那么我们只应说表达发生了变化。如果新工具提高速度，却没有改变可判别问题，也应首先承认效率上的进步。把所有改良都宣布为世界重生，会使真正的对象变化失去辨识度。

这些要求也约束 SDE 解释自身。常规科学史已经能够说明化石地层学怎样发展、仪器怎样改善以及学术共同体怎样形成规范。发生学分析不能把这些事实重新命名后就声称增加了因果解释。它需要展示以往叙述未充分说明的相互规定，例如一个已成立的对象怎样具体选择下一轮观察，而新观察怎样改变对象的可用范围。

各章所能建立的联系并不相同。有些材料足以表明对象身份改变，却没有保存制度随后调整的记录；另一些材料显示的是标准怎样继续有效，而非新对象的诞生。前者需要追查后续采用，后者需要解释维护与修订。把两种任务分开，才能比较变化的具体位置。下一章先考察岩层之间的一条界线：眼前接触、化石标志和公共标准，如何取得各自的身份？

第 2 章 边界怎样从地层接触面成为共同标准

◆ 2.1 接触面不是年表线

走到一面岩壁前，人很容易把颜色突然变化的地方看作一个时代的结束。下部较暗，上部较浅；下部细密，上部粗糙；一道清楚的面将两者隔开。这样的直觉有观察基础，却把三种不同判断压在了一起：岩性发生改变，某个标志事件出现在这里，一条年代地层界线被定义在这里。

岩性接触首先是材料和结构的差异。它可能来自沉积物来源变化、水动力条件变化、后期成岩变化，也可能涉及侵蚀和停积。其时间意义不能仅从颜色读出。同一个沉积环境可以沿海岸移动，于是在不同地点留下外观相近、形成时间却不同的接触。某一地区突然开始堆积砂层，并不意味着全球同一瞬间都进入了一个新的地质时代。

化石标志层也有自己的身份。一个物种在某处记录中的首次出现，首先意味着在现有采样和鉴定条件下，首次于该层找到它。物种何时演化出现、何时扩散到这里、这里何时提供适合保存的环境，并不是同一个问题。化石首次出现可以成为重要的相关工具，但它的含义需要由这些差别加以限定。

抽象时代界线则承担共同参照的任务。它使不同地区的记录能够讨论“这一界线以前”与“这一界线以后”。一旦指定标准点，界线便不随每个研究者的视觉印象任意移动。不过，固定参考点不等于已经知道它在所有地区的精确对应位置。定义和相关是两项工作：定义提供共同锚点，相关检验其他记录怎样与之连接。

三种位置可能接近，也可能不重合。岩性接触需要解释材料为何改变，标志层需要确认某种特征在何处出现，标准界线则为异地比较提供共同参照。即便三者在一处剖面落于相近层位，迁往另一处时也可能分开：沉积环境侧向迁移，生物出现受到局部条件限制，标准的定义却仍被保留。因此，一条线上的重合必须逐项论证，不能作为跨地区使用的默认前提。

2.2 煤山的精确落点

国际地层委员会的全球界线层型剖面和点位表，将印杜阶底界列在中国浙江煤山 D 剖面 27c 层底，以牙形石 Hindeodus parvus 的首次出现为主要标志，并记录批准年份为 2001 年。[A12]这几项信息分别回答地点、层位、识别依据和制度时间的问题。它们合在一起才形成一个能够被他人重新定位的标准入口。

“27c 层底”并不等于“一个叫 27c 的时代开始”。编号服务于剖面内部定位，生物标志服务于界线辨认，批准记录服务于共同标准的追溯。假如只记住数字，却不知道编号属于哪个剖面，同样的编号可以在其他研究中指向完全不同的岩层。科学标准之所以精确，并非它含有很多数字，而是每个数字都被固定在明确的操作关系中。

还要区分标准所对应的古老时间与批准发生的现代时间。2001 年是标准被批准的年份，显然不是岩层形成的年份。官方表把数值年龄与标准位置分列，也提醒我们这两者具有不同的修订逻辑。数值年龄依赖计时方法及其校准，标准点则依赖被批准的定义；一项计时结果更新，不会单凭这一事实自动把点位移到另一层。

官方条目在这里提供的是定义及其批准状态。[A12]定义说明应到哪里辨认共同参照，却不独自回答各候选剖面为何被取舍，也不包办所有地区的相关判断。后两类问题分别依赖评议材料和实际剖面对比。这样区分以后，煤山作为标准的制度身份可以明确陈述，而某个异地标志能否与之对应，仍须由当地记录和连接依据来判断。

从对象发生的角度，煤山点位的重要性在于它把几个原本可分离的任务结合起来。岩层中的一个位置必须能够被重新找到；标志化石必须能够被其他研究者辨认；比较规则必须允许异地材料加入。少了定位，定义无法实施；少了辨认，标志不能稳定；少了相关，标准只能停留为地方名称。

煤山案例中的 E 集中在鉴定与层位解释的相互限定。化石鉴定决定哪个层位可以作为有效的首次出现；沉积连续性又决定一次局部首次出现能否近似反映所需的时间关系。若标本来自再搬运，形态鉴定即使正确，也不足以确定其所在沉积层的年代意义；若层位清楚而诊断特征不足，位置再精确也不能补足鉴定。标本与剖面必须共同进入判断，单独保存任一项都会丢失部分定位能力。

◆ 2.3 怎样搬到别的剖面

标准点不能被物理搬到全世界，但与它建立对应的程序可以传播。设想研究者在另一处剖面发现同类牙形石，首先要问的不是“这里是否就是煤山”，而是“这个标本及其层位，在何种证据范围内支持与标准界线相关”。相关是一项有条件的推断，不是地名之间的替换。

第一种风险是生物事件的地方差异。同一种生物到达不同海域的时间可以不同；海水环境也会决定它能否生存。因而，异地记录的首次出现不必天然同步。若其他化石组合或独立地层信号与这一标志系统性错开，研究者需要判断是扩散、环境选择还是原先相关有误，不能仅因它是标准化石便取消其他证据的地位。

第二种风险来自鉴定。保存不完整、磨损、破碎和个体差异都可能影响判别。鉴定规则必须说明哪些特征具有决定意义，哪些允许变化。若规则随着需要而改变：支持预期相关的标本被宽松接纳，导致矛盾的标本被严格排除，那么相关结论将失去独立性。可重复的鉴定并非简单多数表决，而是他人能够按同样可说明的依据作出或质疑判断。

第三种风险来自记录不完整。某段沉积缺失，会使原本分开的事件在剖面上挤到一起；采样间隔过大，也会制造类似效果。这里要区别自然记录的缺失与研究记录的缺失。前者需要邻近剖面、接触关系等材料约束，后者可以通过加密采样检验。两者即使在表格中都呈现为空白，解决办法也不同。

可以用一个教学情形说明相关的权限。甲剖面中，标志化石出现在某化学变化之后；乙剖面中，两者落在同一采样层。我们不能立即认定甲乙的事件顺序矛盾。乙剖面的采样分辨率也许不足以分开它们。只有在确认取样足够细、记录连续且各标志可靠以后，真正的先后差异才会对同步假设构成更强压力。

这样的分析使标准反过来组织调查。确定需要比较的界线之后，研究者会优先寻找记录更连续、保存更适宜、可获得多项标志的剖面。这是 D=G(S,E)在本案例中的可操作含义：已有对象身份影响下一轮差异怎样取得。它不意味着标准支配自然，而是意味着标准给出寻找哪些反例最有价值的理由。

相关还有一个常被忽略的方向问题。从煤山标准出发识别乙剖面，再从乙剖面连接丙剖面，这条链不一定与煤山直接连接丙剖面具有相同可靠性。每一步都可能引入定位、保

存和标志延续范围的不确定性。如果中间环节只记下一个肯定标签，误差便可能在传播时被隐藏。相关网络应当保留连接依据，而不只保存最终时代名称。

例如，一处剖面有可靠化石，却缺少可直接定年的层；另一处有适合测年的材料，却缺少同样化石。二者可以通过第三种标志建立联系，但第三种标志本身也必须接受检验。最终年龄判断的责任分布在整条证据链中，不能全部归给最后一次仪器测量。这正是为什么一个精密数字，落到不同相关链上会具有不同的地史分辨率。

尺度也会改变所需证据。若只区分相隔很远的两个时期，较宽的相关范围可能足够；若要讨论两个突发事件谁先谁后，同样的范围就可能完全不够。标准在前一种任务中有效，不意味着它能无条件承担后一种任务。认识这种差别，可以使争论从“标准是否可靠”的总体判断转向“它对当前问题可靠到什么程度”的具体检验。

最后，多个剖面的相互一致也不能只按票数计算。若它们都使用同一套鉴定、同一个校准或同一条间接相关，部分误差可能共同存在。真正增加约束的，是来源和机制有所区别的证据仍能在合理范围内汇合。相关的强度来自这种结构，而不来自被涂成同一颜色的地点越来越多。

◆ 2.4 标准保留与修订

如果新的测年工作改变了界线的数值年龄，研究者首先应更新年龄估计及其不确定度，检查它与其他事件的时间间隔是否改变。只要标准定义没有变，地层对比中的参照点可以保留。此时知识确实发生了变化，但变化主要落在时间尺度的校准，而不是界线身份的重置。

如果标志鉴定受到根本挑战，或者标准剖面的保存、连续性和可达性出现重大问题，问题就不同了。需要讨论的不再只是给原来点位配一个更好的数字，而是原有定义及其实施条件是否仍足以承担共同参照。实际的标准修订必须遵循专业程序；本书不能从一个局部疑问直接推出已有标准应被取消。

还有一种较常见、也较易被忽略的中间状态：标准本身保留，但它在某些环境中的相关精度降低。假如标志生物在一种沉积相中稀少，研究者可能需要辅助指标。辅助指标的

引入不是任意放宽，而是要说明它与原标准之间怎样建立关系，以及误差经过几次连接后如何累积。标准的稳定和使用范围的限定可以同时成立。

把这些修订并列，界线、标本与采样之间的关系就更清楚：鉴定改变，相关层位可能重判；采样变密，首次出现的位置可能细化；年龄校准更新，事件间隔的估计可能改变。三种操作所改写的对象并不相同。常规地层学对定义和相关的区分，是这一分析的基础；SDE 在这里组织的是各类修改怎样传递，而不是另立一套界线批准规则。

因此，一项异议首先需要被定位：它质疑的是哪件标本、哪条连接、哪个数值，还是标准定义本身？只有定位到相应层次，复查才有明确对象，局部问题也不至于被夸大成对全部标准的否定。

界线由此显示出一种特殊的稳固性：它不是因为从不受到挑战而可靠，而是因为不同挑战能够被分辨，分别落到定义、相关、鉴定或校准上。把所有变化都叫作“推翻”，与把所有变化都说成“越来越精确”，同样会抹掉科学工作的实际层次。要理解这些层次怎样展开，就必须进一步拆开一项研究内部的几条时间线。

第 3 章 深时、测量时与知识时四条时间线

3.1 古老事件与新鲜读数

实验室里的仪器显示一个刚刚生成的数字，研究者却据此讨论久远的地球历史。这个日常场景包含一项不平凡的跨越：当下发生的测量，为已经不在场的过程作证。数字本身属于现在，它的历史意义来自样品保存的关系以及解释这些关系的模型。

因此，一项定年研究至少涉及自然事件、样品采集与测量、公开论证、共同标准采用四条时间线。第一条回答所指过程何时发生；第二条回答证据何时进入研究；第三条回答结论何时成为可公开检验的主张；第四条回答某种用法何时被共同制度化。四条时间线有交叉，但不能被压成“某年发现某个年龄”一句话。

以陨石材料为例，材料形成并不等于它落到地球，落到地球并不等于被收集，被收集也不等于完成化学分离和同位素测量。即便不列具体日期，这些事件之间的区别仍然重要。地表风化或后期处理可能改变样品状态，影响今天测到的信号能否保留原先过程的信息。研究时间线因此不只是书目资料，而是证据资格的一部分。

更进一步，自然时间内部也可能需要拆分。矿物形成、体系关闭以及后期改造，可以属于不同事件。哪一个被赋予年龄，要看研究模型和样品条件。若笼统写“这块岩石有多少岁”，就会把一个需要论证的问题预先关闭。精确表述应当是：在什么条件成立时，这个年龄估计可解释为哪个过程的时间。

从 SDE 角度，显露的不是孤立数字，而是附有事件身份的年龄命题。差异序列包括样品之间、测量之间和解释版本之间的比较；特征纠缠则表现为计时对象与保存史的互相限定。研究者由此能够知道下一步是提高测量精度，还是换取更适合回答原问题的材料。这两种工作都可能重要，却不能互相替代。

3.2 1955 与 1956 不能互换

帕特森的研究经常在简史中被浓缩为一个年份和一个年龄数字。这样的简写便于记忆，却容易使不同论文失去边界。1955 年的《地球年龄》署名包括帕特森、蒂尔顿和英格拉姆；1956 年的《陨石与地球的年龄》是另一篇论文。[A10][A11]题名相近、问题相连，并不意味着它们拥有相同的作者结构、材料和论证。

1955 年论文的机构库摘要说明了地表铅与铁陨石铅组成差别这一研究入口。[A10]1956 年论文则须作为独立的公开论证来定位。[A11]两者的连续性首先表现为问题相接，不能由此预定表格、样品和计算全部相同。若要判断哪项证据最早支持了哪一步结论，就须把该证据放回实际出现的文献中；只写一个著名年份，会同时抹平研究的发展和论文之间的分工。

例如，若后来论文才增加了某项比较，却被叙述者写成早期研究的起点，就会制造一种过度完整的发现过程。行动者仿佛从一开始就拥有后来全部证据，真正困难的选样、校验和论证发展被抹掉。科学史于是成为结论的倒放，而不再是问题如何逐步变得可回答的历史。

反过来，过于强调一个著名年份，也可能掩盖更早研究已经建立的操作基础。测量方法、样品处理以及不同研究者的协作，都可能跨越发表节点。论文提供公开论证的时间，不必提供最早想到问题的时间。若没有实验笔记、通信或其他明确记录，不能从文章发表日期推定思想在前一天才出现。

“已经知道什么”因而有两重范围。一重是研究者当时可接触的材料，另一重是今天用来复原该阶段的文献。摘要适合确认问题入口，逐样品误差分析则需要正文和数据；题录可确定出版身份，却不说明反例怎样被权衡。后来的回顾把这些环节连成故事时，也应保留它所依据的当年论证。材料的详略会限制历史叙述的分辨率。

综合叙事可以把 1955 与 1956 两项研究写入同一年龄问题的发展，但连接处需要留下明确接点：前文提出了什么比较，后文在哪些材料上继续推进，哪些问题在两次发表之间仍未解决。即使暂时无法重建全部实验过程，公开论证的先后仍可分别记录。这样保留空隙，读者才能分辨知识的延续与文献的独立身份，而不把两篇论文当作同一结论的两种日期。

3.3 数值年表怎样更新

煤山标准提供另一种时间分层。印杜阶底界的官方条目同时列出物理位置、标志和数值年龄，并记录 2001 年的批准信息。[A12]读者如果把它们看作一个不可分割的标签，就很难理解为什么年龄数字可以更新而标准位置仍然保留。

可以设想一种不含真实测年值的教学情形。某次研究把两个层位之间的时间间隔估为十个单位，后来改进校准后估为八个单位；标志层在剖面中的位置并未移动。在这种情况下，被修订的是时间量值及由它推得的速率，而不是原先所指的岩层。沉积过程没有因为估计改变而加快或减慢，改变的是我们目前能够为它承担的数值判断。

但也不能因此说数字更新无关对象。若研究原先把相隔的两次事件误判为一次，新测量足以将它们分开，就可能改变事件的个数和先后。反之，若差别仍在不确定范围内，不能仅凭两个中心值不同便宣布出现了两个事件。判断更新属于精度改良还是对象分化，要看它是否改变了可实施的判别。

标准采用的时间线同样复杂。一篇论文提出新年龄，并不意味着所有年表立即更新；一份年表采用新值，也不意味着所有地方剖面的相关已经重新完成。不同共同体可能以不同速度修订图表、数据库和教学材料。书中若声称某研究“改变了整个学科”，就必须有足够证据支持这种范围，而不能让论文发表自动承担全部后果。

这正是回写关系需要被具体追踪的地方。新年龄可能要求重新估计灭绝持续时间，重新比较两个盆地事件是否同步，或者重新评估某过程的速率。但这些都是不同的后续问题。只有找到实际更新记录，才可以把它们写成已经发生的历史后果；否则应当表述为该结果允许或要求开展的检验。

因此，知识时间不是自然时间的一面缩小镜子。它有自己的版本、延迟和重新组织。不同研究在共同参照中汇合，使古老事件获得越来越细的可比较身份；与此同时，旧数字仍可能残留于图表。研究者需要维护的不只是一个结论，还包括结论与出处、校准和使用范围的连接。

时间线还可以帮助理解“更早证据”与“更早认识”的区别。一份早年收藏中也许已经含有后来用于定年的材料，但当时没有相应测量和解释条件，不能据此说那个年龄结论

已经存在。材料潜在地保留信息，与某项信息取得公共知识身份，是不同的历史状态。后来的研究可以重新利用旧样品，使旧收藏发生新的科学作用。

这种重新使用并非简单从沉睡材料中取出一个始终完整的答案。保存记录是否充分、样品是否被处理过、原始位置是否可追溯，都会决定新的问题能够推进多远。研究者可能拥有样品本身，却缺少事件身份所需的记录；也可能保留旧数据，却发现校准版本已经无法追查。知识延续不仅需要材料寿命，也需要关系寿命。

反过来，一篇旧论文的结论即使不再使用，它仍可作为历史证据。今天判断数值已经修订，与研究当时为什么接受该数值，是两种不同阅读任务。前者比较自然约束和方法，后者还要重建当时能够接触的证据。把过时结论全部从历史中删除，会使我们无法解释后来修订究竟解决了什么。

◆ 3.4 作成可追查时间图

将四条时间线并列，首先可以防止一种最简单却后果严重的错误：把材料所指的年代写成材料被认识的年代。其次，它能帮助检验因果叙述。若某项标准采用早于被声称为决定性原因的研究，就必须重新检查“导致”这个词。后来研究可以加强标准的可信度，却不能倒过来成为早期采用的原因。

还可以通过空白检验叙事的稳固性。某项研究的采样年份暂时不明，并不必然使全部历史论证失效。如果现有证据足以确认采样发生在发表之前，且所讨论的因果不依赖具体月份，就可以保留较宽的时间范围。但若争论涉及谁先使用某方法，缺失的月份或年份就可能成为关键证据，不能用推测填平。

将帕特森研究与煤山标准放入四条时间线，首先要保留其任务差别：前者讨论地球及陨石材料所指的古老形成问题，后者规定年代地层比较的共同界线。采样测量日期依记录填写；公开论证分别保留 1955 和 1956 论文；标准采用把煤山 2001 年批准单列。二者可以比较知识怎样取得公开身份，却没有理由把各自的自然事件排成同一条研究进程。

同样的安排也适用于地图。地层形成、野外调查、图版制作、对外发行和后来修订不是一个时点。所谓 1815 年地图，可以指带有该年份的作品，也可以指某次具体发行的副

本；若讨论设色差别与版本演进，就必须进一步区分。书目上的同一年不能消除实物之间的历史次序。

同一个数字也可能沿着这些时间线改变含义。一次测量最初被解释为形成时间，后续研究若把它重新归于改造事件，数字即使未变，结论已经改变；反之，年龄数值修订而所指事件保留，变化主要发生在量值估计。

时间记录还要保存事件名称的变化。旧文中的同一个年龄，后来可能被解释为另一过程；若数据库只更新数字而不保留解释版本，读者就无法知道变化发生在计量还是事件归属。日期、数值和语义应共同追溯，才能使不同研究真正可比较。

四条时间线的价值终究要在解释中体现。假如分层以后没有改变任何因果判断，也没有改善证据归属，它主要是一种资料整理方法。若它揭示了某个对象先被提出、后来才获得可操作标准，或者某项读数只是在后续模型中才取得事件意义，分层就帮助我们看到了实际的知识发生过程。

深时研究尤其需要这种双重纪律：既承认古老过程不依赖现代发表才存在，也承认它的科学身份需要具体研究活动才得以成立。二者同时保留，才不会把对象发生误解为自然事件被现代人创造，也不会把自然事件的先在性误当作现成知识可以直接从岩石中取出。时间关系澄清以后，下一个问题便是：一处地方成立的判断，怎样取得通向远方的权限？

第 4 章 从地方露头到全国地图外推需要哪些许可

4.1 把地图拆回观察点

全国地质图最容易给人的错觉，是图上每一寸颜色都曾被直接观察。实际上，露头通常是间断的，野外调查也有路线限制。地图之所以能够连续，是因为它把观察和推断共同编排到了一个平面上。若不知道两者怎样连接，就可能把推断的覆盖范围误认成观察的覆盖范围。

史密斯 1815 年地图的题名明确包括英格兰、威尔士与部分苏格兰。[A01]这个范围本身说明了制图的雄心，却不能告诉我们每条边界由多少实测点支持。要评价它的证据效力，需要把连续的颜色还原为不同性质的连接：哪些地方可以沿露头追踪，哪些地方依靠化石和层序相关，哪些地方主要依据地形与地层关系作出推测。

为辨认这些连接，可以进行一次教学性重画。第一层只标出已有直接材料的位置；第二层标明连接所用的规则，例如层面追踪或化石相关；第三层保留尚无材料的区域，并将待检验的延伸与已有连接区分。叠合后，图像依然可以呈现完整区域，但读者能够沿某条色带追问，每一段究竟凭什么延续。这是一种证据分层练习，不能据此复原史密斯各次调查的实际路线。

从离散观察点到连续结构，正是地质制图需要承担的推理。把一百个点放进底图，解决的是它们位于何处；把它们连成一层，则增加了关于未见部分的判断。这项判断可以用于预期下一处露头，因此也必须允许下一处露头将它否定。连接规则越清楚，地图就越能将分散观察转成可继续检验的问题。

如果一条连接跨越了覆盖区，调查者应说明预期在覆盖区另一侧见到什么。可以预期某种岩性组合、某一相对层位或具有辨识力的化石，而不能只预期“那里也应当符合我的地图”。后者缺乏明确的失败条件。前者则允许新地点对旧图产生反向约束。

这里，地层对象一旦获得初步身份，就开始组织新的观察。图上最需要补查的地方，往往不是颜色最复杂或交通最便利的地方，而是不同连接方案会给出不同预期的地方。地图因此既是研究结果，也可以成为下一轮研究的提问装置。是否真正发生了这种回写，仍应由调查记录检验，不能从地图存在本身推定。

◆ 4.2 跨越空白的假设

连接露头有几种常见的逻辑，它们不能混用。沿走向延伸依据地层在空间上的连续性；利用化石相关依据某种组合或序列具有时间辨识力；根据地形推断则利用岩性、侵蚀与地表形态的关系。三者可能互相支持，也可能在关键位置发生冲突。

沿走向追踪看起来最接近直接观察，但它仍然有前提。断层可以错移地层，褶皱可以使同一层重复出露，侵蚀可以使露头中断。一个研究者沿着山坡找到相似岩层，不等于已经证明它们属于同一层。必须检查层序、方向和接触关系是否允许这条连接。

化石相关可以跨越不能直接追踪的空间，却面临另一组问题。若一种生物存在很长时间，相同化石只能提供较宽的年代范围；若它只适应特定环境，缺失也不必意味着年代不同。只有把组合、次序和保存条件一起纳入，化石才可能有效限制连接。它提供的是受条件约束的相关能力，而不是自动同步的标签。

地形线索则可能受到多重成因的影响。坚硬岩层常形成山脊，但相似山脊并非必然由同一地层构成；河流侵蚀和构造也会塑造类似形态。地形适合提示调查方向，却未必能够独立承担地层身份判断。把提示性的证据与决定性的证据区分，是避免外推过度的基本步骤。

设想两个地点都出现厚砂岩，而其下部化石组合明显不同。这是一个教学反例，不是史密斯某次调查的复原。若调查者坚持两地砂岩属于同一层，就需要解释组合差别为何不具有年代意义，并提出独立检验。如果唯一理由是地图已把两地涂成同色，推断便开始用自己的结果证明自己。

更稳健的办法是保留至少两种连接方案，寻找使它们分开的材料。中间露头的层序关系、另一类化石或不整合面，都可能提供判别力。资料多并不等于判别强：在两种方案都

预期出现砂岩的地点继续收集砂岩，可能增加数量而不减少不确定性。外推的科学价值在于能否暴露于有针对性的失败，而不在于色带画得多长。

◆ 4.3 全国性从何产生

一张全国地图需要共同底图、共同位置关系和可共同理解的图例。没有这些条件，地方观察即使可靠，也可能无法顺利汇合。两位调查者对同一层使用不同名称，只是其中一种困难；同一名称被用于不同层位，会造成更隐蔽的困难。全国性首先是比较关系的协调，而后才是纸面面积的扩大。

史密斯地图依托卡里的制图工作，这一合作见于馆藏说明。[A01]地质判断进入公共图面，需要把地方名称、观察位置和地层延伸放入共同的地理关系。否则，同一地点的不同称呼可能被误作两个地点，相邻地点的地层也难以在图上接续。读者由熟悉城镇进入陌生地质序列，所依赖的正是这层转换。底图因而参与了地质材料的可比较性。

共同图例也具有双重作用。一方面，它压缩复杂材料，使远隔地区可以比较；另一方面，它会把地方差异收束到有限类别。类别过细，地图可能难以读用；类别过粗，又可能遮蔽真正的层序差别。这种取舍没有脱离任务的唯一答案。用于全国比较的分类，与用于某一采石场精细研究的分类，可以有不同粒度。

因此，全国尺度不是每个地方尺度的简单相加。它增加了一种能力：在同一参照中检查分布趋势、连接矛盾和区域空白。某个地方看起来合理的层序，一旦与周围地区并列，可能暴露出不可能的重复或缺失。整体视图由此对局部判断产生新的压力。

但整体也不能吞没局部。图上的一致性有时只是制图者提前消除了冲突。如果不一致标本被排除却没有留下理由，地图看起来越整齐，反而越难复核。有效的全国性应当允许局部材料返回整体，并改变整体的连接方式。一个不能被地方反例修订的全国图，只拥有覆盖范围，没有足够的研究开放性。

在 SDE 关系中，S 可以是全国尺度的地层分布对象，D 是露头之间与图地之间的往返比较，E 则是层序规则、符号规则和空间关系的相互规定。颜色必须依赖地层身份才有地质意义；地层身份又必须借助共同表达才能跨地点持续使用。这个整体形成了新的研究能力，但它的可靠性仍逐条依赖具体连接，而不由“全国”二字担保。

跨地区连接还会遭遇边界效应。调查者在自己熟悉区域建立的分类，进入邻区时可能发现某一层逐渐变薄、分叉或变成另一岩性。如果坚持原单位必须以原外观延伸，就会把自然侧向变化误判成记录缺失。另一方面，若所有不合都被解释成侧向变化，分类又会失去可反驳性。需要事先说明哪些特征可以变化，哪些关系是维持身份的关键。

这使“同一地层”具有结构性含义。它不要求沿途每一粒材料完全相同，而要求联系建立在可说明的连续关系或相关规则上。一个地方名称能够延伸到远处，是因为该关系得到支持；若支持消失，名称的地理范围就应收缩或调整。名称跟随证据连接，不能反过来强迫所有材料保持一致。

还有一种外推来自图幅接边。相邻调查区若使用不同层级，边界处就可能出现同一单位被切成两种名称，或不同单位被合并。接边工作由此成为分类检验：研究者必须回到标本、剖面与定义，判断差异来自自然变化还是工作习惯。地图拼得上，不等于地层接得上；真正的协调要求比较规则能够对接。

这一问题让全国图拥有超出展示的研究作用。原先地方内部不显眼的分类差异，在图幅相遇时变得必须处理。整体并非只接受地方知识，它也制造了新的比较条件，使地方工作重新接受审查。这种具体回写比笼统称赞“视野更宏观”更能说明全国性能力的发生。

◆ 4.4 整体图的剩余未知

面对一张成图，读者可以实施一个简单而严格的练习：挑出最可能改变当前判断的一处空白。若两条地层带在覆盖区相遇，一种方案认为它们连续，另一种方案认为存在断层，那么最有价值的调查位置应当能够区分这两种方案，而不是随意增加一个新点。

决定调查位置时，还要区分理论价值与实施成本。一个地点可能具有很强判别力，却完全没有露头；另一个地点理论上稍弱，但能够提供清楚的接触关系。合理路线需要同时考虑能够得到什么以及这些材料能改变什么。这样的选择属于具体研究能力，不能简单归结为愿意走得更远。

外推许可因而包含一条递进链：地点材料是否可靠，分类依据是否明确，连接规则是否适用，中间可能的中断是否被考虑，以及新观察能否反驳连接。链条中任何一环薄弱，

都会影响最终色带的解释范围。但薄弱并不自动意味着整张图无效。可以保留可靠部分，把不确定性定位到具体区域或边界。

这也提供一种评估地图进步的尺度。新图比旧图覆盖更广，可能只是资料汇编扩大；新图若使原先不可检验的连接获得明确预期，并保留反例修订的通道，就增加了研究能力。两种进步都值得记录，却不宜混为同一层次的思想转折。

未测区域还须区分两种状态：一种完全缺少材料，另一种已有两套相互竞争的延伸方案。前者可能需要基础踏查，后者则可直接寻找判别点。二者在成图上都可能留白，却不能安排成同一种补测任务。

图中的问号因此不是制图者无能的标志。一个位置准确的问号，往往比未经检验的实线更能促进后续工作，因为它把研究资源指向真正可能改变连接的地方。

由此可以把地图的证据范围分成可追踪的局部连接、受相关规则支持的异地连接，以及仍待判别的延伸。第 9 章将进一步讨论这些判断怎样通过底图、设色和版本传到读者手中。眼下先要确定一个比较尺度：地图新增了覆盖区域，还是新增了可检验的连接？下一章据此讨论，怎样从一项实际工作变化判断思想转折，而不让成图的完整外观预先决定历史结论。

第 5 章 怎样写一次思想转折

证据、操作与回写

5.1 固定一个前态

从今天回望过去，成功理论往往显得顺理成章。我们知道地层能够用化石相关，知道放射性过程能够参与定年，于是容易把这些知识当作早期研究者也已拥有的共同背景。这样写成的历史有清楚的终点，却失去了真正的起点。行动者面对的困难被后来的答案预先消解了。

判断一次转折，首先要固定前态：当时可以取得哪些材料，能够实施哪些操作，什么疑问尚无可靠答案。对史密斯地图而言，前态不应被写成无人观察岩层、无人绘制地图的空白世界；更准确的问题是，已有调查和制图怎样尚未形成他所推进的那种跨地区地层比较。只有承认已有能力，才能识别新增能力。

对霍姆斯 1913 年的《地球年龄》也如此。这部公版著作讨论了冷却估龄与放射性方法，并注意到样品保存等问题。[A09]阅读它不能直接使用后来同位素质谱的表现替作者作出选择。早期研究使用的测量对象、可达精度和误差来源，必须放回当时条件中理解。今天看来明显不合适的估计，可能仍是一种能够接受检验的初步工作。

固定前态还要求区分实践中已存在的操作与后来给予它的名称。一种调查者能够熟练执行的比较，可能尚未被写成统一理论；某种宏大理论即使已被提出，也可能没有进入日常操作。若只追踪术语的首次出现，会错过实际能力的积累；若只追踪工具的使用，又可能忽略对象身份发生的变化。

前态的范围应由待解释的变化划定。若问题是两地地层何以可以比较，就先检查原有分类和连接规则；若问题是年龄为何有了新的事件含义，就先检查原有选样与解释条件。数据不足、比较规则不足和研究目标未分化，会留下不同困难。将困难具体化，才能判断后来的工作到底解除了哪一项限制，以及哪些限制依然保留。

同时，前态也不能被塑造成方便新理论获胜的弱对手。若旧方法已经能够解释某些关键材料，就应明确保留其能力。新的解释只有说明它增加了什么、在哪些条件下更有效，

才具有比较价值。思想史的公正与方法论的严格在这里是同一件事：不夸大旧世界的贫乏，才能准确说明新世界的发生。

◆ 5.2 转折必须发生了什么

一个名称出现，是可记录的事件；一个新的可检验对象成立，则需要更多证据。对象应当具有能够区分它与其他对象的判据，并能进入某些稳定操作。如果只有新的词语，却无法说明怎样辨认、怎样比较或怎样失败，那么名称的传播不宜被直接当作科学转折。

可以把一次转折拆成对象、操作和共同规范三个方面。对象方面问：原先混在一起的现象是否被分开，原先分离的材料是否获得新的共同身份？操作方面问：研究者是否因此能够选择新的测点、采样或计算？共同规范方面问：他人是否能够按可说明的规则复核，并把结果接续下去？三者有关，但并不总是同步。

以地图为教学例子，新颜色本身不是新对象；如果颜色代表一个具有独立识别依据的层位，并且能够在异地接受检验，变化才超出了装饰。以年龄为例，多写一位数字不是新计时对象；如果研究区分了样品形成与后期改造，并据此更换取样部位，那么对象与操作都发生了变化。

转折也可能不成功。新分类在少数地点表现良好，却无法用于别处；新操作能够产生更多数据，却没有减少竞争解释；共同规范广泛采用，却在特殊环境下出现系统误判。记录这些情况，才能避免把接受程度、效率和真确性混成同一指标。

本书使用 S=F(D,E)追问对象怎样从具体比较和相互规定中成立，使用 D=G(S,E)追问已成立对象怎样选择新的差异序列，使用 E=H(S,D)追问新的对象与操作怎样改变原有约束。三式不提供现成的历史答案，也不是地质过程的定量规律。它们要求作者把原本容易省略的连接写清楚。

例如，“新地图推动调查”仍然过于笼统。需要进一步说明：图中哪个矛盾改变了路线，新增观察怎样判别旧连接，又是否使图例或分类改变。如果史料只支持路线改变，就只记录到这里；不能因为三式形式上能够闭合，便把没有材料支持的最后一环补为事实。

一次转折可以包含缓慢积累、局部重组和后续稳定化。譬如，判别规则已在少数剖面中可用，并不意味着地图和公共年表随即改版；后续采用可能还要解决名称对照、尺度和

材料保存。若把这些阶段压成同一天的突破，就无法解释新能力怎样离开最初现场。统一的比较应保留阶段之间的时差，而非要求各章复现同一条成功曲线。

◆ 5.3 让对手使用同一材料

一种理论最容易获胜的方法，是只让自己解释有利材料，让对手解释残缺材料。本书要采取相反原则：常规方法史与 SDE 重构面对同一组证据，承认相同的材料边界，再比较各自能够说明什么。只有这样，所谓解释增量才具有可讨论的内容。

对史密斯地图，常规方法史可以充分说明化石地层学、测量经验、制图协作与出版条件共同促成作品。这不是一个浅薄解释。SDE 如果只是把经验叫 D、把协作叫 E、把地图叫 S，并未增加任何机制。它需要进一步说明这些因素如何改变彼此的意义，例如相同化石因层位记录不同而承担不同相关责任，或者统一图例怎样反过来暴露原来地方分类中的不一致。

对霍姆斯，常规定年史能够说明放射性发现为年龄研究提供新工具，并分析物理假设和样品误差。发生学分析若要进一步推进，就需要指出“年龄”所指事件怎样参与选择样品，以及新样品如何反过来迫使年龄名称分化。这种分析可以组织已有知识，但是否具有独立的历史因果增量，仍取决于材料是否显示实际操作发生相应变化。

比较时至少要分开四个问题。第一，谁更清楚说明对象身份；第二，谁更准确保留时间先后；第三，谁能够追踪后续操作；第四，谁能够解释没有成功的情况。某理论在其中一项有优势，不等于在全部方面都更强。跨案例统一表达可能改善教学，却未必提高数值预测；找到新的历史因果环节，也不等于发现新的地球物理定律。

负例尤其重要。假如一项研究获得更精密数据以后，旧模型即可解决异常，那么这里可以是旧世界的修复，而不是新世界发生。若 SDE 分析仍坚持把它解释为彻底重生，就暴露了分类过宽的问题。相反，若新操作改变了什么才算同一对象，即使论文标题没有使用革命性词语，也可能存在真正的思想转折。

这类比较还要求理论接受自己选择材料造成的压力。某框架若只能解释资料丰富、结果成功的案例，却无法处理证据分散和结论不确定的历史，就需要缩小适用范围。承认范围并不削弱理论的严肃性；它使理论从一套总能套用的语言，转变为能够承担风险的研究方式。

SDE 在本书中的地位因此是待检验的分析框架。它获得支持，应当因为具体材料显示了对象、操作与相互规定之间值得追踪的生成关系；它受到限制，也应具体指出哪条连接无法确认。读者不需要先赞成全部本体论主张，才能检查这些历史分析是否成立。

比较解释时，还要防止把研究者的事后理由当作完整原因。科学家在序言中强调一种方法，可能准确说明其经验，也可能为了便于读者理解而突出某个方面。历史因果需要把自述与实际材料相互校核。第 6 章将看到，吉尔伯特自己就说明写作次序与调查次序不完全相同。这并不削弱报告，而是提示我们应尊重不同文体承担的任务。

另一种常见误差是把条件当作充分原因。出版合作对地图传播重要，但具备同样出版条件的作品未必都产生相同研究能力；清晰露头有利于结构观察，但并非每位经过的人都会提出相同问题。解释需要说明条件通过什么具体操作转化为结果。否则，只列出仪器、资金和机构，就仍然没有说明特征如何相互规定。

必要条件与促进条件也应分开。某项新测量可能没有它就无法完成特定判别，这接近必要条件；某种路线自主性可能使研究更快，却未必是唯一可行路径，这属于促进条件。两者在叙事中都可重要，但用“离不开”概括所有关系，会夸大因果强度。反事实判断应当与可用材料相称。

本书还把不同时间尺度的变化拆开。一个剖面今天被重新分类，是局部工作变化；若分类进一步进入公共年表，才扩大为标准变化；若它改变许多领域的研究问题，则需要更广接受史证据。层层扩大的叙事范围，应当有层层增加的证据支撑，不能让一个准确案例承担整个时代的全部转折。

这些区别有助于避免框架自我保护。若任何不一致都被说成系统复杂，任何成功都被归给互动发生，那么理论将不再承受风险。更好的做法是明确预计观察到哪种关系，承认什么材料会使解释缩小，并记录实际未出现的环节。理论的生命力不在于永不承认失败，而在于失败能够改变下一轮问题。

◆ 5.4 保存未发生和没成功

历史写作通常保存成功结果，失败过程则容易变成陪衬。然而，如果只保留走向今天的路径，就无法判断某项方法究竟因为什么有效。未被采用的分类、没有产生预期结果的测量，以及后来被修订的地图，都是检验发生机制的重要材料。

本书将三种情况区别记录。其一，异常被已有方法解决，说明原有对象还有修复能力。其二，新概念被提出，却没有带来可辨认的操作变化，说明其发生尚未充分展开。其三，材料不足以确认先后或因果，只能保留候选解释。三者都不能被含糊写成“经过曲折终于成功”。

失败条件应当来自案例本身。地图连接的失败，是新露头与其明确预期不符；层序规则的失败，是今天上下已不能代表原初上下而又未作校正；年龄解释的失败，是样品条件不足以支持所指事件。它们各有特定形态，不能用“仍需更多研究”一句话代替。

六条路径也须在这里受到约束。它们可以帮助作者从显露、差异序列或特征纠缠的不同入口复审材料，却不自动揭示科学家的心理顺序。研究者可能往返、协作，也可能先在操作中遇到困难，再在发表时重排论证。没有时间记录，就不能把本书写作顺序说成发现顺序。

同样，本书的 P1、P2、P3 只是分别追踪对象、路线和相互规定关系变化的分析相位。某章只有 P2 而没有足够材料支持其他相位，并不构成缺陷。真实历史不负责填满我们的表格。允许缺席，才能使框架受到材料约束。

失败记录也要跟踪它的后果。一个反例可能只使某条相关缩短，也可能迫使整套分类重排；前一种修改不能被写成全盘推翻，后一种修改也不能被压缩为数据订正。比较修订涉及的范围，才有助于确定思想转折发生在哪个层次。

共同规程还必须容许各章保有自己的材料节奏。地图通过视觉关系展开，定年通过样品与演化关系展开，天气研究则可能依赖连续记录。统一问题不要求统一叙事外形。只有保留这些差异，比较才有对象，而不是作者把不同历史磨成同一种说明。

吉尔伯特的亨利山调查将把上述区分带入具体现场。他获得了调整路线的自由，但自由的作用仍需由所取得的结构证据来说明：某次转向究竟只是节省时间，还是让原本无法区分的解释获得判别？下一章将把序言中的自述与山体接触、观察位置和报告图示联系起来，检验研究条件如何转化为研究能力。

第 2 编

地层化石与可比较的过去

第 6 章 露头如何变成证据：吉尔伯特的调查自由

第 7 章 斯特诺的层序规则：从空间接触推断时间

第 8 章 化石从标本变成跨地区的时间标记

第 9 章 1815 年地质图：颜色、底图与一个可流通的地层世界

第 10 章 不整合：怎样把缺失读成地球历史

第 6 章 露头如何变成证据
吉尔伯特的调查自由

6.1 一段有方法意义的序言

吉尔伯特在《亨利山地质报告》序言中指出，以往一些西部考察里，地质学家随地理调查者行进，难以决定路线和行进速度。即使山体结构存疑，也可能没有机会到能解决疑问的地点察看。亨利山调查给予他的条件不同：调查次序与方法可以围绕地质问题安排。[A13]地理路线与地质路线因此成为两个可区分的安排。前者使调查队穿过区域，后者还要把研究者带到足以判别结构的位置。

报告所据材料来自 1875 与 1876 年的调查，题名页标示 1877 年。序言同时承认季节限制，降雪使实际工作时间受到约束。自主并未取消自然条件和时间预算，它改变的是如何使用有限的可用时间。选择某个地点，不再只是因为路线恰好经过，而是因为它可能解开结构问题。

吉尔伯特还区分了两种观察：从较高位置获得总体分布，再到最能显露特殊结构的地方进行局部检查。这一安排具有方法意义。全景让调查者形成空间假说，局部接触则检验假说是否成立。仅有全景可能把复杂结构看得过于整齐；仅有零散局部，又可能看不出它们属于怎样的整体。

不过，序言毕竟是研究完成后的自我说明。作者认为调查条件优越，并不直接证明若无这种自由就不可能得到结果。要作更强的因果判断，需要比较关键证据与路线选择之间的关系：哪些地点不是原先行程会经过的，为什么被选中，又怎样改变了模型。现有自述支持调查策略的重要性，却不足以建立严格的必要条件。

因此，本章把自由理解为有证据责任的选择能力。它不是随意走动，也不是拒绝规划，而是允许问题改变规划。当研究者知道两个解释在哪种接触关系上分歧最大，能够把路线转向那里，自由才转化为具体判别力。调查自主性与地质知识在这里互相规定：没有初步结构理解，自主难以形成有效方向；没有调整方向的权限，初步理解又可能得不到关键检验。

6.2 露头的选择不是任意

亨利山报告中一个格外有启发性的细节，是吉尔伯特承认若不知道层序，某些砂岩仅凭岩性难以区分。他也指出风化表面与新鲜断面可以给出不同的结构印象。[A13]这说明露头并非直接展示一个完整身份，材料特征必须放进位置、保存和层序关系中才具有充分意义。

一面岩壁看得很清楚，不等于它对当前问题最有用。若两种模型都预期那里存在同样砂岩，再精细描写也可能不能区分它们。相反，一处不甚壮观的接触，只要显示岩体是截穿层理还是沿层展开，就可能具有更强的判别力。观察价值与景观震撼之间没有必然比例。

报告把侵入体与围岩的空间关系作为关键材料，并通过多种剖面和立体表达帮助读者理解。我们可以据此提出一种现代教学练习：先不看最后的结构名称，只记录岩体边缘与层理之间的相对关系，再问哪种形成过程能够产生这种组合。这样可以防止名称预先替代观察。

同一个接触还需要从不同方向检查。平面上看似沿层的岩体，在另一剖面可能显示截穿；局部拱曲也可能只是更大构造的一部分。报告明确讨论岩脉、岩床与岩盖之间的区别和连续过渡，表明分类不能脱离接触关系简单执行。[A13]一个岩体的不同部分可以表现不同几何关系，名称应帮助理解这种变化，而不是把自然强行切成互不相容的盒子。

调查选择因此涉及尺度转换。从全景看，山体是隆起；从围岩层序看，它可能是若干变形单元；从接触看，又需要区分侵入和覆盖。各尺度不是彼此竞争的照片，而是共同约束结构解释。若局部材料与整体假说矛盾，应当追查是尺度衔接出了问题，还是整体假说需要改变。

在 SDE 语言中，这里的 D 是观察位置与关系差异形成的序列，而不只是走过的距离。E 则体现在结构问题规定何种位置具有意义，暴露状态又限制这种问题能够获得怎样的回答。主体能力同样在其中发生：能够辨认接触与层理的研究者，才会把一处岩壁看作某个模型的检验点；而新的接触关系又可能迫使他学习新的辨认方式。

6.3 从描述进入竞争

吉尔伯特在报告中将熔融物质抵达地表堆积成山，与物质在地下层间展开、抬起上覆岩层的情形作了区别。后一类结构用岩盖概念加以表述。[A13]这使山的形状不再只通过地表轮廓解释，而必须进入内部岩体与围岩的关系。

从外观上看，两种过程都可能造成地表升高，因此“这里是一座山”几乎没有判别力。真正的区别在于岩浆物质处在什么位置，上覆沉积层是否存在，层理怎样弯曲，以及侵蚀以后暴露的接触是否与解释相符。结构模型需要同时解释山体和围岩，不能只说明其中最显眼的一部分。

报告把示意图中某些下部通道标为未被直接观察的假设。[A13]这使剖面具有不同层次：已见接触限制岩体在露头处的位置，多个露头之间的延伸需要推断，更深的输运通道则承担解释岩浆来源的任务。物质必有来路，只能支持存在输运过程，尚不足以固定通道的具体形状。图示保留这一层次，读者才可能分开评价浅部结构与深部解释。

这一区分能够防止模型图的反向误导。读者熟悉一张完整剖面以后，往往把其中每条线都当成同等可靠。事实上，直接接触、由多个露头约束的延伸和纯粹示意部分，可能在同一图中共存。严格阅读需要追问各部分的证据身份。模型越容易理解，越需要防止清晰感替代证据强度。

竞争解释也不应被写得过弱。围岩拱起支持某种内部作用，却未必独自证明一个特定侵入史；侵蚀能够使不同初始形态趋于相似；后期变形可能叠加在早期结构上。更强的论证来自多项关系共同限制：侵入体的位置、接触特征、围岩弯曲与区域层序是否能够由同一过程一致解释。

吉尔伯特报告关于艾伦山西侧若干岩体的描述，展示了这种多项约束的力量。部分围岩跨越岩体边缘，部分被侵蚀切开；相邻岩体与同一沉积层之间呈现不同的上下关系。[A13]这些材料使“内部有岩体”进一步转化为“岩体怎样与已存在层序相互安排”的问题。研究不只寻找岩石种类，还寻找能够限定事件顺序和结构几何的关系。

这里可以看到对象成立对后续观察的反向影响。一旦岩盖成为候选结构，研究者就有理由寻找顶板、底板、边缘和关联侵入体，而不只收集山顶岩石。但是，这种定向也有危险：若只寻找符合岩盖模型的接触，忽视显示其他结构的地点，观察会变成确认练习。对象有助于选路，可靠性却要求路线保留遭遇反例的机会。

按本书的相位划分，P2 在这里追踪的是结构问题怎样改变调查选择。候选岩盖要求检查顶板、底板与边缘，实际暴露状况又决定哪些检查能够实施；找不到预期接触时，调查者需要寻找替代位置或保留未定部分。路线自主性究竟增加判别力还是仅提高效率，取决于新路线是否取得了原有安排难以提供的关键关系，不能仅从观察数量增加来判断。

亨利山材料还有一种值得保留的分类克制。吉尔伯特在讨论区域岩层时，对当时归属不清的层序使用带有不确定性的组合称呼，而没有假装已经与远处单位精确对应。他也承认若干局部地层名称可能只是暂时性的。[A13]这说明命名在调查中可以先承担定位任务，而不必立即宣称解决了全部年代关系。

暂名的价值，是让结构讨论能够继续，同时不掩盖相关尚未完成。研究者可以准确说某岩体位于这个可辨认层组之上，而暂不为层组指定一个更强的全球年代身份。若等所有分类问题完全解决才开始结构研究，工作可能无法推进；若为了推进而冒认所有身份，则会把不确定性传播到后续推理。局部可靠与整体未定可以并存。

报告中的侵蚀条件也发挥双重作用。侵蚀移走覆盖，使内部关系显露；它同时移走部分原始形态，使重建不完整。某岩体保存了顶部接触，另一岩体只剩被切开的核心，两者并不提供同样信息。把多个保存状态比较，能够使一个地点缺失的关系由另一地点约束，但这种互补必须建立在它们确属可比结构的基础上。

由此，理想剖面与现场剖面承担不同任务。理想剖面突出模型需要的关系，现场剖面保留模型实际面对的残缺。前者帮助提出预期，后者检验预期；若只展示前者，读者会高估证据完整性，若只堆积后者，又可能难以理解共同结构。报告的科学表达需要让二者互相照明，而不彼此冒充。

◆ 6.4 可复查性怎样留存

野外调查结束，现场经验并不会自动成为公共证据。观察者可能记得光线、距离和身体位置，读者却只能接触报告中的文字、图示和标本。因此，科学报告需要把个人经历中可供他人检查的部分转化为稳定记录，而不能要求读者相信“我亲眼见过”。

吉尔伯特在报告开头明确说明，写作次序并不完全按照调查次序展开。他选择先介绍结构类型，再详细说明支持材料，以便压缩描述和统一术语。[A13]这段说明对思想史尤

其重要：文章中的论证顺序不能被直接当作现场发现的时间顺序。模型先出现在书中，不证明它先于全部观察存在。

与此同时，先给模型也带来一种阅读责任。读者需要暂时悬置判断，等到证据展开后再决定是否接受；作者则必须提供足够清楚的材料，让模型不只依靠语言的连贯性成立。结构名称可以使描述简明，但不能替代描述。若某段最关键接触没有定位、没有可辨认关系，名字再准确也无法补足证据。

可复查性并不要求后来研究者在完全相同天气和光照下重复每一步，而是要求关键判断具有可追踪的依据。观察地点、相对位置、图示方向、哪些部分属于解释，都影响复查能否实施。自然露头还可能被侵蚀、覆盖或工程改变，因此保存记录本身也是延续研究能力的一部分。

报告还把整体隆起与侵蚀后的现态作图示区分。对同一山体，未被侵蚀的理想外形与实际残余承担不同任务：前者显示结构假说，后者保存观察约束。两者并列，帮助读者看见模型究竟增加了什么。

吉尔伯特所保留的这些限制，也使报告能够被后来工作接续。后人可以重新检查具体地层归属、补测被覆盖部分，或者修改力学解释，而不必把原报告整体接受或整体丢弃。证据关系越清楚，修订越能落到局部；研究能力的延续由此有了物质和文字基础。

亨利山报告将调查选择、暴露条件和结构表达连接起来，同时在图示中保留了部分假设。正因这些层次可以分辨，后来复查才有具体入口：缺接触则补查接触，层位未定则复核层序，深部通道未见则继续作为模型问题。报告未给出每次路线调整的完整因果链；它较明确留下的是一种工作组织，允许结构问题决定去哪里看，也允许未见部分在成图后继续保持未定。

下一章将沿着接触关系再向前一步。吉尔伯特借空间关系解释山体结构，斯特诺的层序思想则让空间关系承担时间判断。一块固体在另一块固体之内、一个岩层位于另一个岩层之上，为什么能够成为过去的证据？关键不在看见上下，而在说明上下是怎样形成的。

第 7 章 斯特诺的层序规则
从空间接触推断时间

7.1 一个包在另一个之内

斯特诺 1669 年著作的题名，把一个看似静止的问题摆在读者面前：固体怎样自然地包含在固体之中？[A03]我们已经习惯看到岩石里的贝壳、晶体和碎屑，很容易以为它们只是不同物质的并置。但是，一旦追问这种并置怎样形成，眼前的空间关系就开始要求一个时间答案。

一枚贝壳的外形被周围材料贴合，并不只是“甲在乙里面”这一几何事实。若周围材料在包裹时能够顺着贝壳表面成形，那么两者在那个阶段的状态就不同：至少有一方已经具有足以限定接触面的形状，另一方仍能够形成或调整其边界。先后判断由形成条件进入，而不是从“里面”两个字直接推出。

这一点需要谨慎，因为并非所有包含关系都意味着内含物比周围材料更早。后期物质可以进入先有岩石的空洞并结晶，替代作用可以使原有材料被新矿物占据。仅凭最终位置，无法把所有情形排序。本章以下有关空洞充填、侵入与变形的讨论，是为澄清现代应用条件所作的分析，不把这些完整判据全部追授给斯特诺。

真正有力的问题是：接触的哪一方保存了对另一方成形的限制？如果砾石的外缘被沉积物围绕，且有证据表明砾石来自已存在岩体的破碎，便可建立“源岩形成—破碎搬运—包裹沉积”的序列。但如果看到的是岩石裂缝中的晶体，就需要询问裂缝、流体与结晶之间的关系。相同的空间词汇，在不同生成条件下可以给出不同时间解释。

这种从形成方式到次序的推断，构成地史研究的一种基本能力。过去不必完整保存在眼前，只要某些事件留下了不能随意交换的关系，就可能恢复局部顺序。一个已被另一个过程切断或包裹的结构，使两过程之间出现了不对称性。时间在此不是由钟表读出，而是由事件关系推断。

这也解释了为什么斯特诺的问题不能被缩减为一条记忆口诀。它改变的是观察所要辨认的内容：研究者不仅看材料是什么，还看材料之间的边界怎样形成。对象从孤立石块转向具有生成关系的复合体。辨认能力、比较操作与对象属性同时进入新的研究单位，这比给岩石增加一个名字更为深刻。

1669 年著作的历史贡献可以由这些接触问题具体辨认，而不必预设现代地层学已完整包含其中。接触两侧的成形与硬化如何先后，属于原文的论证入口；倒转、侵入和保存偏差怎样改变应用，则需要分清历史命题与现代检验条件。两者分列，才能看见一项早期提问如何继续有效，也才能确定后来的地质学究竟补入了哪些判据。

进一步核读 1669 年文本的电子转写，可以看到斯特诺没有只用“自然产生”来结束问题。在提出几个主要命题之前，他讨论单独诉诸自然、太阳或某种形式为什么不足以说明具体物体的生成；材料与位置仍需要进入说明。[R07a]这段论述的历史语言与今天不同，但其问题责任很明确：一个总括名称尚未给出特定形成过程。

随后，固体之间的接触被用来讨论硬化的先后；讨论对象包括晶体、植物部分、骨骼与贝壳等，而非只有地层中的一种化石。[R07a]这一范围说明书名中的固体问题具有统摄意图。不同材料如何在同一空间中形成，成为连接矿物与生物遗存的共同问题。其强处在于把原本可分别陈列的对象放进生成关系，风险则在于不同成因是否真的允许相同推断。

这正好为现代读者提供一个检验点。若同一原则用于贝壳被沉积物包裹与晶体在空洞生长，两种接触的形成方向可能不同。我们应保留他从接触追问生成的开创性，同时逐例检查原则的范围。发生学研究能够从这里学习如何提出问题，却不能因为问题形式相近就取消物理机制的差别。

◆ 7.2 叠置的有效前提

在一套未倒转的沉积层序中，后形成的沉积物覆盖在先有沉积物上方，因而上下可以转化为先后。这种判断的关键，不是空间本身具有时间方向，而是沉积形成机制提供了从空间到时间的桥梁。离开这座桥梁，“上新下老”就只是一个可能碰巧正确的句式。

可以把这一推断拆成三个步骤。先辨认哪些界面属于层面，而不是后来的裂缝或其他构造；再确认这些层的形成具有逐层累积关系；最后判断今天所见的上下是否保留原初方

向。每一步都能失败。把裂隙当作层面，会错误增加地层数；把后期岩体当作沉积层，会误判形成过程；忽略倒转，则会把时间次序完全反过来。

设想一组砂与泥交替形成的层序，随后被褶皱成弯曲形态。只要能够追踪连续层面，弯曲本身并不消灭原来的先后。然而，在强烈倒转的一翼，今天地形上的高低可能与沉积时的上下相反。研究者必须借助能够识别原初顶底的证据，而不能在山坡上由低走高便认为自己沿时间向前。

原初沉积也并非都理想水平。斜坡、局部凹地以及层内构造都可能使沉积具有倾斜。现代应用时，需要区分一整套地层后来共同倾斜，与沉积过程内部形成的倾斜结构。仅凭“这条线是斜的”就推出造山运动，属于证据不足。结构尺度与形成机制在此互相限定。

另一种重要情形是后期侵入。一个看似夹在两层之间的岩体，不一定与上下两层按沉积次序连续形成。它可能后来沿既有层面进入。此时需要检查接触是否显示侵入作用，并寻找岩体在别处是否截穿层序。局部平行关系不足以排除后期事件，这也是第 6 章多地点接触比较的价值所在。

我们还应区分沉积物颗粒的年龄与沉积层的形成时间。一块古老岩石被侵蚀成砾石，再进入较年轻的沉积层，砾石保留的形成史比包裹它的层更长。叠置判断主要排列沉积事件，而不是宣布每一粒物质都在沉积时才诞生。这个区分将在后面讨论矿物定年时变得尤其重要：矿物的数值年龄不自动等于含它的地层年龄。

从逻辑上说，层序规则给出的往往是偏序，而不是完整时间表。若 A 早于 B，B 早于 C，就可推出 A 早于 C；但另一组地点的 D 若没有与其中任何一层建立可靠联系，就暂时无法排序。承认不能排序，不是方法失败，而是方法在报告自己的权限。把未连接的层强行排入一条线，才会制造虚假的完整历史。

甚至在已经排序的三层中，我们也不能由厚度比例直接推出持续时间比例。厚层可以较快积累，薄层可能代表较长间隔，某些时间可能根本没有留下沉积。相对先后与持续时长属于不同问题。层序规则打开了历史，却没有同时给出钟表。

这些限制使规则更有力量。它不再是一条遇到反例就需要补丁的口诀，而是一个由生成条件支撑的推断系统。研究者可以明确指出问题出在层面识别、原初方向还是后期扰

动，并选择相应观察。S 的先后显露依赖 D 中接触关系的比较，也依赖 E 中沉积方式与变形识别的相互规定；已有先后假说又会引导寻找能够推翻它的扰动证据。

原文关于地层位置的讨论，把形成时的下方支承、上部材料状态和侧向范围放在一起。它还提出，看到地层裸露侧面时，应寻找其延续或限制其分散的边界。[R07a]这使层序不只是垂直方向的口诀，也包含横向追踪的研究要求。一个被截断的露头因此能够提出“它在哪里继续”的问题，而不是仅仅结束描述。

电子转写中关于侧向范围的表述，将其他固体的围限与地层进一步延展联系起来；其是否含有“覆盖整个地球”的严格主张，仍须原扫描和可靠译本校勘。[R07a]因此，可以从这一段讨论横向追踪的要求，但不能据此认定现代侧向连续规则已完整出现。现代分析还要分别处理渐变、尖灭与侵蚀截断：同样的露头终止，可以对应沉积逐渐消失，也可以对应后来记录被移除。

关于倾斜的解释，斯特诺还讨论地下作用、空洞与上部地层破裂下陷等可能机制。[R07a]这些机制不是今天的完整构造理论，但它们表明倾斜并非仅被当作初始形状，而被纳入形成后变化。空间图像由此获得了前后状态：目前所见需要与先前状态及变化过程相联系。

◆ 7.3 从一处顺序到异地相关

一处露头解决了上下先后，不等于两处露头已经具有共同年代。甲地有三层，乙地也有三层，不能因为都排成上中下，就认定彼此一一对应。层数相同、厚度相近或颜色相似，都可能来自不同历史。局部排序与异地相关之间，还隔着一项独立的证明任务。

这可以用一个教学例子表示。甲地从下至上是砂岩、泥岩、灰岩，乙地也有相同排列。如果两地相距不远且层面能够连续追踪，连接具有较强依据；若中间隔着大面积覆盖区，就需要其他材料。重复出现的沉积环境可能在不同时间产生相同顺序，故序列外形相同也不是充分证明。

假如两地地层处在相同海拔，情况并不会因此简单。构造运动可以改变地层高度，侵蚀可以削去不同厚度；海拔是今天的空间位置，不是形成时间。把同高度当成同时性，与把地图同色当成自然同一，是同一种未经许可的跨越。

岩性连续也应分层理解。某种沉积从岸边向盆地内部渐变，横向上可以由砂质转为泥质；若只按岩性对应，可能把同一时段的不同相分开。反过来，砂质环境在时间中移动，又可以在不同地点形成外观相似但不同步的层。岩石种类、沉积环境和年代身份因此既有关联，又不能互相取代。

这时，化石的潜力开始显露。若某类生物组合在多个地点以相同次序出现，它可能提供超出岩性的相关约束。但是，这项潜力仍需论证，不能让化石承担预先决定的年代。一种组合之所以具有时间意义，需要排除它只是相似环境的反复出现，并检查保存与采集是否造成缺项。

异地相关建立以后，地方差异仍须保留。甲地的某层较厚、乙地的同一时段记录较薄，可以成为沉积条件的比较对象；若在连接时先把这种差异消除，相关虽显得整齐，却失去了继续研究地方历史的材料。共同年代提供的是比较位置，并不要求两地拥有相同厚度或相同形成过程。

局部层序与化石比较由此形成合作，而不是一方完全依赖另一方。直接接触可以为组合建立先后，组合又帮助把这种先后连接到不连续地点。最可靠的关联来自相互支持而又具有一定独立性的证据。如果化石年龄完全由假定层位决定，层位又完全由同一化石年龄决定，所谓互证就可能只是循环。

为了避免循环，可以把一组地点用于建立初步序列，再用未参与建序的地点检验。也可以使用不同性质的关系，例如明确的切割、不整合或其他独立标志。这里的教学原则并不是把现代统计程序强加给早期研究者，而是说明任何跨地点推断为何需要承担超出自洽的责任。

新对象的发生往往就出现在这个连接位置。原先甲乙两地只是各自的石层，后来能够作为同一地史问题的不同记录共同工作。它们不是因为研究者命名而成为物理上一体，而是获得了可共同辨认、比较和修订的科学身份。这个身份改变后续采集：研究者会有目的地寻找能够填补序列或打破关联的标本。

局部规则因而是通向广阔时间世界的起点，不是终点。斯特诺的问题让空间关系能够承担过去的推断；史密斯的化石工作将进一步发展不连续空间之间的比较。二者的衔接需要保留不同证明任务，不能用“地层学建立了”一句话抹掉其中的距离。

7.4 原典与现代规则分开

现代教科书为了教学，常把漫长发展压缩成几条以奠基者命名的原则。这有助于组织知识，却可能使读者误以为奠基著作已经以今天的形式包含全部规则。思想史需要把命名传统、原典论述与现代应用条件分开，既不削弱早期贡献，也不把后来工作隐去。

原典与现代规则可以按同一问题逐项比较。接触命题首先问边界形成时哪一方已经具有形状；现代应用再查后期充填、交代或变形是否改变了接触。地层位置命题问上下和侧向关系，现代应用则进一步分辨原初顶底及沉积体边界。[A03][R07a]这种比较保留问题的承接，也限定历史归属：后来的判据用于检验早期推断，不能因其今天必要，就断言原著已逐项提出。

历史文本还可能使用与今天不同的分类。一个词在当时涵盖的岩石范围，未必与现代术语相同。若直接换成现代名称，读者会得到更熟悉的句子，却可能失去作者真正面对的问题。较好的写法是保留历史概念，并说明现代分析在何处调整其范围。

原典的总体解释也可能包含后来被放弃的部分。局部推断有价值，并不要求接受作者关于整个地球历史的全部主张。相反，一项方法能够在不同总体理论中继续使用，恰好显示其证据结构具有可迁移性。我们可以继承从接触关系推断先后的能力，同时修订关于沉积普遍性、时间尺度和构造过程的判断。

这使“创造新世界”的含义更准确。新世界不是一个从此永远正确的完整体系，而是一组新的可操作区别开始稳定成立。它可以在后来被扩大、拆分或重新限定。对象的发生与对象的修订不是互相否定的两件事，而是同一研究能力延续的不同阶段。

这也解释了为什么同一原理可以在后来获得更严格的条件，而不因此完全失去历史连续性。连续的是问题与某些可复核关系，变化的是适用范围和形成机制。若把继承理解为原话永远不变，科学史只剩纪念；若把每次限定都理解为彻底替换，又无法说明方法如何积累。地层规则的生命恰在这种有条件的持续中。

图解说明进一步显示了这种历史构造。原书用连续示意讨论地层保持完整、内部空间出现、上部破裂以及后来新沉积等不同状态。[R07a]图像的目的不是记录几个独立景观，而是把目前形态放进一种变化序列。后世读者即使不接受每项动力解释，仍能辨认这种从现态追索形成过程的组织能力。

同时，示意图中的完整性不能当作每一步都被独立观察的证明。地下空洞等机制承担解释作用，它们与可见层位之间还需要证据桥梁。把历史图解理解为模型而非照片，能够同时保存它的创造性和有限性。这种阅读并不以现代结论嘲笑旧图，而是准确辨认旧图尝试解决的问题。

原文相关段落的电子转写支持以上接触、地层位置和状态变化的概括；整部拉丁文及扫描图版尚未完成校勘，故不据此建立精确引文或更广的优先权判断。尤其是从局部接触推向整个地球历史的语句，其适用范围会影响我们评价理论的普遍性。这个尚待核定的范围，应与已经能够讨论的接触推断分别保留。

常规思想史可以从自然史、解剖比较、矿物研究与地层观察的交汇理解斯特诺。这里的 SDE 分析进一步把注意力集中在同一接触怎样成为不同的研究对象：只记两种材料相邻，得到空间描述；加入边界成形条件，得到可检验的先后关系；再据此寻找切割或充填证据，原来的描述便进入后续操作。它的解释作用取决于能否分清这些步骤，而非术语是否覆盖了所有材料。

真正可以检验的增量，是我们能否更清楚地识别失败。若今天的上下被后期倒转，规则不是抽象失灵，而是原初方向与现今位置的关系改变；若内含物由后期充填形成，包含关系就不能支持原来先后；若异地关联缺乏独立约束，局部有效规则不能自动覆盖全国。这些失败各自指出需要修改哪一项互动操作或对象判据。

因此，斯特诺的思想意义不必靠一长串“第一”来证明。它可以落在一个足够坚实的转变上：在这套论证中，固体不再仅作为静止分类的材料；它们之间的接触被组织成过去过程的可讨论约束。这样的过去还没有完整数值，也尚未普遍连接到远方，却已经能够接受有条件的推理。下一章的化石，将把这种局部次序推进为跨地区的时间比较。

第 8 章 化石从标本变成跨地区的时间标记

◆ 8.1 1799 年的表格问题

一只装满化石的柜子，可以按照外形、大小或审美趣味排列，也可以按照采集地点排列。史密斯的工作提示了另一种秩序：把标本与它们所属的地层相连，并让地层按上下次序进入表格。于是，收藏不只展示自然形态的丰富，还开始保存一个能够用于异地比较的过去。

伦敦地质学会的馆藏材料把 1817 年地层化石表与 1799 年巴斯附近地层表联系起来，并保存了不同版本的线索。[A02]表格由此具有两种可比较的次序：行列中的地层次序，以及各版文本形成的历史次序。前者把标本放入地史，后者让我们追问这种排列怎样形成和修订。本章先考察第一种次序所增加的能力，再讨论版本差异能够证明到哪一步。

普通标本表的每一行可以彼此独立：第一行记一种贝壳，第二行记另一种贝壳，交换顺序通常不影响描述。地层表的行序则具有意义。某一组合位于另一组合之上，提供了一项不能任意调换的关系。如果在远处又见到类似关系，地方收藏便可能参与跨地区判断。

因此，表格不是被动容器。它要求采集者同时保存两类身份：标本是什么，以及标本从哪里、哪一层取得。缺少前者，无法比较；缺少后者，比较可能失去地史方向。一件形态完整却层位不明的标本，在分类学上仍可重要，在地层相关上却需要受到限制。科学价值不是标本单独携带的固定数量，而与其能够进入的互动关系有关。

这并不意味着标本的物质特征可由记录任意改变。相反，记录必须能够承受材料的反向检查。如果表格把两件明显不同的标本合并，却没有解释差异，或者把同一采集袋中的材料错配到不同层位，整理过程就会制造伪秩序。表格的整齐不能替代采集与分类的可靠性。

早期表中的地层名可能来自地方岩性或习惯用语，与现代单位的范围未必重合。因此，对照应保留原名，并分别记录它在表中邻接哪些层、包含哪些标本。只有这些关系足以支持对应，才可补充现代名称；仅因两者字面相近而替换，会把原来分类的宽窄及未定部分隐藏起来。表格发展的历史身份，需要由这些具体关系保存。

从 SDE 看，新的显露 S 不是“化石很多”，而是化石与层位共同构成的可比较单位。D 是组合及其出现次序的比较，E 则表现为形态身份帮助限定层位，已知层位又帮助判断保存差异和分类疑问。只有这种关系被保存在记录中，收藏才可能从个人经验转化为他人能够继续使用的证据系统。

1799 年表格的物质历史也值得注意。地质学会的馆藏说明把它记为史密斯口述、理查森书写的文件，并指出原件后来因褪色等原因变得难读，另有十九世纪制作的摹本。[R08a]我们今天面对的证据，已经包含记录、保存与复制的层次。原件并不因历史重要就自动保持清晰，摹本也不因较易阅读就自动等同全部原貌。

口述与书写的合作，说明知识外化本身可以是一项复合操作。知道哪些标本属于哪层的人，与把这种知识排成可读表格的人，共同使经验取得固定形式。表格一旦形成，又允许参与者在不重复全部野外经历的情况下检查顺序。它不是给既成知识套上一个壳，而是改变知识能够被比较和保存的方式。

不过，文件上的次序与自然地层次序之间仍要由材料连接。纸面能够排列得很整齐，不能证明自然曾以同样方式形成。史密斯的关键工作并不只是创造表格，而是让表格持续回到地层与标本。因此，收藏柜、野外露头和纸面行列构成互相校正的三个位置，任何一个都不应独占判断权。

◆ 8.2 组合优于孤立相似

两枚化石看起来相似，可能有多种解释。它们可以属于同一类生物，也可能只是保存不完整造成的近似；即便确属同一类，延续时间也可能很长。孤立相似因此常只能给出较宽范围，不能直接证明两地同时。

组合的力量在于多项限制能够交叉。设想一个纯教学例子：甲类化石在一段较长层序中持续出现，乙类只出现于其中一部分，丙类又限定了另一部分。若三者在同一可靠层位共同出现，可用范围就可能缩小。但这种缩小依赖各分布范围是如何建立的，不能凭三个名称并列便自动提高精度。

如果三类化石都同样受一种环境限制，它们提供的证据未必相互独立。看似三次支持，可能只是同一环境因素的三次表现。组合的优势不在数量本身，而在不同特征能否对

年代与环境施加不同约束。研究者需要知道哪些特征共同变化，哪些变化能够提供额外信息。

可以比较两种情形。第一种，两地岩性相同，化石组合却稳定不同。它提示相似材料未必同时形成，但还需要排除生态环境和保存的影响。第二种，两地岩性不同，化石组合及其上下次序相近。它提示不同沉积相可能属于可比较时段，但也不能忽略组合延续范围。两种情形共同破除了把岩性外观直接等同年代的习惯。

次序通常比单次出现更有约束力。若一组组合在甲地依次出现，在乙地也呈相同顺序，解释必须同时说明多层关系。一个偶然相似比较容易发生，多项有序相似更难由任意连接产生。不过，“更难”仍不等于逻辑上不可能；研究需要检查沉积环境是否以类似方式反复变化，以及分类是否受预期次序影响。

这就触及互证与循环的边界。已知层位可以帮助鉴定，但不能成为唯一鉴定理由。如果研究者因为标本来自预期层位就把它认定为某种，又因认定为某种而证明该层位，证据没有真正增加。较好的做法是保留形态判据，记录疑难标本，并用其他地点或独立关系检验初步组合。

组合还会改变采集策略。为了建立地层相关，采集者不应只挑最大、最完整、最稀有的标本，还需要能够代表组合与层位分布的材料。看起来普通的碎片，如果具有明确位置和可识别特征，可能比漂亮但来历不明的标本更能回答时间问题。对象身份改变了“值得采什么”的判断。

由此发生的科学世界，是具有地层意义的化石世界。化石仍然可以被欣赏、分类或研究生物形态，但它又增加了跨地点连接的能力。新能力不取消其他价值，也不使每件化石都自动成为计时工具。它要求特定比较、记录和相互约束共同成立。

◆ 8.3 保存与采集会制造空缺

在一张化石分布表中，空格容易被读成不存在。然而，空格至少可能意味着四件事：生物在当时当地没有出现，遗体没有保存，保存了却没有采到，或者采到了却没有被正确识别。四种情况在表面上都没有一个有效名称，背后的自然和研究过程却完全不同。

真正的生物缺席属于古生态或演化问题；未保存涉及沉积、破坏和成岩条件；未采到涉及露头、样品量和采样位置；误鉴定涉及标本状态与判据。把它们混在一起，会让研究者用错误操作解决问题。例如，增加鉴定专家不能恢复已被侵蚀掉的层；扩大采样可以检验偶然漏采，却不能保证消除环境性缺席。

首次出现尤其受这种区别影响。观察到的最低层位，是现有记录中能确认该类化石的最低位置；它不自动等于生物在地球上起源的时刻。若下部取样很少，首次出现可能被人为抬高；若标本来自后期扰动，也可能造成不属于原层位的异常记录。研究必须说明所讨论的是局部记录事件，还是更广的生物事件。

这里可以作一个不赋予真实参数的概率教学推演。假如某层确有一种稀少化石，每份独立样品都有非零但较小的发现机会，那么一次未发现并不强；连续多次未发现可以增加缺席判断的支持，但其强度依赖发现机会与样品独立性。若所有样品来自同一狭窄位置，或者保存条件使发现机会接近零，增加样品数的效用就会有限。

这一教学推演说明，负记录的力量依赖于本来有多大机会得到正记录。它没有给史密斯收藏赋予可复算的概率，而是将两种空白分开：在有利保存条件下反复未见，能够限制存在范围；在几乎无法保存的环境中未见，首先限制的是可观察性。后者若仍被用来定位出现或消失，就把记录形成条件误当成了生物分布。

化石保存还会改变外形。同一生物的不同部位可能被分开，压扁和破碎可能使诊断特征消失。层位知识有时帮助恢复判断，却也可能诱发期待偏差。研究者需要在记录中区分确定鉴定、暂定鉴定与无法判断，不能把所有疑难材料强行放入已有序列以保持表格美观。

再搬运则引入另一种时间错位。较老层中的化石可以经侵蚀进入较年轻沉积。其存在不证明新层与原层同时，也不必说明一种生物延续到了后来。需要检查标本磨损、组合混杂与沉积关系等条件。材料自身的历史与它所在层的历史可以不同，这与上一章砾石年龄和沉积年龄的区别相呼应。

这些限制不会取消生物地层比较，反而告诉研究者怎样提高其可靠性。加密采样、比较不同保存条件的剖面、记录疑难标本和寻找独立相关，都可以使解释受到更多约束。

SDE 的回写在这里具有具体形态：已建立的组合身份推动采样改变，而新的分布资料又可能迫使组合范围、鉴定规则或相关精度修订。

这些操作也有各自的有效范围。加密采样主要检验遗漏，重查诊断特征主要检验误认，比较不同沉积相则帮助辨认环境性缺席。若某一关联只在重查鉴定后改变，修订原因就应落在分类，而不必泛称地层序列全部失效。化石序列的可靠性，来自各类偏差被分别识别和处理；不同记录因此可以有明确不同的证据强度。

组合比较还需要保留采样单位的边界。一袋混合了多个薄层的标本，可以增加种类数，却可能制造从未在同一时刻共存的组合。若用这个组合进行精细相关，表面上信息更丰富，实际时间分辨率却降低。采样前需要知道层厚、接触和取样范围，采样后则不能把混合材料重新假装成单层记录。

同样，收藏中的数量不能直接当作古代丰度。容易识别、容易取出和更受收藏者喜爱的种类，可能被过度代表。若要把丰度变化解释为环境变化，必须比单纯确认存在多承担一层采样责任。存在、常见与占优势，是不同命题，所需记录不应被同一个“发现了”替代。

这些区别让表格中的一行逐渐成为复杂而可操作的单元。它不仅包含名称，也包含位置、采集范围、鉴定信心与保存状态。早期表格未必具有今天的全部字段，但它们提出的关联任务为后来这些要求提供了动力。研究发展的方向不是表格越长越好，而是每项判断能够追溯到足够具体的材料。

对于疑难标本，保存图像与采集位置可以使未来鉴定重新进入原问题。如果只留下最终名称，后来分类修订便很难追溯；如果保留材料及其关系，同一收藏就能够承担新问题。档案的价值在于保留重新判断的可能，而不只是证明过去曾经给过一个答案。这一原则也是历史收藏能够持续参与现代研究的基本条件。

当新材料改变一个化石的已知分布范围，相关关系还可能出现连锁修订。若原来把某化石当作短暂标志，后来在更早或更晚层位可靠发现，它的时间分辨力就需重新评估。新的发现没有使旧标本失去身份，却改变了旧标本在年代判断中能够承担的责任。这里改变的是旧标本作为地层证据的时间分辨力与相关权限；物质标本保持连续，它在新的比较关系中获得了重新限定的科学属性。

8.4 1817 表格怎样开放后续修订

1817 年的《有机化石地层系统》与 1816 至 1819 年间的化石图版出版物，是相关但不同的作品。馆藏说明指出，前者更多用文字列示地层化石，并收入地层化石表；后者以图版展示特征标本。[A02]这种分工让我们看到，文字、图像与表格分别承担不同的传播任务。

图像帮助读者辨认形态，表格帮助把形态放进顺序，说明文字则能够交代保存状态与使用范围。任何一种形式单独存在，都可能不足以传递完整能力。没有图像，名称对陌生读者可能过于抽象；没有表格，标本容易重新成为孤立收藏；没有层位和保存说明，漂亮图版又可能遮蔽鉴定的条件。

馆藏中存在不同版本的 1817 相关表格，这为考察回写提供了线索，但版本不同不自动证明理论关系改变。新增一行可能只是补充资料，调整颜色可能只是印制需要；只有在比较实际项目、排序和对应规则后，才能判断它是否改变了读图或采集。思想史不能用“修订版”三个字代替内容分析。

要辨认这种修订，可以先固定同一地层在两个版本中的名称与位置，再检查特征化石、说明和颜色的变化，最后寻找变化所依据的材料。这是一项版本比较方案，不能预定每处差异都有地质原因。仅有版面压缩时，比较单位可以保持；组合范围若被重定并影响异地相关，证据关系才发生相应改变。表格是否较新，与判断是否发生重构，须分别回答。

图像比较还须控制尺度。同一标本在不同版面被放大或缩小，可能给人实际尺寸不同的印象；相似大小的图像也可能来自尺寸相差很大的标本。只有保留比例及原标本身份，图版差异才有机会与鉴定变化分开。

公开表格还使个人经验承受陌生读者的检验。读者可以用自己的标本询问某一项是否可靠，也可以发现名称和层位之间的冲突。但从“具备提出质疑的条件”到“质疑实际改变作者工作”，仍需通信、修订说明或其他记录。出版创造反馈通道，不等于每条通道都已经产生有效回写。

这一区分也帮助理解科学世界的公共性。公共性不是所有人都同意，而是不同人能够围绕可定位的分歧继续工作。某标本应归哪一类、某组合在何处缺失、某条相关是否过

强，都可以成为明确的问题。共同表格使争论有了对象，而不是只有个人见识之间的竞争。

最终，化石的时间标记能力来自一组相互支持而可被拆查的关系。形态、组合、层位、保存和采样共同使它具有地史意义；已有意义又改变下一轮采集。它既不属于标本孤立的内在标签，也不由主体任意赋予。这个可操作的复合体一旦形成，就能够进入地图，让远处的地层在同一平面上获得可比较身份。

下一章将把注意力转到这一转换的视觉和生产环节。化石关联进入地图以后，颜色、底图和发行版本不只是传播外壳。它们决定哪些地层差异能够被他人看见，哪些错误能够被定位，以及一张作品怎样在流通中继续变化。

第 9 章 1815 年地质图
颜色、底图与一个可流通的地层世界

9.1 底图为什么要减法

一张用来旅行的地图，希望读者迅速找到道路、村镇和通行方向；一张用来比较地层的地图，则需要让岩层分布从这些熟悉信息中显露出来。两种任务使用同一地表，却不要求同样的信息优先级。史密斯与卡里的合作，就发生在这种转换之中。

伦敦地质学会对 1815 年地图的说明指出，卡里为它制作了专门底图，对原有地理信息作出取舍，以便更清楚地显示地质。[A01]这个事实看似属于印刷准备，实际涉及研究对象能否成立。地质线若被密集行政边界、道路和文字淹没，即使每条线都画对，读者仍可能无法稳定辨认它们之间的关系。

减法并非简单删除“无关信息”。某条道路既可能在视觉上干扰地层边界，也可能是读者定位露头的必要参照。某个城镇名字占用版面，却能把全国图上的位置与调查者的地方经验连接。因此，底图取舍是在可读性和可定位性之间寻找可用安排，不是让地质脱离地理。

可以设想三幅表达同一区域的教学图。第一幅保留全部道路和地名，地层边界仅用细线叠加；第二幅把地层填色，却不减少原有符号；第三幅保留足够定位点，减弱其他内容，再把地质关系放到视觉前景。第三幅之所以可能更有效，并非提供更多事实，而是使同一组地质事实取得更稳定的辨认条件。

这揭示了视觉研究中的一个重要区别：信息量与判别力不是同一变量。额外细节可以帮助理解，也可以使相关差异难以识别。科学图像不是自然全部特征的缩小复制，它是围绕任务组织的显露。判断其价值，应看哪些关系因此更容易被正确辨认，以及这种组织是否隐去了会改变结论的重要反例。

底图还提供尺度。地图中的一毫米代表多大地面距离，决定哪些局部变化能够被画出，哪些只能概括。一个狭窄地层露头在野外十分清楚，在全国图上却可能小于线宽。此

时需要作出取舍：适当夸示、合并或在说明中交代。比例尺不是装饰性数字，它约束地图允许承担的精细程度。

同样，印在纸上的线有宽度，而自然接触可以复杂、渐变或部分被覆盖。读者若把线的中心当作无误差位置，就会把表达规则误认成测量精度。全国图适合显示区域关系，不宜被直接用作局部工程位置的精确依据。对象的可操作范围必须与表达尺度相匹配。

卡里的底图工作因此不仅给史密斯的发现提供一张舞台。位置、符号和尺度共同参与了地层世界的公共形成。没有地层判断，底图只是一种地理工具；没有可定位的底图，地层判断又难以成为跨地区读者共同使用的对象。这里的相互规定具体发生在读图能力与图面结构之间。

对这一合作的评价，也不应把技术劳动降为天才思想的附属品。测量、雕版、设色和版面安排各自包含专业判断。新的科学对象能够广泛使用，常常正依赖这些判断被可靠地结合起来。思想史若只留下作者名字与最终图像，就会错过对象从个人经验进入共同操作的关键环节。

从底图进入地质图，还发生了观察单位的转换。旅行地图把地点看作到达与离开的节点，地质图则把地点看作连续结构的一处样本。前者关心两城之间怎样走，后者还问沿途跨越了哪些层位。同一条道路因此可以成为研究剖面的线索，它的意义随着任务改变，而不是道路的物理走向发生了变化。

地层分布线与地形等高线也不能混读。前者连接某种材料或地层关系的边界，后者连接相同高程的位置。两条线可能局部平行，但其生成规则不同。若一个地层斜穿山谷，它在地图上的弯曲会同时受层面方向与地形影响；纸上弯曲不必意味着地下层面同样弯曲。这是读图能力从形状识别走向结构解释的一步。

可以用一项教学推演来理解。让一个倾斜平面与起伏地表相交，交线会随着山脊与山谷改变形状；即使地下平面笔直，地表露头线仍可能曲折。反之，图上一条近直线也可能来自多种地形与层面组合。单看轮廓无法唯一恢复三维结构，必须结合倾向、地形和相邻层序。这说明全国图并不消除野外剖面，反而需要它们给平面图提供结构权限。

这种二维与三维的往返，使地图成为训练空间思维的工具。读者学会从一个色带想象可能的地下延伸，又必须把想象带回可测接触接受限制。图像既扩展判断，也可能诱发过度确定；成熟的读图不是把每个色块看成完整实体，而是知道它由什么观察和推断支撑。

◆ 9.2 颜色不是装饰

地图上的绿色并不说明自然界相应岩层就是绿色。史密斯的设色与岩层外观有一定联系，但它首先遵循表达需要；白垩使用绿色，正显示自然颜色与制图颜色之间不是照相式对应。[A01]颜色的地质意义来自图例和层序规则，而不是颜料自身。

这种规则化让视觉差异取得科学作用。读者不必在每个位置重复阅读长篇文字，就能把不同地区的同类单位并置。色带使横向分布一眼可见，相邻颜色又能够提示不同层位之间的排列。地图把逐项阅读转化为同时比较，这是一种具体认知能力的扩展。

但颜色不能独立于图例运作。同一色调在不同地图中可能代表不同单位，同一地图的不同副本也可能因设色、褪色和保存而呈现差别。把照片上的颜色直接当作历史图例的精确值，会混入复制过程的变化。研究原图时，需要同时检查标注、相邻单位与制作身份，而不是只凭屏幕显示判断。

史密斯使用浓淡渐变表达地层关系，较浓的一侧具有特定层序含义；馆藏材料保存了其说明文字的相关内容。[A01]这使颜色不只表示类别，也参与表达边缘与方向。一个看似纯粹美术的动作，实际上承担了地质关系的编码。

我们可以把它理解为一种约定下的操作：先辨认色相所代表的单位，再辨认浓淡在规则中指向何种边缘，最后与邻近颜色、地名和层序表核对。若读者不知道渐变规则，只把浓色看作“岩层更多”或“海拔更高”，就会得到错误解释。科学视觉能力必须学习，它不是自然地从眼睛通向图像。

手工设色让这种编码存在实际风险。不同设色者可以把相似颜料涂出不同浓度，渐变位置也可能不同。如果差异落在没有地质意义的区域，也许只是生产变化；若恰好改变读者辨认边界的方式，就会影响地图的使用。版面品质与科学判断因此能够在某些位置直接相连。

这并不意味着所有色差都属于知识变化。区分生产差异、保存差异和地质修订，是版本研究的基本任务。生产差异来自同一指令的不同执行；保存差异来自后来物理状态改变；地质修订则涉及表达内容本身。三者可以同时存在，不能见到两幅图不同就立即宣布史密斯改变了理论。

一个有效的检查顺序，是先看边界线和文字是否改变，再看色带范围与相邻关系，最后才比较色调。若只有某处整体更浅，而边界和标签一致，科学解释的变化未必成立；若一个原来未区分的单位有了独立边界和对应说明，就更可能涉及分类或相关的修订。这个判断依赖具体副本，而不是作品总名称。

颜色还带来一种容易被忽略的公共责任。地图对不同经验读者开放，意味着作者必须预想误读。熟悉野外的制图者或许知道某条渐变边缘的含义，陌生读者却可能把它当作确定边界。说明文字、图例和图像需要共同工作，使使用者既能读出关系，也能知道表达的限制。

从 SDE 看，图像显露 S 是可按共同规则辨认的地层分布，D 包括野外比较与图地核对，E 则是地层身份与视觉规则的相互限定。图例规定哪种色差具有地质意义，实际读图中的歧义又可能要求调整图例或设色。要识别这一回写，版本材料须同时显示表达发生了何种改变、它回应了什么误读；只有色调深浅的差别，尚不足以确定这条因果联系。

科学地图之美在这里具有实质基础。清晰的层次和协调的颜色可以帮助研究者辨认复杂关系，但美感不能证明地层判断正确。和谐的图面若以抹去矛盾为代价，就会损害研究；能够让差异与联系同时显露的组织，才更接近科学表达所需的美。

◆ 9.3 1815 不是一个瞬间

作品上印着 1815 年，并不意味着全部副本在同一天产生，也不意味着其内容从此固定。馆藏说明区分了图面日期、早期发行和主要发行阶段，并指出地图在持续生产中发生修订。[A01]因而，“1815 年地图”首先是一个作品身份，进一步讨论历史时还需要辨认具体副本。

这与第 3 章的四条时间线直接相连。地层形成远在调查之前；调查和制图经历一段过程；版面刻印与手工设色有各自生产节奏；发行以后，新的材料仍可能进入后续副本。一个统一题名把这些过程组织起来，却不能把它们压缩为同一时刻。

伦敦地质学会以怀特岛白垩露头的变化展示不同阶段的地图，并列出若干副本图像。[A01]本书据此可以确认版本研究有具体空间对象，而不只是书目年份。但要判断每一次边界变化怎样产生，仍需把副本、调查材料和修订记录连接。馆藏标注能够定位变化，不能独自提供其全部原因。

版本差异具有几种不同解释。新露头可能改变对地层延伸的判断；旧材料可能在新层序理解下被重新解释；生产者也可能纠正先前设色或刻线错误。前两种涉及认识关系，第三种可能只是使成品更准确地执行既定判断。它们都能使图像改变，却不属于同一种科学过程。

因此，版本比较应固定同一地区和同一尺度。若一个副本的照片比另一个放大，边界似乎更细，并不证明历史内容更丰富；若纸张伸缩或扫描变形改变位置，也不能直接当作地质修订。现代数字叠合能够帮助对照，但其配准方法本身也要留下说明。

2016 年彼得·威格利在《GSA Today》发表的研究，以数字化方法比较史密斯地图，讨论了把历史图像放入地理信息系统时的投影与配准问题。[R09a]这是一项现代研究，不是史密斯当年的技术。它提醒我们，今天复审早期地图，同样需要处理观察对象怎样经过转换才可比较。

数字叠合首先面临定位问题。历史地图与现代坐标系统并不直接重合，海岸线和城镇位置也可能存在系统偏差。未经处理便将两者叠放，所见差异会同时包含地理定位误差与地质判断差别。因此，投影与配准方法须连同比较结果一起保存，读者才能判断某条边界的偏移属于哪一层。这里采用威格利研究所示的比较思路，具体图像差异仍应回到各副本检验。[R09a]

版本比较也能改变我们理解“发现”的方式。如果一个重要单位在后续图中逐渐稳定，发现就不是一个孤立瞬间，而是一组判断经过制作、使用和修订才成为公共对象。即使作者早已在私人经验中拥有某种认识，公开地图让它承担的责任仍然不同：它现在必须接受他人在陌生地点的检验。

反过来，持续修订并不意味着早期图毫无效力。一个对象可以在一定范围内可靠，同时仍有局部未决。把成熟度理解为非黑即白，会让历史陷入两个极端：要么先驱一开始已经全部正确，要么只要发现错误就取消其贡献。更合理的评价，是逐项说明哪些关联已经稳定，哪些在后来才被调整。

我们还应避免用今天的分类直接裁判所有旧图。若早期单位把现代若干层组合在一起，它可能与当时地图尺度和任务一致；只有在它导致本应能够避免的错误连接时，才构成相应限制。历史比较既要保留自然证据的约束，也要理解不同阶段对象粒度的差异。

1815 标记作品进入公共使用的重要节点，后续副本则继续承载判断的调整。读者取得哪一幅图，决定他面对哪一组边界与图例；作者同一时期的认识，也未必立即进入每个读者手中的成品。于是，发现的公开历史需要同时追踪判断与载体，才能解释同名作品为何在不同地点支持了不完全相同的使用。

对于具体副本，我们还可以把修订分成三种空间尺度。点的修订纠正某个地点的归属；线的修订改变一条接触或延伸；面的修订重新安排某个单位的区域分布。三种改变可能相互牵动，但证据要求不同。一个新露头足以纠正一点，却未必足以让整片区域改色；从点到面的扩张仍需要第 4 章讨论的外推许可。

层序图例的改变尤其值得注意。若只是图上某条线移动，其他区域可能仍可按原规则理解；若某个单位在图例中被拆分，则许多已涂区域都需要重新考虑。后一种变化影响的不只是某地判断，而是可比较单位本身。思想史若只比较视觉差别大小，可能把一处显眼色差看得很重，却忽略图例中一项很小但作用广泛的改动。

读图还涉及反向检查。假设某区域按照图例必须经过甲、乙、丙三个单位，实际沿线却只见甲与丙，读者可以提出几类不同问题：乙是否被侵蚀，是否尖灭，是否被断层省略，或者地图相关错误？图像并不直接给出答案，却把缺项变成能够明确定位的问题。与没有统一图例的零散描述相比，这是一种新增判别。

地图与剖面结合时，空间约束还会进一步增强。平面分布允许许多地下解释，一条可靠剖面可以排除其中一部分；剖面若与图上接触的出露位置冲突，也会受到修正压力。二者不是一幅负责展示、一幅负责真实，而是用不同投影限制同一结构。它们共同形成可以往返检验的地层空间。

副本还可以改变读者的观看方式。分张地图适合逐区检查，拼接的大图容易显露整体分布；折叠使用则便于移动，却可能遮住相邻关系。媒介形式不改变岩层本身，却改变一次观看能够同时纳入哪些地点。因此，作品的物理组织与其研究用途有关。评价不同形式时，应问它们分别支持哪种比较，而不是认为越大或越便携就一定越科学。地质世界的公共实现，也包括这些具体使用条件。

这样的比较也要求记录垂直比例。有些剖面为了显示细节会强化垂向变化，如果读者把图上倾角直接当作真实倾角，就可能误判结构。图示的增强可以有教学价值，但增强规则必须可见。科学表达可以为了辨认而改变视觉比例，却应同时保存读者还原关系的条件。

这些问题都表明，地图的历史不仅是内容不断增多的历史，也是表达权限逐步明确的历史。作品越能进入陌生人的判断，越需要把尺度、符号与证据关系交代清楚。公共性扩大了知识作用，也增加了作品必须承担的说明责任。

◆ 9.4 流通怎样成为回写

地图一旦进入读者手中，新的地层世界便能够在作者不在场的情况下运作。读者可以从熟悉地点出发，沿色带预想远处岩层，也可以把当地材料拿来与图示比较。这种脱离个人现场的延续，是出版带来的重要能力。

但流通不是反馈的同义词。出售一份地图，说明有人取得了作品；使用地图，说明表达进入实践；把发现的问题传回并造成修订，才构成更完整的回写。三者在历史材料中应分别辨认。销售数量不能证明科学判断正确，读者赞誉也不能替代具体检验。

一条有力的反馈链可以这样重建：读者指出某地区图示与露头不合，给出能够定位的标本或剖面；作者或其他研究者检查差异；差异导致边界、分类或说明改变；后续副本使这种改变进入公共使用。这里每一环都有独立证据责任。若只保留来信中的疑问，却没有找到后续处理，就只能确认反馈发生，不能确认已完成修订。

出版合作也会影响这种链条。雕版、设色和销售有各自时间与成本，科学判断即使更新，也未必立即反映在所有副本中。某些读者继续使用旧图，另一些已取得新图，因而同

一作品名下可以并存不同信息状态。共同世界的公共性不是完全同步，而是能够识别版本并追踪变化。

这使档案保存成为科学能力的组成部分。若旧版本被全部丢弃，我们只能看到最终图像，无法知道哪些差异曾经构成问题。若新版本没有标明变化，又容易让读者把自己的副本当作唯一状态。保留副本、日期和修订线索，有助于后来研究者把图像变化重新连接到知识变化。

地图还可能服务于采矿、农业、建筑材料和水源等实践。史密斯附图说明本身表达过这类用途期待。[A01]这些期待说明作者把地层关系看作能够引导行动的知识，而不只是自然史陈列。不过，具体经济成效需要相应案例；不能因作品声称有用，就替所有使用者计算收益。

实践检验也有双向性。成功找到某种材料，可以支持地图在相关位置的判断，却不能自动证明全部地层理论；一次失败可能来自地图错误，也可能来自局部变化、深度估计或执行问题。把结果回写到知识，需要定位哪条推断受到影响。没有这种定位，实践容易只提供赞成或反对的总体情绪。

这里的科学共同体不是站在图外的裁判群体。不同读者带来新的观察位置、使用问题和比较能力，使图中对象的可检验范围扩大。与此同时，他们也需要学习图例、尺度和层序关系，才能提出有效质疑。主体能力与对象公共性在互动中共同发展，正是新科学世界形成的另一面。

一种有效反馈还需要明确指向哪个副本；否则，不同读者围绕不同图像争论，却误以为彼此在讨论同一个判断。

反馈中的位置说明还要能够跨副本复现。若来信所指的“某条绿色边缘”在另一副本中深浅不同，就须借助地名、相邻单位或更明确的位置描述来确认对象。否则，一项地质异议可能其实针对表达差异，修订者也难以知道应调整边界判断还是设色执行。

地图的长期价值因而不只在于某些色带后来被证明正确，还在于它把分散材料组织成能够继续提问的形式。读者可以指出哪里不对、为什么不对、下一次怎样查，而不仅是接受一份权威总图。当这种能力成立，出版便成为知识继续发生的一项条件。

流通、使用和修订因而需要三类记录。发行材料说明地图到达何处，使用记录说明哪些判断进入实践，版本与修改依据则说明问题怎样返回图面。只有三者在具体事项上连接，才有条件讨论对象、操作和表达规则的共同变化。若差异止于设色执行，就应在生产层解释；若新露头改变了相关范围，则需进入地质判断层。两种变化都重要，但它们的因果链不同。

地质图把地层在空间中连接起来，却仍有它无法填满的地方。某些接触不是两段连续记录的普通衔接，而是记录被截断后重新开始。地图可以把这些位置画出来，层序可以说明其上下，却不能只靠连续色带恢复消失的时间。下一章的不整合，将让缺失本身成为需要积极论证的历史对象。

第 10 章 不整合
怎样把缺失读成地球历史

10.1 先读接触，再讲名句

西卡角的力量首先来自接触。较陡、近乎直立的一组地层，被较缓倾的一组地层覆盖，两组层理以明显不同方向相遇。国际地科联的地质遗产说明，以现代地层学将下部归入志留纪、上部归入泥盆纪。[A05]这些现代名称有助于今天定位，但哈顿 1788 年的观察不能被写成他已经拥有后来完整年表。

哈顿 1795 年《地球论》记述了沿苏格兰海岸寻找两类岩层接触的过程。西卡角不是一个没有问题准备的人偶然看见奇景以后立即获得全部理论的场面。他已在寻找能够说明先后形成与后期作用的接触，沿溪流和海岸检查，最后在海水洗露的岩石上获得了格外清楚的材料。[A04]

他的叙述包含几项不同证据：上部砂岩残留在下部陡立岩层端部之上；一些下部岩层突出于后来覆盖物之间；上部材料含有下部岩石的碎片。这些关系彼此配合，使解释不只是“两个方向不同”，而是涉及旧岩层已经形成、受到破坏，再被新材料覆盖的历史。

观察与解释仍须拆开。可以直接描述层理方向、截切关系、碎片与基质；“先沉积、再变形、侵蚀、再沉积”则是一项对这些关系的过程解释。后者有很强依据，但它的强度来自前者和形成机制共同限制，不来自故事听起来合理。

若只看到不同角度，仍需要考虑其他接触类型。断层也可以把不同方向的岩层并置，后期扰动可以改变局部形态。确认上覆沉积与下伏截切面之间的关系，以及碎屑来自已存在材料的证据，有助于排除简单拼接。一个出色案例的价值，恰在于多项关系共同支持了事件排序。

深时的震撼由此有了推理基础。读者先辨认多次形成与改造所留下的不同关系，再理解这种层层嵌套的历史怎样超出眼前海岸的短暂变化。接触没有直接给出年数，但已排除

一次沉积即可形成全部结构的简单解释。漫长历史由此取得可继续追查的形态：还需要哪些过程，哪些阶段能够并行，哪些必定先于另一些发生？

◆ 10.2 把过程逐步拆开

不整合的推理可以从最少事件链开始。下部岩层必须先有可辨认的层状结构，后来它们才可能整体倾斜或褶皱；倾斜层的端部被截断，要求某种移除过程；新的沉积覆盖在这个截断面上，又要求移除发生在覆盖之前。最后，今天的海岸侵蚀使这一组合重新暴露。每一步都依赖前一步留下的条件。

关键不在于列出多少事件，而在于哪些事件不能调换。若把上部覆盖放在下部倾斜之前，就需要解释为何上部没有与下部经历相同程度的变形。某些复杂情形可能存在后期叠加，但最少解释必须回应这种差别。若把截切放在覆盖之后，则难以说明上部沉积怎样直接贴合被截断的下部端面。

碎屑进一步约束次序。上部材料中的下部岩石碎片，意味着来源材料在碎片形成前已具有相应性质，随后经历破碎和搬运，再进入新的沉积。其磨圆或棱角可以帮助讨论搬运与磨损，但不能只凭外形精确计算行进距离或持续时间。哈顿注意到不同地点碎片状态的差别，并把它们联系到类似作用在不同程度上的表现。[A04]

他在 1795 年的叙述还比较了海岸少量碎片与内陆较大规模的砾质堆积。这使证据不只停留在一个著名接触面：同类材料关系可以在不同位置和规模上出现。比较扩大了解释的适用范围，同时也要求判断差异究竟只是程度变化，还是指向不同成因。

过程链可以包含同时进行的阶段。沉积、压实与某些成岩作用未必在时间上完全分离；侵蚀也可能与邻区沉积并行。教学上把它们写成先后步骤，是为了显示不可交换的约束，不意味着地球历史像工厂流水线一样一次只执行一个动作。真实地史需要在确定顺序之外保留重叠与持续。

这种限制还告诉我们，不整合证明的“缺失时间”不是一块可以直接看见的物体。眼前可见的是不同记录之间的关系；时间间隔是为解释这些关系而必须承认的过程空间。其存在与其长度是两种不同结论。几何能够强力支持前者，却通常不能独自给出后者。

可以用一个简单的推演显示这种区别。两套剖面都具有下部倾斜、顶部截切和上部覆盖的关系，但一处侵蚀去掉的材料少，另一处很多。即使知道移除厚度，也需要侵蚀速率及其变化才能求持续时间。厚度相同不等于间隔相同，间隔相同也未必留下相同厚度。时间不能直接从接触角度或岩层缺口的几何大小读出。

局部事件链还可能只是更长历史的一部分。下部岩层在被最终覆盖之前，可能经历不止一次变形或暴露；上部地层形成以后，也可能共同受到较弱变形。最少链条是一种证据要求的下界，不是宣称所有未被保存的过程都不存在。研究者应把必须发生的事件与可能但未被约束的事件分开。

从 SDE 关系看，不整合的显露 S 由接触差异 D 与过程限制 E 共同成立。接触几何规定哪些过程次序可行，形成机制又决定几何差异具有什么历史意义。一旦提出间断，新的调查会寻找覆盖层、碎屑来源和邻区连续记录。对象因此选择下一轮操作，而新剖面可能使原来“单次间断”分解为多次过程。

◆ 10.3 哈顿的循环与边界

哈顿关心的不只是某处岩层发生了什么。他希望说明地表怎样在侵蚀破坏中仍能不断更新：已有陆地提供碎屑，碎屑积累并固结，后来又进入新的陆地。在这个整体构想中，不整合成为地球具有多阶段历史的重要证据。[A04]但局部事件链与无限延续的整体循环，并不拥有相同证明强度。

一处接触可以显示至少两期沉积及其间的改造，不能单独证明过去存在无限多次循环。许多地点出现相似关系，可以加强过程反复发生的判断，也仍不能从有限观察逻辑地推出无限历史。思想史应当理解哈顿总体构想的解释力量，同时保留从地方材料到全球主张之间的推理距离。

哈顿原文提供了比名句更有价值的方法细节。他先说明自己需要寻找什么接触，再向熟悉当地的人询问，并根据溪流与海岸的暴露条件安排行程。某些地方只有地形上的大致边界，具体接触却被覆盖；沿海与河谷的裸露才使检验得以实施。[A04]问题、地方知识和暴露条件在这里共同组织研究。

这项材料直接支持对象对路线的回写。研究者不是均匀收集每块石头，而是寻找能够区分先后形成关系的接触。与此同时，实际路线也不是预设理论的机械执行。哈顿记述了没有看见预期接触的地方，也记述了虽未见接触、却发现更大规模砾质材料的情形。调查中的不满足与意外共同改变了材料结构。

报告还保留了对不充分推理的批评。在讨论别人关于矿脉终止的论证时，哈顿指出：一条矿脉在接触处停止，不必证明它早于另一套岩石；矿脉本来就可能未贯穿所有上覆材料。若后者包含矿脉碎片，才会提供另一种更直接的先后约束。[A04]这一段特别重要，因为它显示哈顿并非只用相合事实堆积理论，也区分不同证据的必然性。

这种区分可以用现代语言重新表述：某观察与一个解释相容，不等于它排除了竞争解释。矿脉止于接触面可能由多种情况产生；被包裹的矿脉碎片则对形成先后施加更强限制。虽然现代科学推理可以给出更复杂分析，这个原文细节已经足以说明，地史解释的力量来自排除能力，而不只是叙事连贯。

哈顿对某些作用机制的认识不等同于现代地质学。我们不应把今天的构造板块、沉积环境和成岩知识写成他已掌握的解释资源。早期地球循环论可以开启深时视野，同时包含后来需要修正的机制。局部推理的持续有效，不意味着总体理论中的每一部分都被保留。

沿着这条调查链，深时问题逐步取得可操作形式。最初询问的是岩层是否经历不同阶段，路线选择将它转成在哪里可能看见接触，现场比较再将它细化为哪些先后能够成立。每一步都缩小问题，却也增加具体约束。所谓对象对操作的影响，在这里可以落到寻找目标的改变；操作对对象的反向影响，则表现为远景预期被近景材料改写。

新地球科学世界由此包含一种此前不易稳定操作的身份：地层之间的缺失，能够成为有过程约束的历史关系。它并不是把空无变成实体，也不是宣称看不见就有无限时间，而是让可见接触支持对不可见过程的有限推断。这样的世界既扩大了能够询问的过去，也规定了不能凭现有证据回答的问题。

哈顿海岸叙述中还有一个适合检验“理论支配观察”的细节。他记述远看某处海岬时，外观使人以为它延续了沿岸常见岩石；靠近以后，材料却显示不同性质。[A04]远景预期与近景检查发生分歧。理论和先前经验确实组织观察，却没有使观察结果完全由理论决定。抵近、辨认和比较能够迫使预期改变。

这使他的路线具有层次：先根据区域分布预期接触，再寻找裸露位置，最后检查具体岩石关系。每一层都缩小下一步选择，却也给下一步留下修正前层的机会。如果把这段过程只写成“哈顿带着理论寻找证据”，会掩盖其中的反向约束；若写成“事实自己说话”，又会抹掉为什么沿这条海岸寻找的理由。

地方信息同样不是无关背景。他人的道路工作经验和对溪流接触的了解，使研究能够更有效地进入目标位置。科学观察并非一个人面对纯自然的孤独场景，而是借助已有地理活动产生的可达性与知识。这里的合作应按实际贡献描述，不必把每位提供地点信息的人都写成完整理论的共同提出者。

天气在原文中也具有具体作用。较好天气使海岸调查顺利，后来的降雨则限制一些观察。这样的细节不是为了让故事生动，而是说明可取得证据受实际操作条件约束。同一个接触可以在某次行程中不可见，在另一次行程中成为决定性材料。自然过程没有因此改变，科学证据的可用性却改变了。

这些材料连接了预期、路线和材料之间的往返：接触问题使某些海岸位置值得抵近，实际暴露决定可以检查哪些关系，近景差异又使区域判断得到修正。循环中的关键并非每一步都有新发现，而是先前判断能够被下一步材料改变。没有见到接触时，结论保持未定；取得碎屑关系时，论证才获得新的先后约束。

◆ 10.4 缺失也可能被误读

不是所有记录空白都属于同一种不整合。没有沉积、已有沉积被侵蚀、观察者没有采到，以及资料整理时遗漏，都可能让表格出现缺层。它们应当分别处理，因为自然间断与研究间断属于不同时间线。

未沉积意味着那段时间没有在该处形成我们寻找的记录；侵蚀意味着曾经存在的材料被移走。两者可以先后发生，也可以在相邻地区同时发生。仅凭某一地层缺席，通常不足以确定哪一种主导。需要借助接触形态、侵蚀残余、再搬运材料和邻区记录，逐步缩小解释范围。

采样遗漏则可能无需诉诸古老自然过程。若一段露头被土壤遮盖，调查者在其上下各取一个样品，表格中的跳跃首先反映观察分辨率。把它直接写成灭绝、侵蚀或环境突变，

会使研究设计的空白冒充自然史事件。加密采样能解决某些遗漏，却不能恢复已经被自然删除的材料，两类行动不可混淆。

角度明显的不整合特别容易识别，也容易让人误以为没有角度差就没有间断。第 6 章涉及的吉尔伯特报告已经明确指出，相邻层理平行不足以证明连续沉积，并讨论了侵蚀面、砾石和其他局部证据。[A13]这使我们看到，缺失的辨认不依赖一种唯一外观。相同几何方向之下，仍可能隐藏不同时间关系。

反过来，角度差也不允许忽视接触性质。若两套岩层通过断层接触，其并置不必对应我们预设的上覆沉积史。需要先确认哪种空间关系成立，再决定它允许哪种时间解释。一个强烈视觉差异可以引发问题，却不能替代地质关系的辨认。

缺失还具有空间上的不均匀性。一处盆地边缘可能长期遭受侵蚀，邻近较低地区却持续积累材料；后来它们被放进同一地层框架，某地的空白可以由另一地的记录部分约束。但这种补足不意味着把另一处岩层原样搬来。需要先证明两地时间关系可比，再讨论哪些事件能够外推。

因此，不整合既是局部对象，也是区域比较的问题。一个接触在甲地表现为明显角度变化，在乙地可能只是小幅侵蚀，在丙地甚至没有相同间断。若把所有表现都强行压成一条全球同时线，就会失去过程的空间差异。区域对比的目标应是解释这些差异为何能够关联，而不是让它们表面一致。

间断的上限与下限有时还能由邻近记录逐步约束，即使暂时没有一个唯一时长。用范围表达并不是不完整回答：它说明哪些历史仍可能成立，哪些已经被排除。地史重建可以先缩小可行序列，再寻求更精密的数值时间。

同一侵蚀面也可能经历多次改造，最后保存的是复合结果。研究者需要辨认哪些痕迹来自最早形成，哪些来自后续再暴露。它提醒我们，科学对象的稳定名称不保证自然过程单一。不整合可以是有明确辨认条件的对象，同时保留内部历史尚未完全分解的状态。

可以把本章结论压缩为三层判断。我们观察到层理方向、截切面和碎屑关系；据此可以在适当条件下推断形成、变形、移除和覆盖的相对顺序；尚不能仅凭这些关系求得全部事件的数值年龄、每一阶段持续时间或全球无限循环。这三个层次若保持分明，缺失就能成为有边界的正面证据。

因此，描述一处不整合至少要保留两个结果：由哪些关系确认存在间隔，以及目前怎样限制间隔的范围。前者较确定、后者仍宽泛，是可以同时成立的研究状态；给接触一个稳定名称，并不要求立即为它指定固定时长。

这种边界也具有研究生产力。知道几何不足以给出年数，就会寻找能够约束上下层年龄的材料；知道一个地点无法区分停积与侵蚀，就会转向邻近更连续的记录。未知不只是等待被填满的洞，它能够组织下一步问题。前提是未知被准确定位，而不是被漂亮的完整故事遮蔽。

不整合最终改变了地史阅读的节奏。地层不再是一部每页连续保存的书，岩层厚度也不再对应均匀流逝的时间。某些最关键的历史发生在两段可见记录之间，需要通过关系推断。地质学因此获得了阅读“删去之页”的能力，同时也承担了说明删去、未写和未读之间差别的责任。

到这里，第二编已经建立几项相互衔接的能力：露头成为可检验的结构证据，局部接触转化为相对先后，化石与地图把分散记录连接起来，不整合则让连接中的缺失具有过程意义。但这些能力还没有提供一个统一的数值时钟。第 11 章将面对另一类压力：当地质历史要求漫长时间，而物理模型给出不同限制时，究竟是哪一种“地球年龄”在发生冲突?

第 3 编

深时间与地球年龄的科学生成

第 11 章 地球年龄之争：从历史年表到开尔文的冷却时钟

第 12 章 渐变与灾变：把两派口号拆成可检验的问题

第 13 章 从霍姆斯到煤山 矿物时钟怎样取得事件身份

第 14 章 帕特森的地球年龄：为什么要离开地球岩石

第 15 章 金钉子与年代标尺 地方剖面怎样承担全球责任

第 11 章 地球年龄之争
从历史年表到开尔文的冷却时钟

11.1 先把各家起点说清

若把关于地球年龄的旧答案排成一条从小到大的数轴，很容易写出一部胜利史：起初人们受权威束缚，只肯承认短暂历史，后来地质学突破限制，物理学又给出不够大的数，最终放射性定年宣布正确答案。这种叙述的问题，不只是遗漏了某些人物。它预先假定，所有人都在给同一个事件计时。事实上，记载最初人类世代的年表、估计沉积岩形成所需的时段、计算熔融球体开始冷却以来的时间，以及追查行星物质聚积的时代，具有不同的起点和证据责任。

一份年表首先需要说明文献之间怎样接续，哪些年代包含在另一年代之内，有无世系遗漏，事件如何转成公历位置。它能够严密或粗疏，却并不因为算术正确便获得对地球固化的解释力。沉积厚度除以年沉积速率，看似更直接地接触自然，仍须说明该厚度是单一地点的净积累，还是把不同地点、可能同时形成的地层相加。岩层缺失的时段不在尺子上，多处地层并存的时间也不能重复相加。没有统一待估量，数字越精确，比较可能越混乱。

冷却时钟则选择一种物理状态作为起点。它问：假定地球从某个温度分布开始演化，经过多长时间，才能出现今日近地表的温度梯度？这里的时钟并非真正埋在岩石中的机械器件，而是初始状态、传热定律、边界条件与观测值组成的反演程序。它所得的时间是否等于地球形成以来的时间，取决于那一起始状态能否代表地球形成，而不是取决于结果以多少位数字印出。

霍姆斯在 1913 年《地球的年龄》中回顾这些争论，把地质时间扩展视为科学史的重要成就。[A09] 他的叙事本身也是历史材料。我们可以使用他保存的论争线索，却不能继承其中把文明排列为先进与落后的全部判断。史学分析需要分开两件事：某种计时方法是否有充分证据，以及一种社会怎样理解自身历史。前者可以被测量和模型反驳；后者还涉及文献传统、制度与价值，不因一个年龄数字的更改而得到简单裁决。

因此，所谓年龄冲突至少有三种。第一种是目标不同：测沉积持续时间和测行星形成时间，数值不等未必构成矛盾。第二种是目标相同而方法不同：两种模型对同一起点给出不相容范围，才需要寻找哪个前提失效。第三种是目标在争论中改变：原来试图估计全部地史，后来发现只能估计某一热事件以来的时间。第三种情形并非把失败藏进术语，而是对象身份受到材料约束后的重新划分。

这一划分也使思想史避免把“更老”自动当作“更好”。一个任意巨大的数字，若没有可追查的测量链，并不胜过一个有条件、能指出误差来源的较小数字。地球科学需要足够漫长的时间解释地层、生物与构造变化，但研究需要不能直接决定自然必须提供多少年。反过来，一个模型不能因为使用了成熟物理定律，就把所有超出它的地质记录判为错误。双方必须在同一问题上相遇：共同讨论的是哪个起点，哪些过程被允许发生，现存记录怎样限制它们。

◆ 11.2 1862 年的定量约束

开尔文当时以威廉·汤姆孙署名。1862 年的学会记录《论地球的长期冷却》从地下温度随深度增加和热量向外传导出发，将地球热状态与有限的冷却历史联系起来。[A08] 这是一项可批评的定量约束，不是凭声望拒绝地质学。若内部温度高于表面，且热量持续流出，就需要说明热量从哪里来、怎样输送、储存了多少、能维持多久。地质学可以质疑某个答案，却不能取消这些物理问题。

为看清逻辑，可以构造一个现代教学模型。设均匀介质占据深度 $z \geqslant 0$ 的半无限区域，最初温度处处为 T_i；从 $t=0$ 起，表面保持较低温度 T_s。介质内部无热源，热只靠传导，热扩散率 κ 为常数，温度满足 $\partial T/\partial t=\kappa\partial^2 T/\partial z^2$。令 $\Delta T=T_i-T_s$，则表面温度梯度 $g=\Delta T/\sqrt{(\pi\kappa t)}$，反解得 $t=\Delta T^2/(\pi\kappa g^2)$。$\kappa$ 的量纲为长度平方除以时间，g 为温度除以长度，因而右端确有时间量纲。这是用于辨认推断条件的标准导热解；开尔文历史计算中的地球参数和边界处理，不在此逐项复原。

这个小公式有四个启发。第一，测得梯度不等于测得年龄。梯度只有与温差、物性及传热方式结合，才能承担计时功能。第二，梯度的变化并不是直接记录地球每一年的温度，而是模型根据现在状态推断过去历时。第三，某些参数以平方进入结果，初温不确定

性可能被放大。第四，公式里没有出现的过程，例如内部产热或大尺度物质输运，也不是已经被证明不存在，而是在模型构建时被排除了。

物理模型的力量恰在于它把这些关系摆到公开位置。原来“地球古老”可以是一种由许多观察共同支持的宽泛判断，现在则必须面对一个能够逐步检查的算式。测温是否准确？近地表梯度能否代表整个地球的热史？岩石在不同温压条件下的物性是否相同？地球内部是否可以用静止、均匀的导热介质近似？这些问题并非模型外的挑剔，而是年龄结论本身的组成部分。

近地表测量还带有空间尺度的限制。一个矿井提供局部温度剖面，全球热史推断却要求判断这处剖面是否具有代表性。地下水运动可以搬运热量，岩性差异改变导热性质，地形起伏也会影响浅部温度场。同一矿井中增加测点，主要改善该剖面的分辨率；跨不同地质条件比较，才可能检验局部梯度能否迁移到整体模型。局部精度与全球代表性因此需要分别论证。

热梯度与导热假设的特征纠缠，就发生在局部读数取得计时意义的过程中。梯度作为“冷却年龄指标”的意义由传热模型规定；模型则必须给出与实测梯度相容的温度分布。改变内部传热机制，同一梯度便可能对应另一段历时；新测量若显示不同岩性或深度的梯度分布无法由同一套参数解释，又会迫使研究者重新划定模型的有效范围。E 在这里落实为可测梯度、材料传热性质与历史状态的相互限定，联系断开后，温度剖面仍然存在，却不再能独自承担原有的计时任务。

冷却推算的历史价值，正在于这些假设能够被逐项提出异议。模型给出的年龄范围使争论有了明确靶点：测量有误、参数选取不当和传热机制遗漏，会导致不同的后续工作。若争论只停在“足够古老”，便难以判断什么新观察足以改变结论。即使最终年龄未被接受，模型仍留下了可以继续研究的热储量、边界通量与内部输运问题。

这个教学模型还可以作一次不带实际地球参数的灵敏度检查。若温差估计增加一成，而梯度与热扩散率不变，推算时间不是增加一成，而约增加两成；若梯度增加一成，推算时间则约缩短到原来的八成多。改变热扩散率又以不同方式影响结果。这样的比例运算显示，一份年龄误差预算不能只看温度计读数，也需要看无法直接观察的初始状态。精确的观测与不确定的初始条件可以同时存在。

还有一个容易混淆的量：导热率与热扩散率。前者联系温度梯度和热通量，后者还包含材料储存热量的能力。两种岩石即使在同样梯度下传递相近热流，它们温度变化的快慢也未必相同。在把实验室物性迁移到整个地球时，密度、比热与温压依赖都成为模型条件。物理计算不是用一个“岩石会导热”的常识替代地球，而是把这些量组织成明确的近似系统。

初始条件与边界条件也承担不同责任。初始条件描述计时起点内部温度如何分布，边界条件描述表面在随后时期怎样与外界交换能量。把表面温度保持不变是有用的简化，却不是从地下梯度本身推出来的事实。若表面条件变化，其影响向深处传播需要时间；某些浅部信号可能主要反映较近历史，某些深部信号则涉及更长尺度。单一剖面中的不同深度，未必都在记录同一时间窗口。

由此可以理解，热学年龄为什么不是一项纯粹的算术竞赛。即使所有人使用同一微分方程，也可以因起始状态、物性分布和边界处理不同得到不同历时。共同的自然定律没有消除具体建模选择，反而使选择的后果能够被计算。真正需要比较的是完整的模型—观测组合，而不是把数字从其条件中剪下来排列。

◆ 11.3 争论进入假设内部

当不同推算不相容时，重新测量往往是最直接的反应。若梯度误差很大，改善测量确有意义；若误差已不足以解释差距，继续在同一模型里精修参数便触不到分歧的根源。霍姆斯 1913 年对佩里质疑的回顾，把地球内部传热性质置于关键位置。[A09] 按这份回顾所保留的问题，争论由“梯度读得准不准”深入到“梯度为何具有这种时间含义”。它涉及内部输运方式，尚不能直接等同于今天的地幔对流理论。

理解这种质疑可以从一个物理区别出发。表面附近温度随深度增加，能够提示热流方向；但同一个近表面梯度，未必唯一决定深部储热与向上传热的方式。一个全部静止、只靠传导的球体，与一个内部热量输送较有效、外层构成主要热阻的球体，可以在某一时刻呈现相近的表面梯度，同时具有很不相同的历史。观测相同不意味着观测产生的历史也相同。反演的不唯一性不是科学无能，而是提醒我们增加能区分模型的资料。

可以用两个教学容器说明。一个容器中热量只缓慢穿过固体，另一个容器的内部液体能够流动混合，但外面包着同样的隔热层。若只测外壳附近的一小段温度曲线，可能无法可靠分辨内部经历了怎样的冷却。这里不能由类比直接证明地球内部一定像第二个容器；类比的任务仅是展示逻辑可能性：一个观测量未必足以确定全部内部状态。要在地球上选择模型，仍需岩石物性、地震、热流和动力学等不同证据。

这个区别还拆开了两种经常被混合的修正。内部热源改变能量收支：系统不再只是消耗初始储热。内部物质运动改变输运方式：热量从深部到表面的路径和效率不同。两者都可能影响冷却推算，却不是同一个原因。若把模型失败全部浓缩为“当时不知道放射性”，就会掩盖即使没有增加热源，也能通过改变内部输运性质重新审查时间估计的可能性。思想史需要保留各条质疑的独立逻辑，不能让最后出现的知识吞并此前已经提出的问题。

从 SDE 分析看，初始显露 S 是一个附带模型条件的年龄。与地质记录和其他推算形成的差异序列 D，使原先固定的传热条件 E 成为争论对象。若目标由“用地表梯度求唯一年龄”转向“判断内部热输运与热史”，同一梯度就从时钟读数变成多种内部模型必须解释的约束。研究者接着需要选择对输运方式敏感的资料，而非仅仅增加同类梯度读数。这是对问题结构的复审；具体研究是否沿此路径推进，仍以论文中的未知量、采样安排和模型变化为据。

问题改道的证据应当来自论文中被改变的未知量、被新增的测量任务、被重新讨论的材料性质。年龄数字突然变大，只能说明结果不同；测量对象和推断责任随之变化，才说明研究组织发生了实质调整。这一区别尤其适合防止过度解释。若后来一篇论文只是更正导热率而保留原来的模型和目标，我们应该称之为参数修订。若它承认表面梯度不能独自计时，并转向多种传热机制比较，才有理由讨论更深的对象重构。

还应保留争论中的不确定性。指出某个年龄上限依赖强假设，不等于已经证明另一个具体年龄。对原结论的削弱，与新结论的建立，是两个任务。佩里式质疑的重要意义可以是解除一个过强的约束，而不是凭一次反驳就完成地球全部年代学。科学史若把每个有效批评都写成新理论的全面胜利，便会低估随后建立独立时钟所需的工作。

从今日方法论看，最能改变争论的新增资料，往往不是再测一个相同地点的梯度，而是能够区分深部模型的另一类观测。若两个候选热史在地表给出同样结果，那么重复地表

测量主要缩小共同结果的误差，却不能解除模型间的不唯一性。只有增加对内部状态、材料性质或独立时间的约束，才可能把它们分开。这里可以看到研究选择怎样由模型差异决定：问题不是笼统地“收集更多数据”，而是寻找对不同候选机制有不同响应的数据。

热流研究还要求区分能量储存与能量输出。地表每单位面积流出的热量，只是系统边界上的通量；要由通量求历时，必须知道内部储热及其变化。一个水库的出口流量不能单独告诉我们水库何时建成，热通量也不能脱离储量和补给独自成为时钟。

11.4 失败模型留下什么

冷却模型失去直接决定地球年龄的资格，并不意味着热流失去研究价值。相反，问题开始分化：地球有多少内部热源，热量怎样在不同圈层间迁移，哪些观测反映局部状态，哪些能够约束长期演化？年龄研究没有把热学赶出地球科学，而是重新安置了它的证明责任。成熟理论常常保留前一理论的方程，却改变方程适用的对象、边界与解释层次。

这一点也改变我们理解失败的方式。完全没有产生可检验后果的想法，可能被遗忘；产生了明确约束却在特定前提下失败的模型，能够留下新的测量议程。它使后来研究者知道应该测什么、哪些量不能互相替代，以及需要怎样的独立证据。失败不是天然值得赞美，但能被定位的失败比无法说明的成功更容易积累。地球年龄之争的知识增长，不仅表现为年数增大，也表现为错误有了可分辨的结构。

以今日审查一项年龄推算为例，可以依次问四个问题。首先，起算事件是否明确？其次，当前观测量通过什么过程与起算事件连接？再次，是否有其他过程能产生相同观测？最后，是否存在不依赖同一关键假设的核验方式？这套审查不偏袒热学或放射性方法。后来的矿物时钟同样必须回答封闭性、初始成分与事件归属的问题。它们不是因为出现得晚就自动免检，而是因为能够提供不同的记录结构，才成为更强的证据链。

区分这些转移，才能判断争论改变了什么。两组结果若因起算事件不同而分离，首先需要对齐对象；若只需更正一项局部测量，则属于观测修订；若原本用来读出年龄的梯度，转而成为研究内部输运的材料，才涉及问题结构的重组。后一种变化尤其值得思想史追踪：一个结论的失效，把先前隐藏在计算中的条件变成了新的研究对象。SDE 在本章的作用，是将这几种改动放在同一条证据链上辨别。

地质学家提出漫长时间，并不是争取一张可以无限填写的空白支票；物理学家要求能量约束，也不是取得裁决全部地史的特权。双方的有效部分在新的研究结构中同时被保留。地史需要物理可行性，物理模型需要历史证据检验。这种相互约束使时间成为能够进入计算、接受反驳、细分事件的科学对象。古老地球并未在十九世纪才开始存在；发生的是一个可以用不同证据相互检验的深时世界。

不过，时间很长还不够。即使各家同意地球足够古老，地层究竟记录长期缓慢变化，还是少数急剧事件，仍不能由总年龄决定。下一章转向渐变与灾变之争，将“长时间”与“慢过程”拆开；随后再看矿物怎样建立另一种时钟，并使过程速度有机会受到更细的数值约束。

第 12 章 渐变与灾变
把两派口号拆成可检验的问题

12.1 四种“相同”不要混用

沿着一段海岸走，悬崖上的碎屑不断落入海中。几小时内，岸线似乎没有改变；一场风暴过后，局部崖壁却可能后退明显距离。若把前一种观察称为渐变，把后一种称为灾变，两者完全可以发生在同一段海岸、同一套物理条件中。它们并不要求不同的自然规律。地质思想史中的长期争论之所以难以用这幅日常图景解决，是因为“过去与现在相同”这个命题，常把多种不同层次的相同压缩在一起。

第一层是规律是否相同。例如研究过去的流水作用，通常仍用重力、物质守恒和运动规律理解它，而不在遇到困难时临时设想一套专供古代使用的自然法则。这是一种使历史推理能够成立的基本要求。第二层是机制是否相似：我们熟悉的河流搬运、火山喷发、地震位移，能否解释某种古老遗迹？机制相似并不要求作用强度、持续时间、材料条件完全相同。第三层是速率是否均匀：同一机制是否始终以相同速度工作？第四层是状态是否稳定：地球的大陆配置、生物组成和气候分布，是否长期维持某种不变格局？

由第一层无法直接推出后面三层。同样的力学规律能够容许平静沉积，也能容许短时洪水；同样的侵蚀机制在不同降水和植被条件下可以具有不同速率；即使局部过程反复发生，整个地球的状态仍可能发生不可逆改变。反过来，发现过去有强烈事件，也不意味着自然规律曾经失效。地层中的突变首先要求查明记录和过程，而不是立刻选择一种宇宙哲学。

莱尔 1830 年《地质学原理》第一卷的题名明确提出，以现在正在作用的原因解释地表过去的变化。[A06] 这是一项研究纲领，其力量在于把研究者带到可观察过程面前。该卷实际讨论河流、洪水、海岸侵蚀、火山与地震，不能把它缩写为“莱尔只承认平缓变化”。同时，他对规律的连续性、过程解释的充分性和长期世界状态的判断，也不能都视

作同一命题。阅读原著需要追问每一处论证究竟支持哪一层，而不是把一个学派标签盖在全书上。

这种区分也改变反驳的方式。若某剖面显示一次快速沉积，反驳的是把该层解释成均匀慢沉积的具体主张；它未必反驳以现代泥石流或浊流理解古代快速沉积的策略。若发现某类古老机制在现代几乎没有可直接观察的同类，说明现代类比的经验范围不够，而不必然说明历史科学失去物理基础。反驳必须命中主张真正承担的责任，否则两派可以在不同层次上互相宣布胜利。

四种“相同”分开以后，历史争论的实质反而更清楚。规律稳定与世界状态恒定是两项不同主张；灭绝事实与一次灾难的全球范围也需要分别证明。关于地球是否经历长期方向性变化、生物替换能否归入既有过程，分歧仍然存在，但每项主张只能使用与自身相称的材料。这样才能辨认一种新观察究竟推翻了哪一层判断，又保留了哪些解释资源。

1830 年第一卷的目录本身就提供了防止误读的具体材料。莱尔在流水作用部分同时安排河道迁移、洪水、滑坡造成的堵塞和湖水溃决，在火成作用部分讨论火山与地震。[A06] 这些过程并非全部缓慢温和。他的研究纲领能够把剧烈事件纳入现今原因的范围。由此可见，猛烈程度与是否采用可观察原因，是两条不同轴线。

该卷还讨论气候变化与海陆配置的联系。无论这些具体解释后来如何修订，其问题结构已经承认地理状态可以不同：纬度相近的地方，因海陆分布、洋流与地势差异而具有不同气候。把莱尔理解为过去每一处都与今天完全相同，便无法解释他为何需要研究这些变化。真正需要逐项辨别的，是他认为哪些总体条件保持、哪些区域状态可以轮换，以及这些判断各自依赖什么材料。

以河流三角洲为例，长期平均增长与一次洪水沉积可以同时成立。测某一年带入的泥沙，不等于测整个三角洲净增长，因为海流可能搬走部分物质，河道也可能改道。若把年输沙量直接乘以漫长时间，隐含了保存比例与流域供沙不变；若只根据一次特大洪水想象全部地层，又忽略事件之间的侵蚀和重工。这个例子显示，渐变与灾变争论需要转成物质收支、保存效率和时间代表性的问题。

12.2 居维叶究竟证明什么

居维叶的力量首先来自比较，而不是来自灾难场景。若出土骨骼与任何已知现生动物都不相同，研究者需要说明差异是否只是个体、年龄或性别变化，是否能够组成同一个稳定的解剖类型，以及这种类型是否可能仍存在于尚未调查的地区。灭绝作为科学结论，需要一条从标本身份到现生比较范围的推断链。单独一块陌生骨头，不能立刻证明一个物种已经消失；但相互关联的结构特征能够逐步排除一些替代解释。

1827 年英译第五版《地球理论论说》包含居维叶正文，也包含詹姆森所写的序言及地质说明。[A07] 这本书是多个作者和版本层次共同构成的文本。书中一句话印在同一封面之下，并不意味着都由居维叶提出，更不意味着可以不加区别地当作更早法文版本的原话。对于灾变论的宗教解释尤其如此：必须分清自然材料支持的判断、作者的历史推论，以及编者附加的解释。否则，科学史实际上是在批评一个由后人拼装的对手。

化石比较能够支持“过去存在过今天没有发现其活体的动物类型”，却不能单凭这一点确定灭绝的速度。一个物种可能在很短时间消失，也可能经历长时间衰退；一种动物的当地消失，还可能只是迁徙或栖息地改变。要从地方缺席推向全球灭绝，必须扩大地理比较。要从灭绝推向突发灾难，则还需要事件同步性、沉积结构或其他环境异常的证据。每扩大一次结论范围，证明责任都随之增加。

从这个角度看，居维叶的研究提出了一个不能由“自然一向如此”消解的问题：过去生物界确有不同的构成。相同器官之间的比较，使已不存在的生命组合获得可辨认身份。那些骨骼不是今天动物的残破版本，也不只是奇异的自然物，而成为能够讨论出现、分布与消失的历史对象。承认这种对象的发生，不等于说古生物是在研究者命名以后才曾经活过；它意味着可检验的灭绝世界，依赖标本、比较解剖和现生调查相互规定而形成。

这里的 S 可以是一个被确认的灭绝动物类型，D 是它与现生及其他化石类型的结构差异序列，E 则包括器官间对应、标本保存和分类标准的约束。若保存破坏使某一关键部位变形，差异的意义可能改变；若多个部位在独立标本上稳定共现，分类判断便增强。分类结果又会引导下一轮寻找哪些骨骼、在哪些地层检验其存在。这一往返能被标本和论文记录追踪，不必借助对科学家心理瞬间的想象。

这一分类对象也可能被撤销。若后来的标本证明所谓新类型只是已知动物的另一生长阶段，原有差异便须重新解释。关键在于独立标本之间能否保持一组稳定的结构联系，而不在于命名时曾有多大新奇感。把标本鉴定、灭绝判断与灾难原因分开，可使有效的比较工作在总体解释调整后继续使用，也使灾难解释承担自己的沉积与年代检验。

居维叶所代表的历史问题，还提醒我们注意证据保存的不对称。骨骼容易保存的动物和环境，可能在化石记录中被放大；不易保存的生命与阶段，则可能留下很少踪迹。看到某类动物在一层之上不再出现，既可能反映真实灭绝，也可能反映保存环境变化。这个限制不是用来无限推迟结论，而是指导我们寻找不同环境、不同地区的相关记录。一种解释若只在最有利的剖面成立，就尚未承担跨地区历史判断的全部责任。

居维叶讨论灭绝动物与现生种是否只是变种时，面对的并非纯粹的名称争执。[A07] 若两个骨骼组合可以通过已知的个体变异连接，它们可能属于同一生物类型；若一系列相互配合的结构差异无法这样解释，就需要考虑不同类型。研究者因此既要看单个部位，也要看各部位如何共同组成身体。比较解剖的力量来自结构关系，而不是把任意一个差异放大为物种边界。

保存还会改变可比较单位。一副完整骨架允许同时观察多个部位，散落骨片则可能被错误组合为一个从未存在的动物。来自同一地点并不保证属于同一个体，来自同一地层也不保证同时死亡。标本采集与组合程序因而参与对象身份。若不记录出处，后来分类者可能拥有更精密的形态知识，却无法补回遗失的关联信息。古生物学的历史因此不仅是理论史，也包含标本的标签、位置与保存史。

◆ 12.3 同一剖面的不同读法

现在构造一个教学情形：某海相剖面下部富含一组化石，上部出现另一组，二者之间变化很突然。研究者甲认为发生了快速生态危机；乙认为中间长期记录缺失；丙认为主要是沉积环境迁移，使当地生物组合变化。三者都可以对同一露头讲出合理故事。但科学比较不能停在故事的可想象性上，必须要求它们承担不同的预期。

如果是快速事件，理想证据之一是在沉积较连续、环境条件不同的多个地点，仍能观察到经过独立相关后近乎同期的变化。若是长期缺失，应寻找侵蚀面、沉积间断、地层厚

度异常或相邻地区保留的中间序列。若是环境迁移，应检验生物变化是否跟随岩相、水深、盐度或基质变化，以及原组合是否在同时期其他地点继续存在。这里并不存在一项万能判据；结论来自多种约束是否指向同一解释。

剖面上的厘米不能自动变成年份。厚层可能快速堆积，薄层可能凝缩很长时间；一条不明显的接触面可以包含大段缺失。若地质学家把单位厚度默认为单位时间，就会把沉积记录的几何形状误读为自然过程的速度曲线。反过来，若任何突变都被解释为“缺了一段”，缺失假说也会失去可反驳性。它需要指出缺失发生的过程和证据，而不是专门用来保护均匀变化的信念。

可以进一步区分事件持续时间与记录分辨率。设真实变化历时很短，但采样间隔很大，观测可能只能把事件限制在一个宽区间内；这个区间是认识的不确定范围，不能当成事件本身的长度。若真实变化较慢，却恰有中段被侵蚀，保存下来的两端可能紧邻，造成急剧转换的图像。正因为两种误读都可能发生，判断不能只靠肉眼觉得变化“突然”或“平缓”。必须同时说明取样密度、沉积连续性和时间约束。

多指标一致也不是自动独立。假如化石变化、化学曲线和某个年代估计都依赖同一条相关线，三者同时落在界线上可能部分来自共同处理，而不全是三个独立证据的会合。分析时需要追问哪一项被用来定义界线，哪一项是在界线确定后才独立检验。如果用化石突变确定事件，再说事件发生时化石突变证明定年正确，就可能出现循环。区分定义证据与核验证据，是将好看的相关图转成可信历史推断的必要步骤。

在这一教学情形中，过程假设与地层分辨率会彼此改变。若怀疑突变来自间断，采样就需集中于接触附近，并向邻区寻找被省略的层段；若怀疑环境迁移，则需沿岩相变化追踪化石组合。新增差异又可能推翻最初的间断或迁移解释。S＝F(D,E)在这里记录的是事件显露、差异序列与保存—相关条件的联合关系，具体判断仍由沉积结构、化石分布和年代约束完成。把这些联系写清，才能看见为何同一批材料会在不同假设下引出不同调查路线。

对历史争论的重新理解也由此展开。所谓灾变论者未必否定所有渐进作用，所谓均变论者也未必否定一切猛烈事件。真正需要审查的是某人面对某段记录时提出什么可区分的预期。如果两套观点在拆分之后分别回答不同的问题，例如一套强调日常作用的累积，一

套强调生物灭绝的事实，那么二者可能互补。只有当它们在同一时间尺度、同一地区与同一待解释现象上产生不相容判断，才形成严格意义上的竞争。

进一步考虑两个剖面显示不同的灭绝顺序。甲剖面某类化石先消失，乙剖面另一类先消失。一个可能原因是真实生态变化具有地域差异，另一个可能原因是罕见类群在取样中更早停止被发现。最后一次观察到某种化石，与它真实存在的最后时刻之间可能有间隔。若采样密度不同，这种间隔也不同。把观察终点直接当作生物死亡时刻，会使地方记录人为地产生不同事件次序。

可以通过加密取样、比较丰度变化和扩大剖面范围来检验这种解释。若某类化石在消失前已非常稀少，缺失更需要谨慎；若多个丰富类群在保存条件稳定的层段同时消失，则突变解释得到更强支持。这里不是要求绝对完整的记录，而是让保存概率进入推断。事件能够被可靠识别，往往因为不同失效方式受到同时约束，而不是因为研究者找到了完全没有缺陷的单一剖面。

也要区分地方灾害与全球事件。一个厚砂层可能记录附近水道的强烈搬运，却不说明整个地球同时发生同类过程。要扩大范围，必须建立跨地区相关，并检查时间误差是否允许声称同步。若几个事件的年龄范围宽且彼此重叠，只能说目前无法区分先后；不能把这种分辨不足改写为已经证明同时。全球性是一项需要专门取得的结论，不是灾难规模在想象中放大的结果。

沉积物的内部结构能够帮助区分一些候选过程。例如颗粒由粗到细的垂向变化，可以提示水流搬运能力在一次沉积过程中减弱；侵蚀底界与混入的旧地层碎块，可以提示先发生冲刷。但这些结构仍需结合区域位置判断：相似的粒度排列可以出现在不同环境，单层内部的快速沉积也不代表相邻整个地层单元都快速形成。事件层与长期层序是嵌套关系，不能互相替代。

沉积之后的扰动还可能模糊原始记录。生物活动、压实和后期变形会改变层理可见性，使本来细分的事件合并为外观均一的一层。观察到均一并不一定意味着过程均匀，观察到分层也不一定意味着每层对应相同历时。地层学的任务，是把沉积、保存与后来改变分开重建。渐变或灾变的判断，只有经过这条记录生成链，才真正指向过去的自然过程。

时间分辨率本身也应进入结论措辞。若资料只能分开百万年尺度的变化，就不能据此判断某个过程究竟持续千年还是十万年。把较细问题保留为未决，并非放弃历史解释，而是使解释强度与记录能力相称，等待更合适的时钟。

◆ 12.4 保留双方的有效成果

现代地球科学能够同时研究缓慢板块运动、短时地震破裂、长期沉积与突发搬运，并不是在两个阵营之间折中取中间速度。它把速度本身变成需要测定和解释的变量。缓慢积累可以使系统接近阈值，超过阈值后又产生快速变化；短时事件造成的地形与物质分布，也会改变随后长期过程的条件。长期与短期不是两种互不接触的世界，而是在具体机制中相互连接。

阈值解释需要指出具体控制量及其累积方式。例如在一个教学模型中，若坡体强度决定失稳门槛，缓慢改变的水分或载荷便可能引起短时位移；位移又会改变后续排水与受力条件。这里可检验的是变化前后的状态及其顺序。仅见突变，尚不足以确定跨过了哪一种阈值。同样，现代过程产生相似沉积结构，只提供机制候选；古代事件的规模、频率和背景仍需由其自身记录限定。

另一组容易混入的争论是水成与火成。某岩石形成于水中沉积、化学沉淀还是熔体冷却，首先是机制问题。沉积可以很快，熔体冷却也可以很慢。把水成论直接归为渐变、火成论直接归为灾变，会把两个不同坐标轴压在一起。机制、速度、历史状态和规律的一致性应分别分析，再讨论它们在某一作者体系中的组合。一个理论可能在机制判断上错误，却提出了有价值的区域层序比较；也可能识别出正确机制，却把它无限推广。

保留双方成果，意味着不再要求历史上的人物以全胜或全败退出舞台。居维叶的化石比较与灭绝问题，不因某些灾难叙事受到修正便消失；莱尔把现今过程变成解释资源的研究纲领，也不因承认快速全球事件而失效。与此同时，我们不能把所有现代结论都回溯为先驱早已知道。历史贡献有边界，后来的综合是在新材料与新方法中完成的，不是把两本旧书并排摆放便自然出现。

两种解释的分界因此能够落实到记录上。一个在粗采样剖面中显得突然的事件，若在连续剖面中展开为长时变化，其速度判断就应调整；一种均匀变化解释，若只能不断加入

无法定位的缺失以回避反证，其适用范围也应收缩。还有第三种结果：两套观点分别解释作用机制与生物更替，分清命题后并无直接冲突。承认这种分工，才能把真正需要年代裁决的争点保留下来。

这也把我们带回时间问题。只知道地球非常古老，不能告诉我们一次灭绝持续多久；只知道化石组合不同，也不能给出两层之间的数值间隔。需要一种与沉积厚度及过程想象不同的时钟，才能进一步约束速度。矿物中的母体与子体比值提供了这种可能，但它也带来新的困难：读数属于哪个事件？样品是否保存了那次事件的记录？下一章将从霍姆斯的矿物比较进入这些问题，再到煤山检验一个高度精确的数字怎样取得地质身份。

第 13 章 从霍姆斯到煤山
矿物时钟怎样取得事件身份

13.1 从衰变概念到可用标本

一粒矿物可以比包围它的岩石更古老，也可以比岩石的主要部分更年轻。前一种情形并不神秘：较老岩石被风化，碎屑经过搬运，进入新的沉积物；后一种情形也不罕见：岩石已经存在，后来流体进入裂隙，形成新的矿物。于是，一项看似简单的要求——测出这块岩石的年龄——在进入实验室之前，就必须拆开。究竟要知道颗粒结晶、沉积物堆积，还是后来一次改变的时间?

这个区别并非今天才回头提出的苛刻要求。霍姆斯在 1913 年《地球的年龄》第十章，已经把矿物与围岩是否同龄作为计时工作的基本问题。他区分火成岩中参与凝固的矿物、沉积岩中先已存在的颗粒，以及后来形成的胶结物或充填物。他讨论的工具和常数属于当时，但“数字必须归属于某种形成过程”的责任已经清楚出现。[A09]

放射性给地质学提供了一种不同于厚度或热梯度的时间线索。时间过去，母体核素减少，衰变产物积累。若知道转化规律，又能判断样品保存了什么，今天的组成便可能约束经过的时间。这不是从矿物表面读出一个年代标签，而是把现在的组成与可能的历史连接起来。

为看清这一步，可写出一个现代教学模型。设单级衰变体系计时开始时母体原子数为 N_0，经过时间 t 后剩余 N；系统封闭，衰变常数 λ 不变，则 $N = N_0\exp(-\lambda t)$。由这批母体衰变新产生的子体数 $D^* = N_0 - N$，故 $D^*/N = \exp(\lambda t) - 1$，进而 $t = \ln(1 + D^*/N)/\lambda$。$D^*$须扣除初始子体和外来子体；若两者可以忽略，测得的子体量才可直接代入。λ 的单位为时间的倒数。这里的 D^*仅为计时符号，与 SDE 中的差异序列 D 分开使用；这一单级模型也只用于说明推断结构。

公式中的星号承担着重要责任。仪器能够测出某种子体的总量，却不因此自动知道其中多少由目标母体衰变产生，多少在计时开始前已经存在，多少后来进入。若把这些成分全部放入 D^*，求出的时间可能很精密，却不是想求的时间。所谓精密，是重复测量之间

的离散程度小；所谓事件归属正确，是测量与目标过程之间的连接成立。两者不能互相代替。

同样，封闭也不是一句“岩石很坚硬”就可以保证的性质。它指在所研究的时段和所采用的计时体系中，相关组分的输入输出是否足以影响时间解释。样品外形没有破坏，不等于化学体系从未变化；样品经历变化，也不意味着全部记录都消失。实际研究要判断哪些部分保留了哪段历史。

霍姆斯列出的前提包括初始铅或氡的影响、后来的增减以及产物来源。他还把新鲜、稳定、原生的矿物与更容易改变的材料区别对待。这使我们看见早期放射性定年工作真正的难处：衰变知识只打开一扇门，标本的形成方式与保存状态决定能否从门内通向某个地质事件。[A09]

在本书的分析中，矿物时钟不是先有一个完整的年龄对象，再替它寻找测量手段。年龄的科学规定性，在目标事件、样品选择、化学分析和地质比较之间逐步形成。例如，把目标从结晶改为沉积，并不是只改题目中的两个字；它会改变可接受样品的范围、对旧颗粒的处理以及年龄结论的表达方式。对象与操作在这里共同发生。

材料的先后关系为这种对象形成划定了边界。一个早于沉积的碎屑颗粒，可以限制物源或给出沉积年龄的上限，却不能直接充当沉积时刻。若改测后生胶结物，所得年龄又可能限制沉积之后的过程。目标事件不同，同一材料所能给出的界限随之改变；研究者的辨认能力，就体现在能否将这些不同时间关系稳定地转为取样与解释规则。

◆ 13.2 一致比值的力量与限制

霍姆斯在第十章讨论了同龄矿物的铅与铀比例以及不同地质年龄材料之间的比较，并指出仅有化学分析、缺少产状与地质关系，会削弱解释。他对早期分析资料的整理表明，化学相似性和地质先后不能彼此隔绝：若一个时钟声称计时，它应当与独立建立的先后关系接受对照。[A09]

这种对照有两种不同用途。第一种是检查同一目标是否能够重复得到接近的结果。第二种是检查不同目标是否按照预期呈现差别。前者检验稳定性，后者检验区分力。一个方法若对所有标本都给出相同年龄，重复性可以极高，却没有重建地史的能力。反过来，结果十分多样，也不意味着识别出了丰富历史；多样可能来自样品破坏或测量偏差。

我们可以构造一个不对应真实样品的教学例子。某岩体中取出三件矿物，分别分析，得到彼此接近的数值。此时至少有三种解释：它们记录同一次结晶；它们原先不同，却在后来一次过程中被共同重置；或者分析共享同一校准偏差。仅凭结果接近，无法在三者之间作出最终选择。重复测量增加了约束，但增加了哪一种约束，必须说清楚。

如果又发现岩体切穿一套更老地层，并被一套更年轻沉积覆盖，那么野外关系给出了不同来源的先后限制。若测年结果与这些限制明显冲突，研究者不能仅凭仪器稳定就驳回地质关系，也不能仅凭传统层序就驳回数字。需要检查各自的对象：切穿关系是否识别正确，所测矿物是否属于原生结晶，覆盖是否真实而非断层接触。冲突首先打开的是核验任务，而不是让某一学科自动胜出。

这正是“特征纠缠”需要比“有关系”说得更多的地方。矿物的保存状态限制同位素比值可以怎样解释；比值与层位不合又迫使重新检查保存状态。野外关系限定事件区间，事件区间反过来帮助识别哪些矿物可能记录另一阶段。两项特征不只是并列存在，而是改变彼此取得证据资格的条件。

还要防止互证退化为循环。若研究者先用预计年龄删除不符合的样品，再以剩余样品证明预计年龄，所谓一致便可能由筛选制造。解决办法不是禁止筛选，而是公开筛选依据，并尽量让依据具有独立性。因明显蚀变而排除某颗粒，与因它不符合期望数字而排除它，并不是同一回事。排除记录也应保存，因为它可能包含另一地质过程的线索。

独立性不是全部具备或全部没有的二值性质。几件样品可以具有独立的野外位置，却共享示踪剂；两种方法可以使用不同核素，却共享同一错误的事件归属。实际论证应逐层描述依赖：取样是否独立，实验过程是否独立，校准是否独立，地质解释是否独立。只有这样，增加材料才可能真正增加判断力。

用 S＝F(D,E)表述本节，S 是获得事件身份的年龄判断，D 是各样品及不同证据之间有次序的比较，E 则是保存、形成与校准特征的相互限定。比值相近、空间位置相邻、共享同一校准，各自构成不同关系，只有其中与目标事件相连的部分才能支持共同年龄。三式在此帮助区分证据的作用，具体年龄仍由核素转化模型计算。

D＝G(S,E)把问题接到下一轮取样。若要区分原生结晶与后来重置，仅增加同类颗粒可能继续得到相近读数；选择不同矿物世代、颗粒内部不同部位，或比较受改造程度不同

的样品，才可能使两种历史给出不同预期。这些是由目标和保存条件提出的候选方案；实际采用哪一种，要由具体研究的样品说明与操作记录确定。

一致性的判断还取决于比较的分辨率。假定两件材料的估计值有轻微差别，但各自的不确定范围较宽，我们不能仅凭末位数字不同就宣布它们记录两个事件。反过来，误差范围相交也不构成“同一事件已经证明”的充分条件。统计相容只说明当前资料未必能区分，事件同一性还需要形成过程与地质关系的支持。

这意味着“尚未分开”与“本来就是一个”必须在书写中区别。一个时代的研究可能只能辨认岩体形成的大致时期，后来更高分辨率工作可能提出多个阶段。但后来的细分并不使早期全部结论失效；早期判断也许在自己的粒度上仍然有效。思想史需要保存这种层级关系，不能把每一次细化都写成全面推翻。

还有一种更隐蔽的情况：平均值很稳定，分布却发生了变化。若两组颗粒的年龄分布不同，简单合并后得到同一个平均值，可能会抹掉值得解释的内部结构。是否应该求一个平均值，必须先问这些颗粒是否有理由被视为同一目标事件的样本。统计操作本身不能替我们完成地质归类。

这些讨论引向一个可以实际使用的判断顺序。先明确目标事件，再说明样本为何可以共同代表它，随后处理读数与误差，最后把结果放回野外关系。这个顺序是本书的建议，不是唯一合法程序。遇到异常时，研究可能反向进行：先由数字差异发现样本分组需要重审。顺序可以往返，但每次跨越都必须交代依据。

真正有用的统一，并不是把所有读数压成一个数字，而是让不同读数在同一解释中获得恰当位置。一个较老颗粒可以约束物源，一个较年轻域可以提示后期作用，一组原生颗粒可以参与结晶时间估计。它们不必相等，仍然可以共同提高对岩石历史的理解。差异不总是统一的敌人；未经区分的差异才会妨碍有意义的统一。

◆ 13.3 蚀变为什么改变问题

一个不符合预期的年龄，可能是实验错误，也可能是另一个真实过程的记录。若这两种情况被混为一谈，研究会在两个极端之间摆动：要么把异常全部清除，要么把异常全部宣布为重大新发现。真正困难的工作是确定异常发生在哪个环节。

霍姆斯对铅与氦计时的讨论已经将记录保存纳入可靠性判断。他关注矿物是否原生、是否发生改变，也讨论氦的保存问题。[A09] 这使材料选择成为计时论证的一部分：即使同受一种衰变规律支配，不同矿物保存产物的能力和经历仍可能不同。核素规律提供共同计算基础，矿物形成与后期变化决定这种基础可以用于哪一段历史。

下面以刚才的简化公式作思想实验。保持母体 N 不变，若一部分放射成因子体逸出，则 D*/N 降低，由公式反算的表观时间变小。反之，若额外子体进入却被误当作原位衰变产物，表观时间可能增大。真实矿物体系的变化往往比这复杂，母体也可能迁移，不同组分可能分属不同域，因此这只是解释误差方向的简化推演，不是某件样品的诊断。

思想实验揭示了一个关键区别：测量设备可以准确报告今日的组成，年代解释仍然可能失败。失败不一定发生在读数阶段，而可能发生在把读数连接到过去的模型阶段。增加仪器精度只改善前者，未必修复后者。若对象身份错误，小数位越多，读者反而越容易获得不应有的确定感。

这也说明为什么一个原先被视为不合格的标本，在问题改变后可能变得有价值。若研究目标是原始结晶，后期改变会干扰目标；若研究目标是后期作用的时间和范围，同样的改变便可能成为研究入口。不过，这种资格转换需要能辨认具体作用的证据，不能把所有异常都笼统称作“重置”。

科学对象的发生由此表现为一组可追查的变化。旧问题将样品分成可用与不可用，新问题重新划分这两类，并要求观察先前忽略的蚀变部位；新的比值分布再与矿物世代和切穿关系比较，改写样品内部的事件序列。完整的回写应当在这条链上留下材料：哪一类样品被重新采用，分析部位怎样变化，结果为何支持新的归属。缺少这些连接，便只能确认问题有了新的提法。

本书用 P1—P3 分析相位复审这类转折时，会问一个更局部的问题：哪些既有规定已经不能同时满足？例如，若坚持所有矿物只记录一次形成，同时又承认经过检验的差异，那么研究可能被迫细分事件身份；若目标身份保留，则可能需要改变取样和处理路径；若路径已经充分核对，可能需要重审保存与组成之间的关联。三种改变不是每次必须出现的三幕剧，而是需要分别取证的候选方向。

这种分析保留了失败可能。若最后发现异常仅来自标本标签错误，纠正标签即可恢复一致，我们就不能据此声称发现了新地质事件。标签管理可以具有重要价值，但它与建立

一种新计时对象承担不同的创新责任。发生学的细致，体现在能辨认变化的层级，而不是把所有纠错都放大成革命。

矿物时钟因此既是一项测量技术，也是一种逐渐形成的地史辨认能力。它要求研究者把手中的颗粒、地层中的位置、实验室中的比值与事件解释放在同一条可复查链上。任何一个环节断开，都可能让一个漂亮数字失去原先声称的意义。

◆ 13.4 煤山的样品年龄与事件身份

在煤山讨论年代问题，最容易发生的错误是把几种不同对象压成一个数字。火山灰中的锆石、火山灰层的沉积、化石组合的变化、主要灭绝区间以及国际地层界线，都可以参与时间研究，却不能因为位置相近就被当成同一事件。

Burgess、Bowring 与 Shen 在 2014 年的研究中，使用五个火山灰层的锆石铀铅定年建立煤山年龄框架。[K02] 五个层位提供了分布于剖面中的时间锚点，使生物记录得以进入数值年代比较。样品年龄到生物事件时间之间，仍须经过层位关系和年龄模型。以下分析集中于这条连接，不重算原论文的单颗粒数据。

把一个火山灰层用于约束附近事件，要经过若干判断。首先，所测颗粒是否代表与该灰层相关的岩浆结晶，而非混入更老材料。其次，颗粒结晶与火山灰沉积之间的间隔能否忽略，或需要另行处理。再次，灰层与目标事件之间是否存在可用的地层关系。最后，若事件位于两个定年层之间，如何从两个锚点估计中间位置的时间。

这些判断共同决定年龄框架的分辨能力。若结晶先于沉积且间隔不可忽略，锆石年龄首先约束的是结晶；若灰层经过再搬运，它与所在层位的时间关系又会改变。只有排定这些关系，火山灰层才能作为附近生物事件的计时锚点。仪器提高的是同位素测量能力，锚点取得地层意义还需要另一组证据。

以现代教学推演说明插值。设下、上锚点高程分别为 z下、z上，对应距今年龄为 t下、t上，通常 z上＞z下而 t下＞t上。事件位于高程 z 时，令 f＝(z－z下)/(z上－z下)。在持续沉积且这段厚度与时间呈线性关系的简化条件下，t事＝(1－f)t下＋ft上，其中 0≤f≤1。于是 f＝0 返回下锚点，f＝1 返回上锚点，中间位置按比例分配历时。若存在停积、侵蚀或显著速率变化，厚度比例就不足以决定时间比例，须另建年龄模型。

提高两个锚点的测量精度，不能自动使线性插值成立。这里至少有两类不确定性：锚点年龄的误差，以及从锚点到事件的模型误差。前者缩小而后者未处理，事件时间的可靠性可能没有按想象增长。若两个锚点共享校准成分，估计时间差时还需要处理误差相关性，不能机械把每项误差都当成相互独立。

因此，增加一个新的定年层不只是往表格填一行数字。它可能检验原先的插值是否合理，也可能显示某段厚度并不对应原先估计的持续时间。此时应具体判断改变层级：若事件定义与先后都保留，只是时间窗口收窄，这是既有对象的精化；若两个原先无法区分的过程获得可辨的先后关系，研究对象的结构才可能出现更强变化。

年龄框架的使用还须区分数据版本。该论文 2014 年的勘误对象是图 3 及其图注。[K03] 更正应按所涉图件和说明处理，不能由此推出全部定年数据失效；引用旧图也会把已修订的表达重新带入论证。本章未逐项复核更正图像，因而只讨论锚点与事件的连接结构，不借该图增加数值结论。

在 SDE 分析中，事件年龄的显露 S，依赖样品比较与层位排序的 D，以及结晶、保存、沉积和事件定义互相限定的 E。形成的时间窗口又使研究者能够提出更细问题，例如两个候选过程在现有分辨率下是否可排序。问题改变可能要求新的取样，但是否真的发生这种研究改道，必须由后续论文和样品记录证明。

煤山案例把矿物时钟的两层工作清楚地接起来：先判断样品记录什么过程，再判断这个过程怎样限制目标事件。年代学的样品审查与误差分析已经承担了这两层任务；SDE 重述所追踪的是两者如何共同改变研究对象。例如测量更精密可能只收窄既有时间窗，也可能使两个原先合并的过程有了可辨先后。后者还需落实到层位和样品记录，才构成事件结构的改变。

从矿物走到事件，计时对象已发生一次重要扩展。但矿物结晶并不等于整个地球形成，现存最古标本也不能独自给出行星的生日。下一章先追问帕特森为什么借助陨石资料约束地球年龄，再由第十五章讨论地方剖面怎样承担全球公共时标的责任。两个问题分别涉及推断对象的层级与共同标准的建立，不能由一个精确数字同时包办。

第 14 章 帕特森的地球年龄为什么要离开地球岩石

14.1 最古岩石不是整体年龄

一颗矿物能够保存很古老的同位素记录，却不能独自交出整个地球的出生证明。这个区别看似常识，实际上决定了地球年龄研究为何需要走出地球岩石。若在地球上找到一件很古老的标本，我们首先知道它所记录的某个事件发生在那时；在合理条件下，还可以说地球至少同样古老。但“至少”与“恰好”之间有一道无法用更精密仪器直接填平的逻辑缺口。

设想一座不断修缮的城市。现存最古建筑的年代约束城市历史的下限，却不能保证城市恰好在建筑落成时诞生。早期建筑可能毁坏，旧材料可能重新使用，城市边界也会改变。地壳更新比这个类比更复杂，但推断结构相似：熔融、变质、侵蚀、搬运和再结晶，使现存岩石成为经过筛选的历史档案。没有找到更古老岩石，并不等于更早时期不存在。样品保存的上限，不能直接改写为对象历史的起点。

即使某颗锆石保存早期结晶年龄，它所见证的也是特定材料在特定条件下形成晶体并保留记录的过程。行星可能在此前已经聚集大量物质，也可能在此后继续获得质量、发生撞击和分异。若把“地球形成”理解为一个有持续时间的过程，就还需要说明采用哪个节点作为代表：最初固体的出现、主要物质聚积、地核分离，还是某套同位素系统开始近似封闭。不同节点可彼此相关，却不必重合。

明确事件类型，使不同年龄能够在同一历史框架中比较。矿物、岩石、地壳和行星各有自己的形成与保存过程，使用相同的“年”作为单位，只解决计量尺度，并不消除对象层级的差别。越接近行星整体，越需要把多个局部记录连接起来：局部结晶何时发生，它与物质聚积或分异相隔多久，哪些早期记录已经丧失？这些连接构成由样品年龄推向地球年龄的实际工作。

帕特森等人的研究因此不是简单地发现一块更老的地球岩石，而是转换比较材料和问题结构。地球表面的岩石屡经改变，某些陨石却可能保存太阳系早期物质演化的信息。如果这些材料与地球具有可论证的共同起源条件，它们就能参与约束地球的形成时代。关键不在陨石来自遥远地方，而在它是否保留了地球样品难以单独提供的初始组成和演化关系。[A10][A11]

这次取样改道值得注意：研究目标仍是地球，证据范围却延伸到非地球材料。外部材料要有用，至少应在初始组成或形成时代上与地球可比，同时保留地球岩石中已难辨认的关系。共同来源只规定可比较的起点，各自后续分异和交换仍需分别处理。因而取样范围的扩大与迁移条件的收紧同步发生；地球年龄的结论能有多强，取决于这些具体联系，而不能由“太阳系整体”这一概称包办。

◆ 14.2 1955 年的方法入口

1955 年帕特森、蒂尔顿和英格拉姆的论文，将地表铅与铁陨石中的铅同位素比较作为估龄入口。[A10] 理解这个思路，需要先区分总铅含量与铅同位素组成。元素相同，不同同位素的来源和演化却可以不同。放射性铀衰变产生的铅会随时间积累；用适当的非放射成因参照同位素归一化，可以使不同材料的组成进入共同的比较坐标。

用现代符号作教学推演，可以看清这种比较怎样取得年龄意义。设 x 为 $^{206}Pb/^{204}Pb$，y 为 $^{207}Pb/^{204}Pb$。一组材料若具有共同初始铅组成 x_0、y_0，并自同一时点开始封闭演化，今天的比值就由初始部分与放射成因增长部分相加而成。^{238}U 与 ^{235}U 分别生成两种铅，增长速率不同；当母体同位素之间的比例受约束时，消去各样品的铀铅相对含量，便可能得到年龄相关的线性关系。以下只推导这一结构，不代入 1955 年的原始数据。

这一关系的核心并非所有样品具有相同比值。恰恰相反，它需要不同材料保留不同程度的放射成因增长，同时这些差异能被一个共同演化模型联系。若所有点完全挤在一起，无论测量多精确，斜率都难以稳定确定。差异不是测量必须消除的杂音，而可以成为时间信息的来源；但只有符合共同模型的差异才具有这种意义。任意收集成分不同的石头，不能保证得到一条有年龄含义的直线。

低铀材料的价值也在于此。若某种材料中铀相对很少，其铅组成在相关时期中较少受到内部放射成因增长改变，就可能接近初始组成的参照。可是“较少改变”需要检验，不能因为样品叫铁陨石便自动成立。材料是否经历后期交换，是否混入外来铅，相关矿物相是否可靠保存早期信号，都影响它能否承担基准责任。

这个推理与上一章的单矿物时钟既连续又不同。连续之处在于都依赖核素转化、初始条件和保存史。不同之处在于，研究者不只是问一件样品积累了多少子体，而是通过多材料之间的关系重建共同演化约束。初始成分不必在每一件材料上独立知道，但它不能完全不受约束。用比较消去一个未知量，是一种数学能力；证明可以这样消去，则是材料科学与历史解释的任务。

测量污染在这种研究中尤其值得警惕。若样品中目标铅极少，微小外来铅就可能改变比值。这里污染不只是把数值弄得“有点不准”，还可能把本来属于不同形成史的成分混合，生成看似有结构的分布。空白控制、化学分离和仪器校准因此不是高深理论之后的杂务，而是决定观测对象究竟是什么的一部分。没有这些操作，所谓原始铅可能主要是实验过程带来的铅。

多样品拟合使共同演化假设接受联合检验。若某些样品偏离拟合关系，可以检查它们是否经历不同交换、是否存在污染或初始差异；若样品全部接近直线，还须排除混合形成同样图形的可能。矿物学、各材料的形成史与独立计时因此承担不同角色：前两者约束线的来源，后者检查时间解释。直线是一种关系证据，其年龄含义来自这些限制的共同成立。

初始铅在这种比较中取得了新的数学位置。对单个样品，x_0 和 y_0 是影响年龄反演的未知量；对一组符合共同模型的样品，它们规定拟合线必须经过的共同初始点。低铀参照可帮助限制这个点，多种母体含量则帮助限制线的方向。两种材料功能不同，缺少前者会增加初始条件的自由度，缺少后者则会使年龄斜率难以辨认。数学消元的可用性，由这种有分工的材料组合支撑。

用两点总能画出一条直线，却不等于有充分的演化检验。增加具有不同组成的样品，可以检查它们是否仍符合相同关系，也可以暴露某些样品具有不同经历。特别重要的不是样本数量本身，而是它们覆盖了怎样的组成范围。十件几乎相同的材料，可能不如少数具

有适当差异且保存史清楚的材料提供更强斜率约束。采样设计因此从“多取一些”转为“让不同材料承担不同辨识功能”。

参照同位素的作用也不应被误解为绝对不变的天然基准。归一化使组成能够比较，却要求分析过程准确区分各同位素，并控制干扰和污染。若参照量误差同时影响横纵轴，点的偏移方向就有关联。一个看似狭长的点群，可能同时反映自然演化与共同测量结构。读图时需要把误差椭圆、样品选择和模型方向放在一起，而不是只看相关系数高不高。

此外，封闭性具有时间范围。样品在最近阶段没有明显交换，不足以证明整个早期历史都保持封闭；相反，一次足够早的共同重置可能使时钟记录重置以后而非最初形成以来的时间。事件身份仍然跟随记录机制。将这种可能性提出，并不等于否定每一项年龄，而是要求不同材料、不同体系和地质背景共同决定时钟在何时开始具有所需意义。

◆ 14.3 1956 年的承重论证

1956 年《陨石与地球的年龄》把这一问题推进为更完整的论证。原论文比较不同陨石的铅组成，并讨论其他放射性计时方法的关系，给出约 45.5 亿年的年龄估计；它同时明确讨论化学分异、共同初始条件和从陨石向地球迁移结论的限制。[R14a] 这里最值得学习的不是把结尾数字背下来，而是看到作者没有把陨石年龄与地球年龄当成无需说明的同义词。

本书为此将证明责任拆成三层。第一层是分析层：各样品的同位素组成是否可靠，误差和空白是否足以改变分布？第二层是演化层：样品间的关系是否适合共同起始、不同铀铅比例和后续封闭演化的模型？第三层是迁移层：这组陨石所约束的时间，为什么可以代表地球形成时代？前两层即使非常成功，也不能代替第三层。把某个外部样品测得很精确，不会自动使它成为地球的代表。

为了理解第二层，可以继续前节的教学式。若两条衰变链的衰变常数分别为 λ_1 与 λ_2，母体同位素比例受到约束，那么铅比值图中的斜率与两种指数增长项之比有关。年龄由斜率反求，意味着不同材料之间的协变结构成为时钟。读者不必在这里记住具体常数，但应看见一个重要区别：样品点的离散不全是误差，沿着模型方向的变化提供识别力，偏离模型方向的变化则可能暴露误差、混合或不共同的历史。

横纵轴共享^{204}Pb，误差分析便须保留两者的关联。参照同位素读数若偏高，两个比值可能一同偏低；点的误差范围因而具有方向，不能只由各轴独立的误差棒概括。某类样品的共同污染还可能整体移动一组点。拟合时需要判断这些偏移方向与模型斜率是否接近，否则狭长点群可能给出过于乐观的年龄精度。这是对关系图的现代分析，原论文的具体处理仍应按其自身方法评价。

第三层涉及地球物质史。地球经历分异，地表铅并非原始整体铅的透明窗口。若某样品富集或亏损铀，其后续铅组成会沿着不同路径变化；若它与深部储库交换，单一阶段模型也可能不够。陨石研究提供的是一条有关早期共同条件的比较线索，地球样品需要在可解释的演化模型中与之相接。能够相接的范围，决定“地球年龄”这个说法有多强，而不是反过来由题名决定材料一定支持全部含义。

一个反事实可以帮助看清边界。假如各陨石点不能由共同演化关系解释，而显示多个明显不同的初始铅组成，那么直接拟合一个共同年龄就失去根据。研究者可能仍可测定个别样品年龄，但应缩小整体结论。再假如陨石关系稳定，而地球材料无法在合理模型中与之连接，那么结果首先是陨石群的年龄，不应仅凭数值符合期待就升级为地球年龄。这种反事实不是否定研究成功，而是说明成功实际跨越了哪些门槛。

还应警惕“多方法一致”的独立性问题。不同衰变体系如果共享某项校准、某种形成假设，或者一种方法被用于调整另一种方法的常数，它们的一致性就不能全部算成相互独立的重复证明。这并不使比较无用；它要求把依赖关系写清楚。独立性是证据网络中的结构性质，不是论文列出几种方法就自动拥有的数量。若一套方法在共同条件之外仍能提供不同失效机制的检验，其支持才更有分量。

这里可从 E→D→S 复审推断：共同初始条件与保存条件，限定哪些材料差异具有时间意义；这些差异构成序列，进而使行星形成时代成为可辨认对象。也可从目标 S 反向追问，为什么需要选择能够约束初始铅的外部材料。这两种入口分别突出证据资格与采样需求，合起来才能检查年代结论是否依靠了未经说明的迁移。它们是本书的分析次序，帕特森实际研究步骤则须由相应历史记录重建。

新的研究世界由这些联系构成：地球岩石提供后续演化记录，陨石保存可比较的早期信息，化学分离与同位素量测辨认组成，形成模型规定何种差异可以计时。样品由此获得不同的证据角色，研究者也获得选择、排错和比较的新能力。初始条件、后期分异与行星

形成时代得以分开提问。发生学分析若能指出其中某条联系怎样建立或失效，便能说明这一能力的历史结构；笼统地说“认识更新”，则会遗漏实际的材料与操作变化。

将关系写成教学计算链。取样品 i 现今的 $u_i=^{238}U/^{204}Pb$，并设各样品现今的 $r=^{235}U/^{238}U$ 相同且已知。在两条衰变链适用、^{204}Pb 稳定、母体与子体封闭保存的条件下，令 $A(t)=\exp(\lambda_{238}t)-1$，$B(t)=r[\exp(\lambda_{235}t)-1]$，便有 $x_i=x_0+u_iA(t)$，$y_i=y_0+u_iB(t)$。对 $t>0$ 消去 u_i，得到 $y_i=y_0+[B(t)/A(t)](x_i-x_0)$。A、B、$u_i$ 及斜率都无量纲。共同初始点与共同历时规定线的位置和方向，母体相对含量的差异使各点沿线展开；用斜率求 t，仍以共同条件成立为前提。这是教学推导，不是原论文计算表的重制。

现在分别改变一个条件。若不同样品有不同 x_0 与 y_0，点群不再必然落在同一条线上；若某样品后期加入外来铅，它的移动方向由外来成分决定；若两种端元混合，也可能生成线性趋势，但线的意义属于混合比例而非共同年龄。因此，图形相同可能对应不同生成过程。年龄解释需要把图形放回矿物、化学与形成史中，这正是不能从回归成功直接跳到历史结论的原因。

这组关系还说明，地球年龄研究中的“外部”并非空间上随意找来的外物。真正有用的外部材料必须在某些特征上与目标共同，在另一些特征上形成可辨差异。共同性提供迁移的基础，差异提供识别演化的杠杆。若只有共同而没有差异，比较缺乏辨识度；若只有差异而无共同条件，比较又难以取得统一时间意义。行星年代学因此把相似与不同组织成同一个推断结构。

这种结构也使研究结果能够引导下一步采样。如果某类材料偏离共同关系，最有意义的后续工作可能是调查其矿物相和交换史，而不是立即把它删除；若偏离与某种独立特征相连，它甚至可能提供另一阶段的时间信息。异常样品既可能是分析问题，也可能保存真实的复杂历史。如何区分，取决于增加的证据是否改变了它的事件身份。

在行星层级，样品之间的空间距离也不直接等于历史距离。两颗相距很远的陨石可以保存相近的早期形成条件，两块今天相邻的地球岩石却可能经历完全不同的结晶和变质事件。取样的代表性不是地图上分布均匀就能保证，而要看样品在物质演化网络中处于什么位置。这种区别使年代学从简单的地点目录转向材料历史的比较。

还可以把地球形成看作质量随时间增长的过程。若用达到某一质量比例作为定义，所得时点与最初固体结晶时点通常不同；若用某套同位素系统停止显著交换作为定义，又取决于系统对过程的响应。这样的定义必须服务于明确科学问题。例如讨论早期热演化，需要知道能量输入与分异阶段；讨论地壳记录，则关心晶体形成和保存。一个通用年龄提供时代坐标，专项研究再把坐标分解为过程。

因此，帕特森研究的成功不应被理解为从此不必解释“地球”一词的历史含义。它使行星年龄成为更强的可测对象，同时打开形成过程内部的细分空间。数字的稳定与问题的增长可以并存，甚至前者正是后者的条件：有了可靠的大尺度时间框架，研究者才有理由追问较小的先后差别，而不必每次重新争论全部历史有多长。

这种成功也显示，外部比较与内部证据需要循环接续。陨石提出早期时间约束，地球材料检验自身演化能否与之相容；不相容之处再要求改进地球物质史或重新评估材料共同性。结论并非从天外一次性输入，而是在多类记录的相互限制中逐步稳定。

材料的代表性还需要与问题共同更新。适合限制共同初始组成的样品，不一定适合记录后来的撞击；适合记录撞击的矿物，又可能失去最初结晶信息。采样策略因待估事件而改变，不能让一套成功材料永久承担所有年代问题。

◆ 14.4 结论强度与未决对象

一个稳定的地球年龄数字并不意味着形成过程已经被压缩成瞬间。可以把它理解为对早期行星形成时代的强约束，同时继续追问各个阶段的先后与历时。太阳系早期固体形成、行星体聚积、地核分离和地壳结晶，涉及不同材料与过程。若新的高分辨率方法把原来不能区分的阶段分开，这是对年龄对象的细化，不必被写成旧结论整体崩溃。

这与测量精度提升有微妙区别。把同一事件的误差范围缩小，是原有对象的更精确认识；发现原来混在一起的两个事件需要分别定年，则改变了待估量的结构。两者都重要，却不应混为同一种创新。一个广为使用的年龄能够长期保持实用价值，同时允许研究者不断重审其中“形成”的具体含义。科学知识的稳定性不要求语义永远停止发展，而要求每次细分都能说明与此前结果的关系。

对这类年龄应同时报告时间值与其条件层级。共同初始组成可由材料比较限制，封闭性可由保存史审查，陨石与地球的迁移还要由形成模型承担。某一条件需要修订时，影响未必扩展到其他层：个别样品失效不必推翻全部关系，迁移模型调整也不等于同位素读数失效。分层说明条件，使一个长期使用的年龄既能保有可靠部分，又能接受新的地球形成问题。

本章也给出一种研究设计上的原则。面对整体对象，最直接取样不一定最有解释力。地球岩石离我们最近，却可能在关键初始条件上保存较差；陨石来自外部，却可能提供更有效的比较。选择材料的依据不是空间亲近，而是它在推断链中承担什么功能。这一原则不能无限推广：外部材料若缺少可验证的共同性，就只能提供类比，不能提供年龄迁移。真正有力的改道，同时包含扩大取样范围与收紧迁移条件。

从本章回看冷却争论，可以看到两种时钟的差别。冷却模型主要从当前热状态反推热史，放射性系统则利用材料内部积累的核素变化，但两者都不是脱离历史假设的自动读数。新时钟的强大，不是因为它取消模型，而是因为它提供了新的可测关系、不同的失效机制以及更多交叉检查机会。成熟的年代学不是从有前提走向无前提，而是从难以约束的前提走向能够分别检验的前提。

最后，行星年龄并不能单独组织全部地层历史。知道地球何时形成，还需要知道哪些界线供不同研究共同使用，地方剖面如何对应全球时间，数值年龄更新时名称与边界是否随之移动。下一章转向金钉子与年代标尺，讨论一个地理位置如何承担公共坐标的责任。时间知识从材料内部的时钟扩展到跨地区的共同标准，并不减少自然证据的重要性，而是增加了相关与修订的公共规则。

第 15 章 金钉子与年代标尺
地方剖面怎样承担全球责任

15.1 公共坐标解决什么困难

设想两组研究者在不同地区工作，都报告发现了二叠纪与三叠纪之间的转折。第一组依据岩性变化划线，第二组依据某类化石的首次出现。两份报告可能都观察认真，却仍不能直接比较，因为“界线”在两份报告中可能不是同一对象。若不先解决这一点，更高的测量精度只会使不同对象的描述越来越精细。

公共标准首先要解决指向一致的问题。它使来自不同地点、使用不同材料的研究者知道，彼此试图相关到哪个共同位置。共同位置建立以后，相关是否正确仍然可以争论。标准没有消灭争论，而是减少争论中对象偷换的机会。

这里至少要区分三个字段：界线在哪里定义，以什么标志帮助识别，数值年龄估计是多少。它们经常被压缩在同一张年代表的一格里，所以看起来像一件事。实际上，它们各有证据责任：定位要落到剖面，识别要落到标志与比较，定年要落到样品和时间模型。

一个界线点的定义可以保持，而对其数值年龄的估计发生更新。这个逻辑并不奇怪。某地点的位置可以由一套基准规定，对其到另一地点的距离仍可以不断改进测量。测量更新不意味着地点重新诞生，但会改变可以利用这个地点完成的比较。地层标准还具有自己的自然历史与相关困难，类比不能代替专业论证，却能帮助我们避免把定义与估计混为一谈。

共同标准也不是越复杂越好。若一种标准需要只有极少数实验室才能识别的特殊材料，它可能具有局部精度，却难以传播；若标准过于粗糙，又可能把不同事件压在一起。标准设计承担一种张力：需要足够明确，使不同团队能互相纠错；又需要足够可用，使标准不只活在少数标本中。

从本书的立场看，公共时标构成一种新的互动能力。研究者不必共享全部经验，仍可以沿同一标准交接问题。甲在某地建立的相对关系，乙可以检验它是否与另一剖面相容；

新的不一致也可以反过来暴露原先标志应用的边界。这个共同世界不是把资料放进同一个文件夹，而是资料之间获得了受到约束的可比较性。

因此，制度性与客观性不能被摆成只能二选一的两端。批准程序不能决定古代生物何时出现；古代生物留下的痕迹，也不会自己起草研究者应该怎样引用和修订标准的规则。一个可用时标同时需要自然材料的约束与共同操作的组织。前者防止任意，后者使不同地点的工作能够真正相接。

◆ 15.2 煤山 27c 底的公共界线

国际地层委员会登记的印杜阶底界位于浙江煤山 D 剖面 27c 层底，主要相关标志为牙形石 Hindeodus parvus 的首次出现，批准年份为 2001 年。这些信息分别定位了地点、层位、标志与制度时点。2001 年不是岩层形成的时间，批准也不是那类生物首次出现的自然事件。[A12]

这个案例首先要求我们抵制视觉压缩。把界线画成一根细线，十分有利于阅读，却容易让人误以为自然历史在任何地点都留下了同样清楚、同样连续的一根线。图上的线是标准表达；不同剖面上能否找到对应位置，要另外建立证据。

同样，灭绝与年代界线也不自动重合。灭绝研究要讨论特定生物记录如何减少或消失，涉及采样、保存和事件划分；地层界线要提供共同的时间定位。二者可以互相约束，却不能用“这是大灭绝附近”替代精确的界线定义。第 13 章讨论的样品年龄，则为这些对象提供时间限制，而不是取消它们之间的区别。

这些对象的时间关系应分别表达。颗粒结晶年龄通过保存与来源判断约束灰层沉积，灰层再通过层位关系约束生物记录；标准界线则需要说明与哪些记录相接。两处数值相近，可以来自测量分辨率有限，也可以来自过程确实接近，须由上述连接加以区分。把数值放回关系链，才能知道一个时间究竟是在定义位置、估计事件，还是检验二者的相关。

阅读剖面对照图时，可以把层位记录、相关推断和标准定义分别标明。记录说明某标本取自何处，相关推断说明为何将它与另一剖面相连，标准定义说明比较最终指向哪个位置。若其中一项变化，另外两项未必随之改变。例如标本重新鉴定可能使某条相关线失

效，却未触动标准剖面的定位。这样的读法有助于追踪论证，并不构成对煤山原始资料的独立复核。

在 SDE 语境中，界线作为 S 显露，并非一块岩石孤独携带的标签。D 包含不同样本、层位和剖面之间的辨认次序，E 包含化石鉴定与保存条件、沉积连续性与比较任务的相互规定。某个化石标本能否用于识别界线，取决于它被怎样鉴定及所在记录是否可靠；被赋予的相关作用，又会规定下一步需要检查哪些形态和层位细节。

不同证据在界线识别中有主次之分。主要标志提供首要相关依据，其他材料检查其鉴定、保存与区域适用性。若某剖面缺少主要标志，辅助资料可以缩小可能区间，却未必达到同样定位精度；若主要标志与独立层序发生冲突，又需要审查它是否被误认或重搬运。E 的具体内容就在这种分工与制约中，而非把所有因素赋予相同权重。

煤山因而不是凭借一个地名就成为全球时间的缩影。它的标准地位来自可定位的定义及可执行的比较安排。自然材料使这种安排受到限制，公共采用使它能够跨研究传递。一个地方承担全球责任，靠的不是地方包含所有历史，而是地方能够为分散研究提供共同起点。

◆ 15.3 怎样搬到别的剖面

标准无法像石块一样搬到别处，能够传播的是识别与相关的工作。另一剖面发现了同名化石，还不能省略对标本身份、保存条件和出现位置的检查。尤其是“第一次采到”与“当地第一次存在”，“当地第一次存在”与“全球第一次出现”，是三种不同陈述。

以下是一个明确的假想案例。甲剖面连续取样，乙剖面取样较疏。某化石在乙剖面较高位置才被发现。若直接把这个高度差解释成地质事件不同时发生，就忽略了采样差异；若反过来强迫两个出现位置绝对同时，又可能遮蔽真实的区域差异。需要增加取样、检查保存以及寻找其他可比较线索，才能判断哪种解释更有支持。

这个例子还揭示“没有”的证据问题。样品中没有某化石，可能因为它当时不在这里，也可能因为未保存、未采到或被误认。每一种解释对下一步工作的要求不同。扩大样

品量能处理部分漏采风险，却未必修复彻底缺失的沉积记录；重新鉴定能修复分类问题，却不能制造不存在的标本。

共同界线还改变下一轮工作的重点。以它为目标，研究者会优先寻找有相关判别力的层段和标本，并检查同名化石的出现是否受保存或环境限制。这可用 D＝G(S,E)描述：界线的显露 S 与现有特征约束 E 共同选择比较路径 D。2001 年的批准是否改变过某团队的具体路线，则属于另一个历史问题，须比较批准前后的采样计划和研究记录。方法上的引导与实际发生的行动，在这里各有证据。

标准还可能使原先被忽略的资料重新变得重要。一个旧标本若缺少准确层位，即使鉴定无误，也难以承担精细相关任务。研究者可能因此改进采集标签，记录更细的位置和保存状态。若这种变化有前后版本可查，便构成 E＝H(S,D)的一种具体候选：新的共同目标与实际工作改变了特征能够互相限定的条件。

记录改善首先增加复核能力。准确标签使后来者能够回到原层位，区分重新采样所得差异与原来定位不清造成的差异。只有这些复查进一步改变了相关关系或事件判断，才增加新的地史内容。因而标准实施可以有几种不同成果：定位更稳定、复核更便利，或原有解释被修订。将它们分别说明，比用“标准促进进步”概括全部变化更能看清实际贡献。

跨区域相关的科学性，最后体现在它能否接受不一致的材料。若一种标准应用方式要求把每个反例都解释成样品错误，且没有允许这种解释失败的条件，它便失去纠错能力。有效的标准既能稳定比较，也能明确指出某些剖面暂时不能可靠相关。承认暂不能定位，不是公共标准的崩溃，而是保住标准的证据含义。

公共比较还需要处理分辨率不对称。甲剖面保存细密、取样充分，乙剖面只能提供较宽的时间区间。若把甲的高分辨率结果直接赋给乙的某一点，便把相关能力写得超过实际证据。更稳妥的表达可能是一个区间，而不是一条看起来精确的线。标准要求共同定位，不要求每份记录都有相同清晰度。

这种不对称在教育中尤其容易消失。整齐的年表和剖面对照图倾向于把各地界线画在同一水平线上。图示可以表达一种相关判断，但读者还应知道哪些连接受到直接标志支持，哪些依赖间接约束，哪些仍有较大范围。教学上的整齐应当服务理解，不能用来替代研究中的不齐整。

由此还可以区分“标准是否有效”与“某次标准应用是否有效”。一项应用失败，不必导致全部标准失效；但若多项应用持续暴露同类困难，也不能只责怪使用者。需要问问题出在样品、地区条件、标志选择，还是标准承担了超过其能力的任务。责任可以被定位，才可能发生有针对性的修订。

这种分析也使共同体的作用更具体。共同体不是一个可以凭意见创造自然事实的总主体，而是通过可共享的记录、比较规则和批评程序，使不同研究者能够彼此校正的组织过程。某项结论被很多人重复引用，只证明它流通广；当后来的人能够沿证据重新检查并有条件地修订它时，共同性才取得更强的科学含义。

标准的实施还有一个具体检验：两个研究者是否能够依据公开说明回到同一层位，并区分定位分歧与年龄估计分歧。只有这种差异可被分开记录，标准才真正提高协作的可复核性。

◆ 15.4 标准保留与修订

面对新材料，研究共同体可以作出不同层级的修订。第一种是保持界线定义，改进数值年龄估计。第二种是保持标准地点，调整某些辅助标志的使用。第三种是发现某项相关方法在特定地区不适用。还有更强的情况：定义或标准本身需要经过正式程序重新审查。这些动作不能都用一句“年代表改变了”概括。

修订层级不同，证据门槛也不同。更精确的测年可以支持更新年龄估计，却不当然要求移动界线；异地某化石出现偏晚，可能限制它在当地的相关用途，却不当然推翻标准地点的定义。若不区分这些层级，就会把局部问题放大成全面危机，或把真正触及定义的问题缩小成小数位调整。

我们可以再提出一个假想的版本比较。第一版年表把某界线年龄写成 t_1，第二版写成 t_2。如果两版都指向同一剖面位置，变化主要属于估计层；若层位位置也改了，则必须追问对象身份是否同步改变。只比较两个数字的差值，无法识别这两种更新。版本记录应同时保留定义、标志、数值、方法与理由。

本章对“新世界”的主张也因此受到约束。一个标准建立后，若不同研究确实获得此前缺少的共同定位、可迁移操作及纠错渠道，我们可以讨论公共时间世界的发生。若只统

一了称呼，而实际比较仍不可执行，便不能把命名的成功当成研究能力的生成。若旧机制早已完成相同任务，新标准的价值可能在于维护与便利，而非本体层面的创新。

用 P1—P3 复审标准变化，应追踪张力究竟落在哪里。若差异序列与相关条件不能支持既定界线身份，问题指向 S 的重定；若界线身份与条件保留，现有采样却不足以完成比较，则需重组 D；若新结果与既有比较持续冲突，保存和相关条件 E 可能需要调整。这些分别对应 D/E、S/E、S/D 之间推动另一项变化的候选相位。一次批准并不必然包含全部三相，具体判断应由前后版本和操作记录支持。

标准地层学已经有成熟的定义、相关、保存与修订方法。发生学重述应把这些具体方法与共同体的辨识能力联系起来：一项规定怎样使异地研究者找到同一位置，一次失配怎样促成新采样，一组新材料又怎样改变标志的适用条件。只有这些连接在历史记录中可追踪，跨时代比较才获得内容；仅把既有做法换成 SDE 词语，仍只是另一种概括。

一枚金钉子的象征力量，容易遮住背后漫长而细密的工作。真正使它成为知识基础的，是研究者可以回到具体剖面、具体标本与具体定义，检查某次跨地域连接是否成立。标准最深的价值，不在于禁止后来的人重问，而在于使后来的人知道应该从哪里重问。

至此，样品计时、行星年龄与公共时标已经取得不同而相连的意义。接下来转向空间：山地如何被分解为构造与侵蚀对象，地图如何借助投影、尺度和调查形成可比较的地表世界。时间与空间的表达共同稳定以后，我们才进一步走向海洋，追问海底信号怎样取得运动的历史意义。

第 4 编

地表测绘与空间尺度

第 16 章 山地不再只是风景：吉尔伯特怎样分解地形

第 17 章 地图投影：一种空间为何需要多种合法表达

第 18 章 尺度怎样改变可问的问题，而不改变山河本身

第 19 章 地貌循环与动态平衡：两种解释怎样接受同一地形检验

第 20 章 丁文江的下扬子调查：地方文献怎样成为海岸变迁证据

第 16 章 山地不再只是风景
吉尔伯特怎样分解地形

16.1 一幅图中的两个山体

吉尔伯特 1877 年《亨利山地质报告》的卷首图，把同一座山拆成两半：远半部表示只有抬升造成的形态，近半部表示抬升与侵蚀共同作用后所见的形态。[A13] 图中并不是并排放着两座现实山峰。一半是依据结构关系重建的模型，另一半接近当前观察对象。它的力量来自这种不对称：研究者用一个未被直接看见的状态，与一个能够在野外检验的状态相比较。

若只欣赏山的轮廓，两半都是图像；若追问山怎样形成，它们承担的责任便不同。现状部分需要与实际山体的谷地、岩层露头和地势关系相符；重建部分需要解释为什么认为某些岩层曾经连续、为什么推断它们经历拱起，以及哪些后来被侵蚀的部分能够恢复。图画得逼真不能代替这些证据。一幅图可以把推理表达得很清楚，也可以把未经检验的推理伪装成直接观察。

这幅图使“山的高度”不再是一个足够的解释对象。现在山顶在哪里，与构造抬升曾经把地层抬到哪里，并非同一问题。侵蚀既降低地表，又暴露内部结构；研究者因此能够看见原本被覆盖的岩体，但也失去了部分原始形态。证据获得与证据毁损发生在同一自然过程中。地貌思想由此不能只问山怎样被造出来，还必须问如今能看到的山，经过怎样的选择性保存。

报告将构造讨论与地表雕刻分章组织，这种安排不是简单的写作分类。它要求不同解释各自承担可检验工作：构造模型解释层位怎样被移动，侵蚀模型解释哪些部分为何消失或保留。两者相互关联，却不能互相包办。若把所有形态都归于抬升，就无法说明岩性不同造成的保存差别；若把所有形态都归于侵蚀，又不能解释岩层本身为何倾斜和拱起。

研究对象在这种区分中获得了内部结构。位移、层序、接触、抗蚀性与排水关系，使同一座山可以分别作为构造与侵蚀问题加以比较。复原拱起需要追踪同层位置，分析残余

山形需要判断哪些部分易被移走；两项任务又借助同一批露头相互限制。新的科学身份因此表现为提问、测量与核验方式的分化：原本只可整体描绘的轮廓，现在能够逐部分追究成因。

16.2 从形容词到可比特征

“陡峭”“破碎”“雄伟”能够准确传达观看经验，却不足以区分成因。同样陡峭的坡面，可以由坚硬岩层差异侵蚀形成，也可以与断层位移或快速切割有关。要进入解释，观察需要转成能够比较的特征：层面倾角、地层连续性、不同岩石的接触关系、谷地位置及露头高程。转换并非取消人的感受，而是为不同调查者提供能够相互核验的接口。

其中最关键的一步，是区分地形坡度与岩层倾角。一条陡坡上可以露出近水平岩层，倾斜地层也可能被削成较平缓地面。若把两者混同，视觉上最突出的形态就会支配地下结构解释。调查者必须寻找层面，确定其走向与倾斜，追踪它在相邻地点是否仍为同一层。这样，一个看似简单的山坡才分解为地表形状与岩层姿态两个对象。

吉尔伯特在报告中也明确保留了地层辨认的困难：一些砂岩在岩性上近似，失去层序位置后难以可靠区分。[A13] 这是一条重要的方法线索。颜色和硬度不是随处同样有效的标签，它们与上下层关系、化石和变质状态共同限定层位身份。某块岩石被辨认为哪一层，影响构造连线；构造连线又提示下一处应该寻找什么层位。这个往返有生产力，也有循环论证风险，所以需要把实际见到的露头与推测连接分开记录。

侵蚀使某些边界清楚显露，同时让其他边界中断。研究者不能因为两侧露头相似，就无限延伸一条看不见的线；也不能因为中间没有露头，就断言地层必然断开。缺失信息需要用不同线型、可靠性说明或文字限定表达。地图和剖面图的诚实，首先体现为它不把推测与观测涂成同一种确定程度。

在 SDE 分析中，D 不是“不同山峰”这样宽泛的集合，而是具有问题指向的差异序列：同一层沿不同方向的高度、倾角和连续性如何变化。E 则是层位身份与几何关系之间的相互规定。没有可靠层位，倾角变化可能是在比较不同岩层；没有空间追踪，岩性相似也不能确定同一性。研究者辨认能力随着这套关系发展，山体的构造特征也在可检验的互动中取得科学身份。

这些特征须有区分成因的能力。若两种模型都容许同样的层面姿态和接触位置，继续记录同类形态只会加密描述；应寻找两者预期不同的位置。若测量误差已足以覆盖所谓倾角变化，则需先提高测量或层位辨认能力。构造解释的复杂程度，由可辨差异所限制，而不是由图上能画出多少条线决定。

◆ 16.3 构造解释需要什么

亨利山报告的一个中心问题，是地下侵入体与上覆地层拱起的联系。若熔体进入层间并使上覆岩层隆起，就应出现与这一过程相容的空间关系：侵入体、围岩接触和地层变形不能彼此任意配置。研究者需要让地表露头组成能够约束地下形态的证据网，而不能仅凭山呈圆顶形就宣称地下必有某种岩体。

吉尔伯特讨论侵入体时，不只列举形态，也分析密度与穿透阻力等可能影响熔体运动的条件，并对哪些条件具有决定作用提出可比较的预期。[A13] 我们不必将其中每一项十九世纪机制解释都视为现代结论，才能承认这种论证的重要性。关键是作者把一个总原因拆成可询问的机制：物质为何向上或沿层面扩展，上覆岩层如何抵抗变形，不同层位为何可能产生不同结果。

若某解释强调特定岩层构成阻碍，就应检查侵入位置是否集中在相应层位；若强调熔体与围岩密度关系，就应比较材料性质与侵入高度是否相符。两者可能同时起作用，单一现象也可能不足以判别。思想史需要保留这种不确定性，不把最后被教科书记住的术语当成原报告所有推理已获确认的证明。

观察与模型之间还有一个尺度转换。单处接触能够说明局部岩体关系，却未必代表整个地下体；一张区域轮廓图能够表达整体形态，却可能掩盖局部不相容之处。可靠构造解释需要在局部与整体之间往返：整体模型提出应当出现的接触，局部观察则决定模型哪些部分需要修正。这一过程不是把零散事实堆成全景，而是让全景持续承担被局部事实打断的风险。

可以构造一个反例来检验。如果山体表面呈拱形，但地层并未向外倾斜，所谓拱起解释就需要另找证据；如果侵入体存在，但其时代晚于某项主要变形，也不能把所有变形都

归因于它；如果侵蚀差异足以产生同样地形，现今轮廓就不再是独立的构造证明。这些反例说明，存在某种岩体、存在某种形态、二者具有因果关系，是三个不同层次的判断。

空间模型可以进一步转化为调查路线。如果两种地下连接在已知露头处都成立，却在山体另一侧给出不同层位预期，那里便成为更有判别力的补测位置。路线安排、交通和经费影响能否到达该处，地质结论则取决于所见接触是否支持预期。操作条件与地质证据在同一调查中相接，却承担不同责任：前者使比较可执行，后者决定哪种解释应保留。

报告的构造分析还有一个值得保留的顺序：先识别可见关系，再讨论造成关系的力学原因。若把力学上可能的过程直接当成这里实际发生的过程，就会遗漏历史科学最关键的一步。熔体能够使覆盖岩层弯曲，不代表任何拱形岩层都由熔体造成；只有具体接触、岩体分布和相对时代相互支持，机制可能性才进入地方历史。反过来，地层几何能够强烈限制位移，却未必足以在几种力学机制之间唯一选择。

侵蚀的作用也具有选择性。坚硬岩体可能在周围较软地层被移走后突出，但一个突出体的现今轮廓不必等于它最初的地下边界。裸露部分只是整个侵入体被切开的窗口。若据此恢复地下体积，需要考虑尚未暴露部分与已经消失部分，不能把现今地形直接当作原始岩体的完整模具。图像重建因而是一种有约束的补全，而不是将眼前山形简单向地下延伸。

这类区分对于后来的地质剖面极有生命力。剖面把看不见的地下连接表达出来，使不同观察点能够组成假说；同时，每一段地下线都应能追溯到露头、层序或其他约束。它既扩大观察能力，也扩大推断范围。好的剖面不是推测最少的剖面，而是推测身份清楚、能够指出下一处关键观测位置的剖面。吉尔伯特的图像组织由此成为一种研究工具，而不只是报告的装饰。

岩层被侵蚀切开的窗口还具有方向性。同一构造从不同方位观看，露头形态可能相差很大；只沿一条谷地调查，容易把剖面特征误当整个山体形状。多方向追踪能够检验三维解释，使立体图中的连接不只依靠单一视线。

16.4 图像产生的新问题

理想抬升图与实际山体之间的差异，不应被当作画家必须抹平的缺陷。它恰好形成后续研究议程。某一处预期岩层没有出现，可能意味着侵蚀更深，也可能意味着原先的层位连接有误；某条谷地偏离模型预期，可能与岩性、排水历史或后期构造有关。残差只有在模型足够明确时才显露，因此模型的贡献不只在解释已经匹配的部分，还在使不匹配部分获得清楚身份。

不同形态应引出不同的补测任务。系统拱起的层面，要求沿同层追踪倾角，并比较侵入接触与变形时代；坚硬层形成的突出山脊，要求检查抗蚀差异能否解释其坡位和保留范围；穿过预期构造高部的谷地，则要求重建排水路径与构造变化的先后。三种任务分别检验位移、选择性保存和河网继承，把图中的不匹配转为可执行的观察，而非为每种轮廓预先指定唯一原因。

图像还有传播功能。不同读者可以在同一幅图上指出不同缺口，野外调查者也可以据图选择补测地点。但传播得广不等于解释得到证实。图中的模型成分越容易被遗忘，传播反而越可能把推断固化为“人人看见的事实”。因此，模型图的图注应保留它表示哪一阶段、哪些部分为重建、哪些关系具有直接依据。清楚的视觉表达与清楚的证据分层应该同时存在。

从三方程的复审角度，S 是能够分别检验的构造与侵蚀对象，D 是重建形态与实际形态之间有方向的差异，E 是层序、岩性、位移与暴露的相互规定。新露头若改变同层连接，重建的拱起范围随之改变；拱起范围改变后，原先归于侵蚀的地形差异也可能需要重新分配。这样的连锁修订才是条件回写的具体内容，图件前后版本及其野外依据应能显示每一步改动。

山地从风景进入科学，并未失去整体性。相反，构造与侵蚀的区分使整体拥有更丰富的内部关系。一座山既有被抬起的历史，也有被削去的历史，还有研究者能够辨认二者的条件。接下来还有一个看似更基础的问题：这些地点如何进入共同空间？同一地球表面在平面图上必然经过变换，地图投影将使我们看到，科学表达的差异也能够受到严格的几何约束。

第 17 章 地图投影
一种空间为何需要多种合法表达

◆ 17.1 图上相同不等于地上相同

两幅世界地图上，同一片陆地可以显得面积悬殊；一条在某图上笔直的航线，在另一图上可能弯曲。若没有说明投影规则，这些差别容易被解释为制图者随意摆弄地球。但地图的根本困难首先是几何的：弯曲表面不能在保持所有内在长度与角度关系的同时，完整铺展成平面。选择一种投影，就是选择哪些性质优先保留、哪些失真接受控制。

用球体近似下的两个教学公式，可以直接检查变形。设经度为 λ、纬度为 ϕ、中央经线为 λ_0，球半径为 R；图面 x 向东、y 向北。取赤道为标准纬线时，等距圆柱投影可写为 $x=R(\lambda-\lambda_0)$、$y=R\phi$；球面墨卡托投影则为 $x=R(\lambda-\lambda_0)$、$y=R\ln\tan(\pi/4+\phi/2)$。角度均用弧度，且 $-\pi/2<\phi<\pi/2$。以下比较是公式的解析推演，其距离和面积结论限定于这一球面模型。

在纬度 ϕ 处，球面微小东西距离为 $R\cos\phi\,d\lambda$，南北距离为 $R\,d\phi$。等距圆柱图的对应长度为 $R\,d\lambda$ 与 $R\,d\phi$，故东西向比例因子是 $1/\cos\phi$，南北向因子为 1；两者在高纬不再相等。墨卡托图有 $dy/d\phi=R/\cos\phi$，使两个方向的比例因子同为 $1/\cos\phi$，局部角度因此保持，面积比例因子则为 $1/\cos^2\phi$。在北纬或南纬六十度，$\cos\phi=1/2$，所以局部长度放大两倍、面积放大四倍。这些是局部变形，不能直接用同一倍数乘遍跨纬度的整条路线。

这个例子已经足以反驳“图上面积相同就代表地上面积相同”。它也说明失真不能只靠视觉感受判断。某图看起来舒适，可能因为读者习惯了它；某图看起来陌生，却可能更适合面积统计。地图评价需要回到任务所要求的量测关系。若比较的是地区面积，应检查面积保存；若判断局部方向，应检查角度；若估计某条路线长度，则需要确认该方向和位置的距离变形。

同样，等距投影不意味着任意两点距离都正确。通常只有特定方向、线或相对于指定中心的距离得到保存。把名称中的“等距”理解为全图万能尺，会把一个局部性质扩大成不存在的全局承诺。投影术语必须与具体公式及有效区域一起阅读，不能把名字当作已经完成的几何检验。

地图空间由位置对应关系与失真结构共同定义。地面上的同一条河流，在不同图面中仍可通过坐标转换识别；但沿线量长度、统计邻域面积或判断方向，需要采用各自适用的规则。因此，换投影可以改变图面上可执行的操作，却不改变山河本身。这种可转换而操作条件不同的关系，使多种地图能够共同服务于同一地理研究。

◆ 17.2 用途如何参与定义

斯奈德 1987 年的《地图投影工作手册》系统组织不同投影、使用条件以及正反变换公式。[A15] 它代表一种成熟的知识形式：研究者不必在每次制图时重新发明所有方法，但必须能够根据任务选择并检查方法。手册不是一份最佳投影排行榜，因为没有脱离用途的唯一最佳投影。

设想两个团队使用同一批地理位置。一个团队比较不同区域的土地覆盖面积，另一个团队需要表达航向。前者若直接在强烈面积失真的图面上数像素，就可能使高纬区域被过度放大；后者若只选择等面积而不检查方向性质，也可能误解图上直线的航向含义。这里的差别不是两个团队拥有不同地球，而是他们需要从同一空间中保留不同的操作关系。

地图任务参与定义对象，首先表现为“相等”的含义不同。面积研究要求面积相等在变换后仍可比较；角度任务要求局部相交角度得到保存；区域制图可能把有限范围内的总变形控制作为首要目标。这些目标受到同一几何条件制约，并非价值偏好可以随意满足。选择只在可行集合内发生，不能通过声明重要性取消不可能性。

需要进一步区分形状与角度。等角性质是局部的，意味着足够小范围内的方向关系保持；一个横跨巨大区域的大陆轮廓仍可能显著改变。把“保角”说成“大陆形状完全不变”，就把微分尺度性质误扩大到有限区域。类似地，等面积不意味着轮廓不变，也不意味着所有长度都能直接量取。每一种保真都需要同时声明它不保证什么。

图幅范围也会改变选择。同一种投影用于适宜的区域可以有较小变形，扩展到远离设计中心的地区则可能变形明显。于是地图对象不仅附带投影名，还应附带中央经线、标准纬线或比例因子等参数。两个文件使用同一投影类别，却采用不同参数，不一定可以直接叠加。名称相同不能替代数值定义相同。

在 SDE 框架下，E 是任务所需的量测性质与几何可行性之间的相互规定。任务要求保面积，便限制可接受的图面变换；几何又反过来限制该地图能够支持哪些其他任务。S 是具有明确保真目标的地图空间，D 则是不同投影下可比较的长度、角度和面积差异。这个解释的价值在于把“地图有立场”转成可核验问题：具体保留什么，牺牲什么，错误使用会产生何种偏差。

投影变化是否改变研究能力，可以由量测流程判断。若只更换屏幕显示，而面积始终在同一曲面模型上计算，所得面积对象与算法并未改变；若换用等面积图面后，像素计数成为可靠的区域比较方法，则新增了可用操作。两种情形视觉上都可能变化很大，科学意义却不同。应追踪的是误差分布、算法条件和可辨模式怎样变化，而非仅看图形轮廓是否新颖。

等面积的教学例子进一步显示这种操作条件。在球面圆柱等面积投影中，取 $x=R(\lambda-\lambda_0)$、$y=R\sin\phi$，则 $dx=R\,d\lambda$，$dy=R\cos\phi\,d\phi$，微小图面面积为 $R^2\cos\phi\,d\lambda\,d\phi$，恰与球面面积元相同。东西向比例仍为 $1/\cos\phi$，南北向比例却为 $\cos\phi$，两者乘积为 1。面积保存来自两个方向的精确补偿，而非各方向都不变形；计算面积可以使用图面，比较形状和长度仍需另行处理。

对一个高纬小区域而言，东西方向被展开，南北方向被压缩，面积可以保持而形状明显变扁。若研究者从这种图面直接估计东西向与南北向的真实距离比例，就会误读。相同面积与相同形状之间的区别，可以通过极小矩形的两边分别计算出来。读者不需要接受某种地图美学立场，只需要检查变换对两个方向的比例因子。

◆ 17.3 从手工困难到计算可用

投影公式可写出来，不等于已经方便使用。实际制图需要把大量经纬度转成图面坐标，必要时还要从图面位置反求地理位置。正变换和反变换的复杂程度可能不同，有些逆

运算需要迭代，有些方法还需处理特殊区域和数值稳定性。计算机使更多方法在日常工作中可用，但并没有使几何约束消失。[A15]

手工制图中，计算成本会影响方法的实际采用；数字系统中，转换可以快速完成，研究者便能针对不同区域和任务重复比较。这里发生的是可操作方法集合的扩展。一种投影并非直到计算机出现才在数学上存在，却可能到计算工具成熟后才成为大量项目可以承担的日常选择。数学可定义、算法可实现、机构可持续使用，是三个不同的门槛。

然而，自动化也会隐藏新的错误。软件可以对输入坐标顺利执行某个公式，即使用户把经纬次序写反、角度单位弄错，或者给出了错误的参考系统。计算成功只表示程序完成运算，不表示输入的物理含义正确。结果显示在熟悉的底图附近，也不保证所有参数都匹配。错误有时很大，容易发现；有时只有几十米，却足以影响工程或地质边界解释。

因此，一次可信的转换至少需要三类检查。第一类检查定义：单位、轴顺序、参数与坐标参考系统是否明确。第二类检查数值：选取已知控制点，比较预期位置与转换结果，必要时做正反变换往返检查。第三类检查任务：计算出来的面积、距离或角度是否采用了适合的规则。往返误差很小可以证明算法自洽，却不能证明起初给错的参考系统已经变对，这一点尤其容易被忽略。

球面模型与实际测绘还隔着参数和数据两层选择。精密任务可能需要椭球模型，计算还须指定其参数以及坐标所用基准；边界线若来自有限精度的采样，解析公式再准确，也无法补回未被记录的曲折。制作区域比较时，应分别报告转换误差和原始边界误差，并检查二者是否已经小于任务容许范围。这样，前节得到的局部比例关系才有明确的实际应用位置。

技术变化还改变协作。过去一张印刷地图的投影说明可能由读者在图廓外寻找；数字文件则可以把参考系统与坐标一起保存。但元数据缺失、格式转换和软件默认值，仍可能切断这条联系。真正的技术进步不只体现在图像出现得快，也体现在定义能够随数据传递，使不同人不必猜测同一组数字代表哪里。若便利性提高而可追查性下降，新的计算能力也可能同时增加误用。

从 D＝G(S,E)的角度，地图呈现 S 会影响使用者选择量测操作，进而组织可比较的差异序列 D；投影和任务的约束 E 则决定这些操作是否有效。图上容易画直线，可能诱发把

它直接当作航线；图面容易数格，也可能诱发把格数直接当作实际面积。若工具在量测时保留所用投影、曲面模型和单位，并提示当前算法的适用范围，这些操作便能回到可检查的条件中。便利性由此与可靠性相接，而不只依靠使用者记住所有隐含约定。

数字工作还使动态重投影成为常见可能：同一份地理坐标数据可以按需要显示在不同投影中。这样做的好处是减少复制数据，风险是屏幕看上去统一，而后台量测规则未必与显示一致。有的软件按数据坐标计算，有的按显示坐标，有的可以采用曲面算法。使用者需要知道所调用的计算究竟是哪一种。显示正确与分析正确是相邻但不同的判断。

一个简单的自检是选取跨越不同纬度的同类小区域，分别比较图面面积与曲面面积。如果差异随纬度呈现系统变化，应该先检查投影比例，而不是立即解释为自然分布变化。对于路线，选择已知位置的控制点并检查方位与长度，也能暴露单位或轴序错误。这样的检查不需要重做全部制图，却能够针对最可能影响结论的环节。

坐标转换的可追查性还包括时间。某些位置参照随地壳运动和测量更新而变化，精密任务需要知道坐标所对应的实现与历元。本章不展开现代大地测量细节，但这个事实进一步说明，坐标不是脱离操作史的永恒数字。对于一般教学地图，这种差异可能远小于图面精度；对于精密定位，却可能进入主要误差。任务尺度决定需要保留多少条件。

投影参数还会影响图面中心附近与边缘的误差分布。把标准线放在哪里，是将较小变形安排到哪些区域的决定；它可以依据人口、研究区范围或测量任务选择，但应公开理由。同一投影名称之下的不同参数因此不是无关紧要的技术细节，而是实际精度如何分配的组成部分。若跨区域比较需要一致误差，应检查各区是否处在相当的位置条件中。

对跨区域比较，误差是否相近有时比各区图面是否同样精细更重要。若一地区靠近标准线，另一地区位于高变形边缘，直接比较图上长度可能把位置条件误作区域差异。误差分布应与数据本身的精度一起进入解释。

◆ 17.4 共同坐标不抹平差异

坐标基准、投影与地图图例经常在使用中被混为一谈。基准规定位置参照的地球模型及其实现；投影规定由曲面到平面的变换；图例规定符号如何表示河流、岩层或行政边

界。它们解决不同问题。两张图颜色相同，并不保证分类相同；投影相同，不保证基准相同；坐标一致，也不保证所绘对象来自同一时间和尺度。

跨图叠加因此不只是把图层拖进同一个窗口。必须先判断坐标是否可比较，再判断对象分类是否相容，最后检查时间与精度是否足以支撑问题。一条地质界线与一条道路看似相交，可能只是图层的空间误差；不同年代的岸线有偏移，可能包含真实变化，也可能包含基准、测绘与投影差异。把所有错位都当作自然变化，会把表达差异误写成历史过程。

共同坐标的价值，是使这些差异能够被分开讨论，而不是使一切差异消失。研究者可以将旧图配准到新框架，明确哪些控制点可靠、哪些区域变形较大，再限定可以讨论的尺度。保留不确定区域，并不等于无法研究；它让后续调查知道哪些地方需要独立补证。越是强大的叠加工具，越需要避免视觉重合带来的虚假确定性。

回到本章开头的两个任务：土地面积比较应在适当的等面积表达或可靠曲面面积算法中完成；航向判断则应根据对应路线和投影性质进行。若同一幅图不能同时满足，使用两幅图并不是科学不统一，而是对几何条件负责。统一性来自共享的位置定义和可转换关系，多样性来自任务不同，和谐性来自不同表达的边界明确而能互相接续。

地图投影由此提供一个特别清楚的发生学案例。科学世界可以有多种合法表达，而合法性并非任意。它由公式、任务、参数和验证共同构成。自然山河不因地图而移动，地图中的可量测空间却确实随着这些条件发生。下一章将进一步追问：即使投影不变，仅改变尺度、分辨率和综合规则，研究者能够看见的对象又会怎样改变？

第 18 章 尺度怎样改变可问的问题，而不改变山河本身

18.1 从原始线条到可读图面

假如把一个地区所有道路、沟谷、露头边界和建筑都缩进一张小图，线条便可能相互覆盖，主要关系反而难以辨认。史密斯地质图对底图的处理，提示早期制图已经需要让地质信息从其他地理内容中突出。[A01] 历史制图与现代数字综合都面对取舍，但操作形式不同：前者通过图面组织突出关系，后者还可将选择、简化与合并写成可重复执行的规则。本章由这种共同问题进入现代教学分析。

比例尺与分辨率相关，却不是一回事。比例尺描述图面距离和地面距离的比例；分辨率说明数据能够区分多小的空间变化。把低分辨率数据放大显示，不会凭空增加细节；把高分辨率数据缩小，也不保证所有细节仍能读清。屏幕缩放只是显示动作，数据采样与综合规则才决定哪些差异被保留为可辨对象。

制图综合包含选择、简化、合并、夸张和位移等操作。选择决定哪些对象进入图面；简化减少线条细节；合并把小对象组合成更大单元；夸张使重要而狭小的对象可见；位移缓解符号冲突。每种操作都有代价。若河流符号被适度加宽，它仍可表达河道位置，却不应再被直接当作实际河宽；若一组小湖合并为湖区，图中连续蓝色区域也不必意味着水面真实连通。

USGS 关于多尺度综合的研究，将可读性、位置、形状与用途联系起来。[A16] 这些要求有时彼此牵制：一条狭窄河道保持真实宽度，可能在小图中不可见；将符号加宽，又会掩盖岸边其他对象。合理综合需确定任务首先保留哪种关系，再说明其余关系受到怎样的影响。以下以河网、湖泊和网格的教学例子展开这一问题，不使用真实区域实验数据。

科学表达中的选择并不自动意味着偏见。没有选择，有限图面无法工作；没有公开规则，选择又容易把结论藏进过程。合理做法是声明保留什么、删去什么、为何这样做，并

允许其他研究者用不同规则检查主要结论。地图的客观性不在于从未经过加工，而在于加工过程受到任务和材料共同约束。

◆ 18.2 固定资料，只改规则

设想一份包含河网、湖泊和岩性边界的原始数据。我们要制作区域调查图与全国教学图。为了比较尺度效应，应先固定原始资料版本，而不是一边换尺度、一边更换调查年代和分类标准。否则，图面差异可能来自真实更新、定义变化或数据质量不同，无法知道究竟是哪一步造成。

在区域图中，小支流可以单独表示；在教学图中，保留全部支流会挤占图面，于是需要选择主干或按一定等级保留。选择规则应能解释一条支流为何留下、另一条为何删去。如果依据长度，就可能忽略短而水量重要的河道；如果依据汇流面积，可能更适合某些水文任务，却不一定适合展示地貌密度。规则没有绝对中立的形式，只有与任务相符且能被审查的形式。

湖泊的处理又提出另一类问题。几个相距很近的小湖，在小比例尺图上可能被合并成一个符号。对于概览湿地分布，这种处理可以有用；对于判断船只能否从一个湖驶入另一个湖，它就不够。对象身份在这里发生了变化：从若干独立水体，转为一个区域性的水域集合。若不声明这层转换，使用者会把集合符号误当成连续水面。

线条简化也需要区分几何误差与关系错误。一条河弯被简化，位置偏差可能在允许范围内；但如果简化使河流跨过分水岭、与另一条河错误相交，便改变了拓扑关系。前一种误差可能影响局部测距，后一种可能使整个流域归属判断出错。综合质量不能只用线条与原线的平均距离评价，还要看任务依赖的连接、包含和邻接关系是否保持。

这种教学比较应固定输入资料，仅改变表达规则，才可分辨规则本身的作用。先记录原始对象及其连接关系，再分别应用调查图和教学图的取舍方案，最后检查哪些差异只影响外观、哪些已经影响任务结论。例如支流删减可以保留主干流向，却可能改变可见的汇流等级；若用简化后的等级判断侵蚀过程，便须返回原始河网验证。输入、规则和任务结果因此需要同时保存，真实区域应用还须公开数据来源与处理参数。

SDE 分析在此指向一种具体的回写关系。若图面显露 S 使读者误判河流连通性，就要同时检查原始连接差异 D 与综合条件 E：是原数据已经断连，还是符号宽度和合并规则制造了连接？只有后者成立，调整综合条件才是相应修复；修订后还需回查 D 所保留的交点、支流与分水岭关系。E 因而是几何容差、连接约束与任务要求的相互规定。修改前后的规则和图件，能够显示修复究竟改变了哪一层。

一个纸上可复核的离散例子能显示合并的后果。设剖线由八个等宽小单元组成，依次为湿、干、湿、干、湿、干、湿、干。每两个相邻单元合成一个大单元后，若规定“含湿即湿”，四个大单元全被标湿；若规定“全湿才湿”，四个大单元又全被标干。原始湿地比例均为一半，输入没有改变，二值规则却给出全湿与全干两种区域描述。这个教学推演说明，粗化后的类别比例和连续性可能主要反映阈值选择。

较合理的做法可能是保留湿地比例，用颜色深浅表示每个大单元的一半面积为湿，而不是强行二分。但比例图仍然不能恢复内部排列：湿地位于左半还是右半，对连通性和水流路径可能很重要。一个汇总量可以准确，却不足以回答另一个问题。尺度转换不仅可能丢失数值，还可能丢失排列关系；后者在许多地球科学问题中恰是机制的关键。

◆ 18.3 哪些新模式是假的

粗化有时会让区域模式更清晰。散乱的小片区合并后，一条大范围带状分布可能显露，帮助研究者提出新的构造或环境问题。但同样的操作也可能制造虚假连续性。两种结果在视觉上都可能很整齐，因此不能用图像的美观判断哪一种更接近自然关系。

考虑一个教学例子：山谷中分布着许多不相连的湿润斑块。若把每个网格内达到一定比例的湿润地表都标为湿润，改变网格大小和阈值后，斑块可能连接成一条连续带。连续带也许反映真实的区域水分梯度，也许只是网格吞并了狭窄的干燥间隔。要区分，需要回看原始资料，检查不同合理参数下是否仍然出现类似结构，并寻找独立水文或地形证据。

这里有两个不同的问题。第一个是稳健性：模式是否对合理规则变化保持基本结构？第二个是解释性：即使模式稳健，是否真的由所提出的机制造成？某条区域分界可以在多种算法下都存在，却仍可能与数据采集制度而非自然过程有关。例如调查范围或分类习惯

在行政边界两侧不同，就可能形成稳定但非自然的界线。参数稳健性能够排除某些表达伪影，不能单独完成因果证明。

统计汇总也有类似风险。把小单元合并成大区之后，相关关系可能增强、减弱甚至反向，因为不同组内部与组之间的差异被重新组合。地图颜色若只显示区域平均，局部极值便可能消失。资源调查关心少数高值地点，教学总览关心广域趋势，两者对平均的要求不同。不能因为区域平均稳定，就推断区域内每一处都具备相同性质。

更微妙的是，综合规则可能以目标结论为依据。如果研究者先相信存在一条连续构造带，再优先保留使其连通的线段，最后把成图后的连续带当作新证据，便出现处理与结论之间的循环。合理做法是先明确规则，再检验它对相反假说是否同样公平，或者比较多套规则的结果。选择过程不必完全自动，但人工判断也应有记录和理由。

在本书的相位语言中，若跨规则的差异序列与原有连接条件无法共同支持一个区域模式，D/E 之间的张力可推动显露 S 重新规定，对应 P1 的候选过程；若已有模式与新比较持续冲突，并迫使修改综合条件 E，则可检验 P3 的候选过程。只有图面外观变了，还无法完成这种判断。需要记录改变前后所保留的空间关系，以及独立资料是否支持重新规定的对象。

参数敏感性给模式的解释划出边界。若网格大小稍变，湿地带就断开，而独立水文资料也未显示连续路径，图上的连通关系只能作为表达结果保留。若多种合理网格都显示同一广域湿润梯度，区域模式便有更强支持，但局部是否连通仍需细尺度资料。两层判断可以并存：粗化揭示区域趋势，同时损失决定水流路径的细节。对象在操作中显露，也在不同尺度的互查中获得适用范围。

线性对象还有一个特别的尺度效应。曲折河道粗化后通常变短；若上下端高程不变，用这条简化线计算的平均坡度会增大。将该坡度代入侵蚀模型，制图操作就可能造成过程强度的假变化。类似地，岸线长度依赖所保留的细节，两份综合程度不同的线长无法直接证明岸线增长或缩短。比较应固定量测尺度，或把不同尺度下的结果作为一组关系报告。

另一方面，保留每一处细节也不总是更好。原始线中可能包含测量噪声，过细的曲折会夸大长度并制造假交点。适当平滑能够提高任务可靠性，但需要以观测精度和问题尺度

为依据。由此，信息损失与误差减少可能同时发生。评估综合不能只问删去了多少点，还要问保留的关系是否更接近任务所需的稳定特征。

边界附近的对象尤其容易受网格起点影响。同样大小的网格，仅平移半个单元，就可能使一个狭长区域从被完整保留变成分散在多个格中而被删去。若结论依赖这种偶然配准，应该报告敏感性，而不是只公布最清楚的一幅图。比较不同网格起点和方向，可以帮助识别由格网制造的形态。

对于线划资料，也可以比较不同简化容差下关键交点是否保持。若主要结论在宽范围内稳定，而局部细节变化，则区域解释可能可靠，局部判断仍需原始图。这样的分层结论比对整幅地图简单标注“准确”或“不准确”更有用。

图件说明因此应让使用者判断：哪些结论在规则变化后仍保持，哪些只在当前容差下成立。它记录的应是解释范围，而不只是软件名称或出图日期。

◆ 18.4 尺度选择的责任

资源调查和课堂总览需要不同地图。前者可能关心局部岩体边界、采样位置与交通条件，后者需要让读者看见大尺度分区。把所有细节塞进课堂图，会妨碍理解；用教学图安排现场采样，则可能遗漏关键边界。地图好不好，必须与将要作出的判断联系。一个图可以在自己的任务中优秀，却在另一任务中危险。

因此，尺度选择应写成有理由的决策：研究需要辨认的最小对象是什么，允许多大位置误差，必须保持哪些连接关系，哪些细节可以删除？这些问题能够帮助使用者理解图的边界，也帮助制图者决定是否需要多幅图协同。统一并不要求所有人使用同一尺度，而要求不同尺度之间有可追查的对应。

小比例尺地图尤其不能支持未经验证的局部反推。区域上被标成某种岩性的地块，内部可能包含未表示的小型岩体；图上没有河道，不等于现场没有季节性水流；一种颜色覆盖某地，也不意味着每个位置都经过同等密度调查。地图的空白与概括都需要解释，不能把缺少符号当成自然缺席的证明。

研究者还应区分数据本身的年代与图面制作年代。新制作的地图可能使用旧调查，旧底图也可能叠加新观测。若两个时期的图有差别，必须先排除分类与综合变化，才讨论地

表真实改变。第 20 章将看到，这一点对历史岸线研究尤其重要：文献、地图和地形需要互相约束，不能仅凭两条线不重合就计算海岸推进。

尺度问题最终使我们获得一种更成熟的整体观。整体不是保留所有细节，也不是删去所有差异，而是在特定任务中组织哪些差异需要共同出现。粗图可以揭示广域关系，细图可以检验局部机制，两者通过明示规则互相校正。下一章将把这种表达纪律带入地貌解释，比较阶段序列与动态平衡如何面对同一地形，而不是让各自挑选最有利的图景。

第 19 章 地貌循环与动态平衡
两种解释怎样接受同一地形检验

19.1 戴维斯提出怎样的次序

山地形态可以被讲成一段历史，也可以被分析成一组正在运行的过程。前一种方式问它处于怎样的演化阶段，后一种方式问什么相互作用维持了当前坡度和河网。这两种提问不必互相排斥，但它们选择证据的重心不同。戴维斯的地理循环与吉尔伯特的侵蚀反馈，提供了比较这种差别的历史入口。

戴维斯 1899 年《地理循环》开篇把地形与结构、过程和时间联系起来。[A14] 这个起点并不是只看形态、不管机制。结构提供材料与初始配置，过程改变地表，时间使改变得以累积。地貌的历史分类由此不同于仅按外观列举山、谷、平原：同样低缓的地面，可能具有不同起点和经历；同样陡峭的谷地，也可能处于不同的区域历史中。

阶段模型的吸引力，在于它把分散形态组织成有方向的序列。一个教学情景可以说明其一般逻辑：地面抬高以后受到河流切割，早期谷地加深而分水岭仍较高；随后谷坡扩大、分水岭降低，长期侵蚀使区域起伏减小。这个序列将不同地貌部位的变化联系起来，使“早”“晚”有了过程含义。它是阶段式解释的简化重建，范围限于所设条件，并非戴维斯 1899 年原文图件的复制。

序列成立有条件。若抬升持续、基准面变化、气候改变或岩性差异显著，形态不一定沿同一顺序发展。一个陡峭谷地不能仅凭“看起来年轻”就取得数值年龄；一个低起伏面也不能仅凭“看起来老”就证明长期稳定。阶段词首先是模型中的关系位置，必须与独立地质证据相接，才成为地区历史判断。否则，研究者会用形态定义阶段，又用阶段解释形态，形成一个无法被新观察打断的圆圈。

阶段解释也有正当用途。它能提出相对次序问题：河流切割是否晚于某地表形成，阶地是否见证多次下切，残留面是否曾经更广泛分布？即使最终不接受单一完整循环，这些问题仍能组织野外工作。批评一个过度推广的模型，不意味着把它曾经产生的全部研究能

力一起删除。思想史应当区分某种表达的启发力、具体解释的证据强度和后世教科书的简化。

19.2 吉尔伯特的反馈链

吉尔伯特 1877 年的报告提供了另一种入口。他讨论流水搬运、磨蚀、坡度、水量与岩石抗蚀性的相互制约，并明确使用动态平衡的思想。[A13] 研究的重心由地形在序列中处于哪一阶段，转向各过程怎样彼此调整。反馈与相互依赖已在这套地貌分析中展开；本书的发生学复审将追踪它们如何决定观察对象和比较路线，具体机制仍由水沙与岩性关系说明。

考虑一条沿程流量近似不变的河段。若某处坡度较大，水流搬运和侵蚀能力可能较强；若这使河床更快降低，局部坡度又可能随之减小。相反，物质堆积可以提高床面，改变水面坡降并影响后续搬运。这里的坡度不是只作为原因出现在方程右边，它也受到侵蚀与沉积结果的改变。这种双向性，使地形不能被当作始终固定的容器。

河流携带的负载同样重要。一条河可用于搬运的能力，与实际进入河段的泥沙量不一定相等。若来沙超过可搬运量，部分物质堆积；若搬运能力有余且河床可侵蚀，则可能发生下切。负载还会改变磨蚀作用，水流与颗粒并非两种可以完全分开的因素。实际河流更复杂，但这一结构已说明，不能从流速或坡度一个量单独推断全部地貌变化。

吉尔伯特还考虑支流汇合。两股水汇合后，流量增加，主流与支流之间的适宜坡度关系会改变。河流系统因此不是若干独立沟槽拼在一起，而是通过水和碎屑的输入输出相连。一个河段快速下切，会降低上游出口条件；增加的碎屑又可能使下游暂时堆积或减缓侵蚀。局部变化沿网络传播，形态调整具有方向和时间差。[A13]

动态平衡不等于静止。河流可以持续搬运物质，而某些统计关系或相对形态保持近似稳定；地表也可以整体降低，同时局部坡度关系趋于协调。若把平衡理解为“没有变化”，就错过它的过程含义。反过来，平衡倾向也不保证自然总能达到平衡。外部条件变化过快、极端事件频繁或调整时间过长，都可能使系统长期处于过渡状态。

报告中一个特别重要的细节，是吉尔伯特没有把理想分析与观察完全一致当作写作前提。他讨论分水岭附近坡形时，指出实际顶部存在理论尚未解释的凸曲部分。这种持续出现的差别不能简单归为偶然，于是成为遗漏因素的线索。[A13] 理论的成熟在这里表现为

能够指出自己解释不了什么，并把这个缺口定位到具体空间部位，而不是用含混的复杂性一笔带过。

从 SDE 角度，这条反馈链具有明确的特征纠缠 E：坡度影响侵蚀能力，侵蚀又改写坡度；上游输出限定下游负载，下游水位条件又影响上游调整。差异序列 D 须同时包含沿程位置与调整时滞，否则同一时刻的上下游观察可能被误当作彼此独立的状态。显露 S 则是据这些关系辨认的地貌过程状态。这样的分析要求将测点放回水沙传递网络，才能判断局部异常是原因、响应，还是两者之间尚未完成的调整。

吉尔伯特关于负载的讨论还包含一个对日常观察很有启发的区别：搬运较多细颗粒，与搬起较大颗粒，不完全是一种能力。水流增加时，若有足够细物质供应，负载可以增加；若细物质不足，更大的颗粒可能开始被移动。[A13] 因此，河床上看见粗颗粒，并不能单独告诉我们长期平均输沙量。颗粒大小、可供物质与流动条件必须一起解释。

他也指出，塑造河床坡度的可能主要是高水位时期，而不是调查者最容易进入河道的低水位时期。[A13] 这使观测时机本身成为科学问题。一次晴日调查能够精细测量河床，却未必直接看见最有效的搬运过程；洪水期间过程强烈，测量又更困难。地貌知识因此需要把不同条件下取得的资料联系起来，而不能将便利观察到的状态当作整个系统的代表。

岩性与坡度的联系同样不是简单的一因一果。坚硬河床抵抗下切，可能保留较大坡度；较大坡度又增加侵蚀能力，使不同岩段的长期作用趋于某种协调。若仅看到硬岩对应陡坡，便说硬度决定了一切，就遗漏了调整过程；若只看到坡度越大侵蚀越强，也会忽略为何较陡地段仍然突出。反馈分析把两条单向关系接成可以演化的系统。

◆ 19.3 给两种解释同一道题

比较两种框架，需要让它们面对同一组资料。设一个教学流域，上游较陡，下游较缓，中段有明显陡坎；不同岩层沿河出露，附近保存一组阶地。阶段式解释可能把陡坎看作新一轮下切的前缘，把阶地理解为过去河床；过程式解释可能先问陡坎是否对应更抗蚀的岩层，以及坡度是否与流量、来沙和岩性相协调。两种解释都合理，但它们应提出不同的新增观察。

如果陡坎主要反映沿河传播的历史调整，调查应关注陡坎与构造边界是否分离，不同支流的陡坎高程与位置是否呈现有意义的传播关系，以及阶地年代能否约束下切次序。如

果陡坎主要由岩性控制，则需要检查它是否稳定对应特定接触带，同类岩层在其他河段是否出现类似坡度变化，跨接触处的侵蚀抗力是否足以解释差异。这里没有宣布其中一条一定正确，而是让同一形态产生可区分的测试。

两种原因还可能同时存在。一个向上游传播的下切扰动遇到坚硬岩层，可能减慢并停留更久，于是今日陡坎位置同时反映历史扰动和材料阻力。若硬要在“阶段”与“平衡”间选唯一赢家，反而会把真实复合过程切碎。更好的比较是问各自解释了哪部分资料，剩余部分需要什么机制。模型组合应有明确分工，而不能在遇到反例时随意添加原因。

独立年代在这里特别关键。若阶段完全由形态判定，再用阶段证明形成次序，就没有独立检验。阶地或沉积物的年龄、交切关系与构造活动证据，可以限制历史解释。过程测量也同样需要时间尺度匹配：一次低水位流速不能代表长期塑造河床的有效流量，一次暴雨的泥沙输出也不能自动代表多年平均。历史模型不能逃离年代，过程模型不能逃离代表性。

还可以把物质守恒写成教学关系。设固定河段内的固体泥沙储量为 M，单位时间进入和离开的固体质量为 Q 入、Q 出，并暂不计侧向输入，则 dM/dt＝Q 入－Q 出。两项 Q 均为质量流率，不能与时段累计泥沙量混用。若要由 M 转为床面高程变化，还需河段几何、颗粒密度与孔隙率等条件。输出超过输入时，差额须来自段内河床或岸坡储量；若岸坡被划在系统边界之外，它的供沙就必须另列输入。守恒关系虽不决定输沙率，却可排除物质来源不明的解释，也要求阶段重建说明被侵蚀物质的去向。

同一流域中的坡面与河道还具有不同响应时间。河道下切后，坡脚条件改变，坡面可能经过较长调整才达到新的形态；于是某个时刻的河道与坡面不一定显示同一“年龄”。如果把整个流域一律归入青年或老年，可能抹去这种异步。过程式分析能够揭示异步，但若完全忽略过去扰动，也可能把暂态误认成当前条件下的稳定格局。两种框架的互补，来自对这些盲点的相互纠正。

比较的结果可以是一幅有分工的解释图：某些河段主要受岩性控制，某些部位保留历史下切信号，另一些尚不足以区分。每一分区都应对应判据及待补资料，例如接触带附近需要抗蚀性比较，离开接触带的陡坎需要年代和传播关系。这样，两种模型的组合便具体

到可再访的地点。SDE 记录对象、差异与条件怎样随比较而改变，水沙动力学、年代学和构造关系则决定这些改变是否成立。

这个教学流域还可以用于检验模型的时间方向。若今日坡度与岩性配合良好，过程模型可能认为当前调整较充分；但阶地显示过去经历明显下切，并不因此与之矛盾。过去扰动之后，系统可以重新接近某种协调。反过来，今天的形态若与当前来水来沙不协调，也不必说明过程定律失效，可能是外部条件刚刚改变，地形尚未完成响应。历史资料帮助判断离平衡的原因，过程资料帮助判断历史遗迹会被保留多久。

相同形态还可能来自不同路径。一个宽谷可以由长期侧蚀形成，也可能继承过去较大流量时期的河道；一个低缓面可以是侵蚀面，也可能与沉积充填有关。阶段分类若只依轮廓，容易把它们归为同类；过程解释若只看当下流量，又可能低估遗继作用。要区分，需要谷底沉积、基岩形态、年代与流域变化等证据。对象的同一性因此不能仅由外观相似决定。

不同尺度的平衡也应分开。一个河段的输入输出在多年平均上接近，并不意味着整个流域没有净侵蚀；坡面物质可以持续进入河道，再由河道传出。局部储存稳定、区域物质流失与整体高度变化可以同时发生。若不声明系统边界，“平衡”便会在不同意义之间滑动。相同术语在不同边界上有不同计算责任，这是过程模型能够与阶段历史接合的重要前提。

如果要把两种解释落实到野外路线，选择也会不同。阶段研究可能优先访问保存高位面的山顶、阶地和交切关系清楚的谷口；反馈研究可能沿河连续测量坡度、粒度、岩性和支流输入。两套路线各自容易遗漏另一套重视的材料。把它们结合，不是简单增加行程，而是安排能够连接历史次序与当前过程的关键地点，例如有年代约束的阶地与现今河床之间的关系。

一个可操作的共同问题是：某段河谷为何比相邻河谷宽？历史解释可提出其形成更早或经历较大流量阶段；过程解释可提出岩性较弱、侧向侵蚀条件不同或来沙较多。比较需要控制流域面积、岩性和基准面，再寻找年代或沉积记录。若这些条件不受控制，宽窄差异可以被任何框架解释，模型比较就只是口头竞赛。好的共同题目必须让至少一种可能观察对某个解释不利。

局部水系的改变还可能来自河流袭夺，使某条河突然获得更多水量而另一条减少。若只按现今流域面积解释旧谷地，就可能出现显著不匹配。历史路径分析可以寻找遗留谷口与排水方向，过程分析则检验新增流量怎样改变坡度和输沙；两者在同一问题上形成可检验的接续，而不只是概念上的兼容。

◆ 19.4 历史发展不必单向取代

若只按概念的流行次序写地貌史，就容易把阶段模型放在前面，把过程模型放在后面，再宣布后者全面取代前者。吉尔伯特的反馈分析早于戴维斯 1899 年的论文这一事实，已经提醒我们：思想不是按一条从叙事到定量、从简单到复杂的直线排队。不同研究任务、材料条件和表达方式可以长期并存，并在不同场合获得作用。

阶段模型帮助人们组织过去的次序，过程模型帮助人们分析当前及长期调整的机制。若一个问题主要是“这条阶地相对于哪次下切形成”，时间关系不可替代；若问题是“为何同一河流某段更陡”，岩性、水量与负载关系可能更直接。两者回答不同层次问题时，应当写成分工，而不是制造推翻情节。只有当它们针对同一条件给出不相容预期，才需要通过证据选择。

同样，过程语言并不天然比阶段语言更科学。若“系统达到平衡”只是给任何现状贴标签，没有说明输入、输出、反馈与时间尺度，它与空泛的“成熟地貌”一样缺少判别力。相反，一个明确提出初始条件、演化顺序和独立年代预期的历史模型，完全可以接受严格检验。评价标准应落在解释能否承担反证，而不是词汇属于哪个时代。

这种比较同样约束 SDE 分析。若将本应随侵蚀变化的坡度固定下来，便会切断反馈；若把上下游相互影响的样点当作独立重复，便会高估证据数量；若忽略调整时滞，又可能把过渡状态误作平衡。这些都是可以定位的错误，而非三个字母本身提供的新机制。常规地貌学若已处理这些关系，SDE 的作用主要在于把其对象形成方式与本书其他案例对照。

地貌思想的丰富，来自它能同时容纳历史与运行。山河既有不可逆的过去，也有持续调整的现在；现今形态既是遗迹，也参与形成后续形态。下一章把这些问题带到中国下扬子调查。那里不仅有岩层和河流，还有县治年代、海塘遗址与地方文献。研究者需要把不同性质的记录接到同一空间过程上，并让它们能够相互修正。

第 20 章 丁文江的下扬子调查
地方文献怎样成为海岸变迁证据

20.1 工程问题怎样成为调查

1916 年，丁文江应上海黄浦江水道总局之邀，对扬子江下游开展地质调查；1919 年，《芜湖以下扬子江流域地质报告》出版。[A17][A18] 这条简短时间线包含三种不同活动：工程机构提出调查需求，研究者在区域中组织观察，报告将材料转为可公开讨论的论证。不能把报告出版年当成所有观察发生的年份，也不能仅凭工程任务推断每项科学结论都由工程目的直接决定。

下扬子区域把多种问题聚到一起。河流进入海洋，沉积物堆积与搬运改变河口，海岸推进或后退影响港口和土地，低平地形又使一些变化难以仅凭露头判断。这里需要的不只是识别岩石，还要理解水道、地形和时间。某些过程超出一次调查能直接观察的范围，历史记录因此成为潜在证据。

地方志记载建县、迁治、修塘和水患，本来服务于行政、工程和地方记忆。地质研究使用这些记录时，必须改变问题方式：不是问某个县历史多么悠久，而是问它在什么年代、什么位置出现，以及这一位置对当时陆地与海洋关系提供怎样的约束。文献中的社会事件与自然过程之间需要建立桥梁，桥梁本身就成为论证的一部分。

报告的公开片段显示，丁文江实际比较县治年代的空间分布，又使用海塘年代与距离讨论海岸变化。[A18] 这不是简单在地质章节前添加文化背景，而是使文献进入推断链。一个年代可以限制某处至少何时已具备设治或筑塘条件，一段海塘位置可以限制某时期的防护边界。与此同时，地形和沉积关系也能反过来检查文献解释是否合理。

不同材料的接合也使调查组织具有具体的科学任务。地方文献需要保留原文和日期性质，野外观察需要记录位置与地质关系，二者才能按共同问题比较。机构能够汇集资料，尚须靠考证和定位把旧地名转为可用地点，再由地形与沉积关系检验其自然含义。调查能力因而表现为这条工作链能否持续运行，其成效应从材料怎样被连接来判断。

◆ 20.2 县治分布不是天然海岸线

原报告第 73 页公开片段，以一条比较线两侧的县治年代分布讨论区域历史。[A18] 这里的基本思路是空间上的较早与较晚：如果较晚设立的城市更多分布于某一侧，可能与土地形成和开发的方向有关。但行政年代首先是行政事件的时间，它与自然陆地形成之间的间隔必须另外分析。

一处土地可能早已出露，却很久以后才建立县治；也可能先有季节性利用，再出现稳定聚落；行政机构还可能因人口、交通或政治安排迁移。因此，设县时间通常不能直接等于海退到该处的时间。它更适合提供一个有条件的界限：若能确认县治当时确在某位置，并具备稳定使用条件，那么该处在那时至少已经能够承载相应活动。这样的界限比一个看似精确但含义错误的出露日期更有价值。

地名同一性是另一道关口。同一个名称可能在不同年代指向不同地点，旧县辖境也不等于县治所在地。若把辖境中心、现今政府位置和历史城址混用，地图上就会生成虚假的迁移。研究者需要查明记载指的是建置、迁治、改名还是辖区调整，再把地点落实到可比较的坐标范围。文献考证在这里不是外加的人文修饰，而是空间资料能否成立的条件。

即使地点可靠，县治分布仍可能受到交通与地形的选择。人们可能偏好较高的天然堤、旧沙洲或河道附近，而不是按陆地出现顺序均匀布点。若将所有城市位置连成岸线，就会把选址制度投射为自然边界。更稳妥的做法是把每个点当作局部约束，再与地貌单元、沉积环境和历史地图比较，逐步限定可接受的海岸位置。

这一过程能够形成真正的交叉检验。若文献声称某地很早已有稳定城址，而沉积资料显示相应年代仍为深水环境，就需要检查年代、位置、沉积解释或城址含义。不能预先规定自然资料一定正确，也不能以古籍权威压过地质证据。冲突的价值在于使各自隐藏的前提显露，促使研究者寻找能够区分解释的新材料。

用 SDE 术语说，D 是按文献年代和空间位置排列的差异序列；E 是城址定位与陆地环境判断的相互限定：历史城址若无法定位，沉积解释便失去对应地点，地层若显示当时不具备设治条件，又要求重审地名和日期的归属。S 是由这些互查支持的海岸变化阶段。这种证据组合能够工作，正因为文献与地质各自保留了纠正另一方的可能，而非仅在同一地图上同时出现。

◆ 20.3 海塘距离怎样变成速率

海塘比县治更接近海岸过程，却仍不等于天然岸线。筑塘地点取决于防潮需要、地基条件、工程能力和土地利用；一道海塘可能位于当时潮水能够达到的范围附近，也可能有意退后。旧塘与新塘之间的距离，只有在明确各自功能和位置以后，才能参与海岸推进的估计。

原报告第 75 页片段指出，不同地区海塘的年代并不相同；第 77 页则对特定塘段比较年代与距离，把 1472 年至 1884 年的时间和约六英里的间距联系起来，得到约六十九年一英里的平均值。[A18] 这项历史推算的对象是特定区段与时间范围。它不能未经转换就扩展为整个江苏海岸的固定速度，也不应拿来预测今天的岸线变化。

这里的算术并不困难。以距离除以时间得到平均速度，或以时间除以距离得到每单位距离所需年数，都可以。困难在于分子和分母是否对应同一过程。若距离沿海岸方向量取，而想解释的是向海推进，方向就不匹配；若两塘并不平行，平均间距需要说明如何取得；若时间一个是始建、另一个是重修，历时也可能失去统一含义。算式越简单，越容易让这些问题从读者视线中消失。

原报告的片段还区分前后塘之间不同的距离与时间关系。这提醒我们，平均值会掩盖阶段差异。某段时期推进较慢，另一时期较快，可能与来沙、河口位置、围垦或风暴侵蚀有关。即使总平均可靠，也不能由它恢复每一年的岸线位置。线性外插是一项额外假设，需要证明相关条件在外推时期仍可近似保持。

海岸变化还具有方向和空间不均一性。一处向海淤长，邻近地区可能受到侵蚀；沉积物从一个区段转移到另一区段，并不要求整个区域同时增加陆地。若只选取推进明显的塘段，就可能低估这种差异。区域结论需要覆盖不同地貌位置，说明哪些区段有记录、哪些没有，以及记录的选择是否偏向容易保存的地点。

速率的可靠范围同时受位置与年代影响。若推进距离为 L、对应历时为 Δt，教学式 $v=L/\Delta t$ 给出平均速度；L 的不确定性来自塘线位置、量测方向和几何近似，Δt 的不确定性来自日期类型与记录精度。相同的距离误差，在短距离塘段中占比更大；相同的年代误差，在较短历时中影响更大。因此，区域比较不能只排列平均值，还应比较各段可达到的分辨率。现有公开片段不足以量化这些误差，此处只说明它们进入推断的位置。

这个案例特别清楚地显示，文献并非只提供时间，地图并非只提供空间，工程史也并非只提供背景。塘的工程身份决定哪一条线应被量测，量测方向决定距离含义，年代解释

决定历时。三者相互规定以后，才产生可讨论的速率。新对象是一个具有位置、区段、时期和误差条件的海岸变化量，而不是孤立的一串数字。

原报告第 77 页的公开片段允许核对这一步除法：1884 减 1472 为 412 年，以约六英里作间距，得到每英里约 68.7 年，与报告的约六十九年相符。它是平均速度的倒数，单位为年/英里；若改写为速度，则为约 6/412 英里/年。两种表达须保持方向和单位一致。算术相符说明时距换算能够接续，却尚未验证塘线定位、测距方向以及两次日期是否标记可比工程状态。

比较片段中的前一阶段，1472 年至 1590 年约 118 年，对应约一英里的间距；后续阶段的平均变化显然不同。即便距离都是近似值，这个对照也说明一条总平均隐藏了阶段不均匀。研究者应进一步问，差别来自自然淤积速度变化、塘线功能变化，还是量测和文献范围不同。原文保存的分段信息，比一个被后人广泛转述的区域平均数更适合提出新问题。

此外，陆地向海扩展与地面沉降并非逻辑上互斥。只要沉积补给足够，局部海岸可以在地面下降的同时向海推进。原报告公开片段也讨论了这类关系。[A18] 这提醒我们，岸线位置是多过程共同作用的结果，不能凭推进方向直接判断地壳上升或下降。垂直运动、沉积厚度与水平边界变化必须区分，才能让海岸史接入更大的地球动力过程。

海塘作为工程还会改变后续记录的保存。塘内土地受到保护以后，原先经常被潮水改造的表面可能更稳定；塘外则继续经历淤积与侵蚀。后来调查者看到的沉积差异，可能部分由筑塘本身造成。工程年代因而既帮助标定时间，也帮助解释为何同一时期两侧留下不同记录。只有把这条作用路径纳入分析，才能避免把人类工程造成的保存差别全部归为自然环境变化。

同样，地方志对重大水患和修塘往往记载较多，对平稳年份记载较少。史料丰富度并不等于事件频率，缺少记载也不等于海岸没有变化。将文献与连续沉积记录相接，能够部分补偿这种选择性，但仍应保留每类资料的时间分辨率。

汇总这类材料时，宜将记录的事件类型与观察时间分开：修塘年份说明工程动作，受灾年份说明一次水患，调查年份说明某条遗迹在何时仍可辨认。三类日期在同一时间轴上，却未必测量同一过程；保留其身份，才能使后来者重新组合证据。

20.4 让材料相互反驳

把县志、海塘和地质观察放在一起，最有价值的结果不一定是彼此吻合。假如三者一致，研究者可以增强对某种变化方向的信心；假如它们冲突，就有机会发现旧图配准、地名迁移或沉积解释的问题。证据组合的力量，不在于来源越多越壮观，而在于不同来源具有不同的失效方式。

设一份县志记载某城早已建立，一张旧图将城画在今天近海位置，而地形资料显示现址周围陆地较年轻。首先应核查旧图是否按现代几何精度绘制，城名是否经历迁移，以及记载是否指辖区而非城址。如果这些解释仍不能消除矛盾，再检查地质年龄与环境判断。这样的顺序不是固定仪式，而是依据每项资料最可能的误差来源设计的排查路径。

若海塘之间的距离与文献中的推进描述不符，也应区分塘线重建、局部改线和天然岸线变化。工程设施本身参与地表过程：围垦、防潮与水道整治可能改变沉积环境。因而海塘既是记录，又可能是原因。把它只当一把被动标尺，就会遗漏人类行动对所测对象的回写。这种双重身份应落实到具体工程及其影响，不能仅凭“人地互动”四字就宣称因果已经解释。

在调查协作中，原文出处、历史地名、定位理由、日期类型与地质判断应沿同一记录链保存。若一种解释要求补查旧志，另一种要求重测塘线，两条路线的差别就能追溯到各自不确定的环节。这里可以讨论研究问题怎样引导操作；至于机构是否据此修改工作规范，还需前后版本、档案或实际记录。报告中的科学推理与其可能造成的制度变化，分属不同层次的历史结论。

丁文江下扬子研究的意义，正在于地方记录能够在现代地质问题中取得新的证据身份，同时保留接受反驳的条件。传统知识与现代科学之间不是整块替代，也不是无条件融合，而是逐项转换：一个日期变成时间界限，一条塘线变成空间约束，一处地层观察变成对文献解释的检验。转换成功之处产生新的历史对象，未能转换之处则应保留为空缺。

从时间标尺到地表测绘，前面两编展示了科学对象的多种构成：有的依靠矿物内部的变化，有的依靠公共界线，有的依靠投影和尺度，有的依靠文献与自然遗迹相接。下一编把比较范围推向海洋两侧。大陆边缘看似相似，究竟只是形状巧合，还是能够支持大陆移动的历史证据？同样的问题将再次出现：什么可以比较，哪些共同性具有机制意义，哪些差异要求新的观测条件。

第 5 编

大陆移动争论与海洋证据条件

第 21 章 海岸线相似：从拼图直觉到可检验的大陆组合

第 22 章 魏格纳的大陆漂移：证据综合为何没有立即形成共识

第 23 章 固定大陆学说：局部有效为何不能保证全球成立

第 24 章 海洋观测基础设施：新证据为什么终于能够进入争论

第 25 章 未被接受的思想如何存活：延续、修订与真正的失败

第 21 章 海岸线相似
从拼图直觉到可检验的大陆组合

21.1 从现代海岸到大陆边缘

把南美洲东岸和非洲西岸放在同一幅地图上，一个很容易提出的问题是：这两条曲线是否曾经相接？问题的吸引力来自形状，但问题能否进入科学，并不由吸引力决定。同一张地图既可以引出有待检验的古地理重建，也可以引出无穷无尽的视觉联想。两者的分界在于，研究者是否预先规定了可以移动的对象、可以改变的边界和必须承担的不吻合。允许任意拉长一个大陆、剪去碍眼的突出部，再补上缺口，几乎任何两块图形都能拼接。只有限制这种自由，吻合才具有信息。

现代海岸尤其容易被误认为大陆天然的固定边界。海岸实际是当下海面与地表相交的位置。海面改变，低缓岸段就可能大幅移动；河流输送沉积物，三角洲便向海推进；侵蚀削去陆地，岸线又向内退缩。即使大陆整体没有水平位移，海岸轮廓也能够变化。因此，不能把现代海岸的每一个曲折都当作大陆分裂时留下的断口。用今天的水陆分界追索古老大陆接触，相当于把一个随时间变化的交线当成永恒轮廓。

向海下寻找大陆边缘，会改变拼合问题，但不会自动消除它。某一水深的等深线是一项可以公开复现的操作定义，却不必处处与同一地质结构重合。陆架宽度、陆坡沉积、裂解时的伸展以及后来发生的构造改造，都会改变轮廓。选择等深线的优点，是将边界由随手勾画转为可报告的规则；它的局限，是几何统一未必意味着成因统一。一个严谨的组合必须说明，为什么这条线可以近似代表需要重建的接触，而不能只说它拼起来最好看。

拼合还依赖地图投影。把球面大陆画到平面，各种投影都会改变某些距离、面积或方向。如果拿纸片在平面上任意平移旋转，残差中可能混入投影变形。大陆在球面上的相对位置，应以球面旋转描述。这里不是要求每一幅教学拼图都进行专业计算，而是要求研究者知道教学操作省去了什么。纸片拼图能够说明一种可能性；定量古地理模型必须使几何条件和地质主张相配。

最容易忽略的另一项条件是尺度。大陆边缘上的小海湾与跨越数千公里的整体弧线，不能简单按同一重要性计分。小尺度起伏受到局部侵蚀和沉积的影响更强；大尺度形状又可能因过度平滑而失去判别力。研究者必须报告保留了哪些空间波长、舍弃了哪些细节，以及这种选择是否在看到拼合结果之前已经决定。否则，所谓相似只是不断调整尺度之后被挑出来的一种显露。

本章所说的古大陆对象，不是从地图背后直接取出一块已经标好名称的整体，而是在边界定义、旋转规则、误差许可与外部证据之间成立的研究单位。它指向真实发生过的地质历史，但其可辨认身份需要这些操作共同承担。若换一条有同等地质理由的轮廓，组合便完全崩溃，这个对象还没有稳定；若合理变化只改变局部残差而保留整体关系，它才获得继续检验的资格。

◆ 21.2 数值拼合与剩余空白

Bullard、Everett 与 Smith 于 1965 年发表的大西洋周围大陆拼合研究，把数值方法和最小二乘判据明确带入这一问题。[B19] 这个节点的意义并非计算机宣布大陆漂移正确，而是把“看上去很像”转化为“在明确规则下，哪一种相对位置使偏离较小”。算法为争论提供了共同入口，也让争论对象更加具体：边界从哪里来，哪些点参加计算，距离怎样定义，何种残差应当被视为重要。

用一个球面拼合的教学问题可以说明这种变化。先在两侧大陆边缘选取点，并规定对应方式；给定一个球面旋转后，计算对应边缘之间的距离，再将距离的平方加权求和。寻找使总量较小的旋转，就把肉眼判断转成了拟合问题。点位、对应方式和权重须与旋转参数分开：若对应点也可随意更换，优化便同时改变了“拼在哪里”和“什么算吻合”；若某段陆缘采点特别密集，未作权重处理时，它会主导结果。这一简化说明只解释拟合逻辑，不复现原论文的点集与计算。

平方残差的作用之一，是让较大偏离付出更高代价；但它也可能让少数边界异常主宰结果。一个大河三角洲造成的突出，未必应当决定整块大陆的旋转。解决办法不能是看到哪个点不顺眼就删哪个点，而应依据独立地质信息识别后期沉积覆盖，或比较不同权重方案下结果是否稳定。如果剔除规则仅由拟合成功与否决定，算法只是把视觉偏好藏进了数据清理。

残差图比单一拟合分数更有历史意义。一个组合可能具有较小的总体误差，却把全部重叠集中在同一长段陆缘；另一个总体误差略大，但偏离主要分布在受到明显后期改造的区域。两者不能仅靠一个数字定胜负。连续的系统偏离可能指出边界选择有误，可能提示大陆曾发生内部变形，也可能说明所设定的刚性组合根本不适用。分散的小偏离则可能来自测量与近似。残差不是等待擦掉的污点，而是模型必须继续解释的地质信息。

空隙和重叠也不能视为完全对称的失败。空隙可能涉及后来被拉伸变薄的陆缘、无法保留的部分或边界位置误差；重叠可能涉及沉积加积、缩短变形或错误旋转。这里只能列出候选机制，不能把每处不吻合都自动分配一个便利解释。提出修补之后，还要追问其独立证据在哪里：是否存在相应的伸展构造，是否有沉积体范围，是否有能够支持缩短的剖面。补丁若没有独立约束，拟合就会失去失败的可能。

真正提高可信度的，是敏感性检验。保留一个地区而拟合其他地区，再检查被留出的轮廓；使用另一种合理的平滑程度，再比较旋转是否明显变化；把误差较大的点位权重降低，而不是直接删除。结果若在这些变化下维持同一基本组合，说明它不是某一套人为选择的偶然产物。这类检验所增加的是稳定性判断，不能替代跨洋岩带的年代或化石材料，却能明确几何证据究竟承担多大责任。

数值拟合所改变的，是轮廓假说的公开方式：从一幅可以反复摆动的拼图，变成含有边界定义、旋转规则与剩余偏差的组合。它仍留下地质解释的任务。同一缺口，若由陆缘伸展造成，应与伸展区的位置和尺度相容；若由轮廓选择造成，改用另一条合理边界时应有相应变化。由此，计算的终点成为地质检查的起点。总体误差的下降只回答几何改善了多少，局部结构与年代材料才说明这种改善是否具有历史意义。

还可以把拟合看作模型复杂度的选择。只允许两块完整大陆相对旋转，参数较少，残差可能较大；把每块大陆再分成许多可以独立转动的小块，残差容易下降，却增加了结构假设。两种方案不能只按误差大小比较。每新增一个可动块体，都应回答其边界在哪里、变形发生于何时、哪些地质记录支持独立运动。否则，细分只是用更多自由度换取更好看的拼图。

残差还应区分随机散布和有方向的结构。若沿一段边缘持续出现同向偏离，可能提示模型遗漏了某种系统过程；若偏离在误差范围内无规律变化，则不必为每个点发明地质故事。研究者要防止两个相反倾向：把真正的结构信号当噪声抹去，或者把测量噪声解释成

一连串重大事件。地质判断与统计判断在这里互相约束，任何一方单独工作都容易过度解释。

一个可供他人复查的拼合结果，因而应同时展示最终位置、参与边界、旋转参数、误差尺度和剩余区域。它不像一道只有唯一标准答案的拼图游戏，而像一份对古地理负责的模型说明。读者能够看到模型在哪些地方成功，也能据此选择下一处最值得调查的地点。几何对象由此获得了继续生长的接口。

因此，拟合结果的可重复，不是任何人凭直觉都能摆出相似图案，而是不同研究者使用相同规则能够得到相近结果，并知道结果对规则变化有多敏感。公开这种敏感性，才把私人眼力转为共同的地质判断。

◆ 21.3 让化石与岩带独立发言

几何拼合最有力量的下一步，不是继续寻找更多相似的曲线，而是提出尚未参与拟合的地质预期。假如两个今天相隔大洋的边缘曾经相接，那么某些岩带的方向、形成时代和变形历史应当能够延续；某些古生物分布与沉积环境的组合也应获得较连贯的说明。这里的“应当”不能扩大为所有地层都逐层相同，因为一个完整大陆内部同样存在环境差异、沉积间断和不同构造区。好的预期既有明确内容，也有允许自然差异的边界。

岩性相似本身很弱。花岗岩、砂岩或玄武岩在世界上多处出现，单靠颜色与名称不可能确认它们曾属于同一地质带。证据需要逐层增强：先比较形成时代，再比较岩石组合、变质条件、构造方向和后续事件。若各项特征在重建位置上形成具有顺序的对应，偶合的解释负担才会增加。相反，两侧都有一种常见岩石，只能构成检索线索，不足以承担大陆组合的证明。

化石材料同样要经过对象鉴定。相同属名可能覆盖较大的时间跨度，鉴定标准也可能在研究过程中改变。若两个地点的化石其实分属不同年代，空间上拼得再近，也不能支持同一次地理连续。对于陆生或淡水生物，跨越宽阔大洋的困难可能提高大陆相接解释的吸引力；但这种论证必须考虑生物的生态能力、漂流可能、当时海陆格局和记录缺失。不能仅凭今天的生活习性，为古老生物划定绝对不可逾越的路线。

陆桥解释因此应作为真正的竞争方案进入，而不是一句“旧人迷信固定大陆”便被排除。若假设曾存在连接陆地的桥梁，它可能解释某些物种分布；但也要说明桥梁的地质基础、存在时段以及后来消失的机制。一个陆桥解决一种化石，另一个陆桥解决另一种化石，若不断增加却没有独立痕迹，复杂性便逐渐上升。大陆重组也必须承担相同纪律：不能遇到不吻合便任意旋转小块，直到任何分布都得到解释。

独立性尤其关键。如果先根据某条古老造山带确定大陆旋转，再把同一造山带的接续列为外部验证，实际上只有一次信息输入。正确写法应是：该岩带帮助规定了模型，随后由另外的年代关系或未使用的地层区间检验。独立证据不是来源文章数量，而是约束链条是否重叠。两篇论文可能使用同一批标本；同一篇论文也可能包含具有不同误差来源的几类检验。

可以把研究分为建模资料、检验资料和解释剩余三栏。建模资料决定候选组合；检验资料在组合确定后检查预期；剩余则保存未被解释的地区与事件。这样组织资料，有助于防止成功故事吞掉反例。若所有不吻合都被放进“后期变化”而不再追查，模型得到的并非更强支持，而是更大的免责范围。一个有价值的古大陆组合，应当使至少一部分原本陌生的地质对应变成可预先提出、可以查错的任务。

◆ 21.4 何时需要改写大陆身份

大陆组合成立之后，研究者工作的单位发生变化。现代行政地图和水陆轮廓把南美洲、非洲分为两块遥远区域，古地理重建则把它们的部分边缘放进同一历史结构。新增的不是一张更漂亮的合照，而是一组可以互相提出要求的属性：某条岩带若在此终止，彼处应寻找什么；某次沉积转折若跨越组合区域，其年代如何对读；某处偏离若来自后期伸展，邻区应留下怎样的结构记录。

在 SDE 分析中，显露 S 是受到约束的古大陆组合，差异序列 D 包括轮廓选择、旋转比较、残差检查及地质对读。特征纠缠 E 具体发生在边界精度、可接受变形与跨洋连续性之间：边界精度决定多大的不吻合值得解释，地质对应则反过来限制哪些轮廓能够代表原始接触。若一条较浅等深线拼合良好，却割断了年代相容的构造带，几何最优便需要重新评估；若岩带接续只靠任意扩大边界误差维持，地质相关也不能单独保证组合成立。

组合稳定后，调查可集中于最有区分力的连接地段。若两种旋转都能容纳大部分轮廓，却使同一岩带分别接到不同地区，优先比较这两处的年代与变形记录，比继续加密已经吻合的海岸点更有效。新资料由此可能加强原组合，也可能改变旋转参数或缩小其适用范围。这是由拼合关系导出的调查逻辑；将它落实为某次历史调查的因果解释，还须有候选组合、调查安排与修订结果之间的连续记录。

失败条件也由此变得清楚。若组合依赖任意修改轮廓，外部岩带年代持续冲突，或者所谓预测全部是在结果出现以后重新表述，古大陆身份就没有得到充分支持。反过来，即使总体组合得到支持，也不意味着每一处边缘都曾按刚体方式相接。区域变形可以成为模型的一部分，但须由独立材料规定，不能成为无限自由度。

几何反演确定候选位置，地质相关检查历史连续性，两者共同规定古大陆组合的证据强度。SDE 把这组关系接到后续操作：哪一处不吻合会改变组合，哪一种新比较能区分其原因。古大陆因此获得了可调查的身份，但尚未回答移动如何发生。进入魏格纳争论时，需要进一步区分大陆相对位置是否改变、移动机制是否可行，以及当时材料是否足以支持某个具体版本的漂移理论。

第 22 章 魏格纳的大陆漂移证据综合为何没有立即形成共识

22.1 三次出版节点的论证差异

大陆漂移史常被压缩成一句戏剧性叙事：魏格纳发现大陆能够拼合，同代人拒绝他，后来科学证明他正确。这句话既保留了一项重要转折，也抹去了转折的实际结构。它把不同年份的文本并成同一个主张，把不同学科的反对理由并成同一种偏见，又把后来的板块理论当成早期漂移理论的完整复制。若这样叙述，历史只剩先驱与阻力，科学判断怎样改变反而无法解释。

阿尔弗雷德·魏格纳研究所的官方传记提供了三个清楚的时间锚点：1912 年的公开报告，1915 年《大陆与海洋的起源》初版，以及 1929 年的第四版。[B20] 这些节点首先证明思想经历了持续公开表达和修订，而不自动证明每一版增加了哪些证据。论证史必须进一步把题名相同的书拆成具体版本。首次报告提出一个问题，与一部多年修订的著作对这一问题作出的综合回答，不能互相替代。

版本差异为什么重要？因为一项理论并非只有“相信大陆移动”这一句话。它至少包含移动的对象是什么，曾经怎样排列，什么时候分离，凭哪些证据确认，以及可能由什么力量推动。某一版本可能加强古生物材料，却保留有困难的驱动力说明；另一版本可能修改重建图，却没有解决测量争议。理论在这些维度上并不总是同时进步。若只记录出版次数，就无法分辨知识增长与重复传播。

1912、1915 与 1929 三个节点应分别对应当时可见的论证状态。可从这些出版与报告记录确认的是，魏格纳持续将形状、地质、古生物和古气候中的关联纳入大陆相对位移问题。[B20] 各版具体增加或撤回了什么，仍需逐段对读。版本问题会直接影响解释：若一项材料在批评之后才进入文本，它可能承担回应作用；若早已存在，则不能仅凭后来被强调，就把它当成争论新产生的证据。出版年表只有与主张及材料的年表相接，才成为思想发展的历史。

读者可以据此理解一项看似繁琐但关键的区分：一个命题的公开时间，不是它所有支持证据的获得时间。地图上的地质对应可能来自更早的区域调查；古生物鉴定可能后来修正；用于现代解释的海底资料又可能远晚于作者的生命年代。把这些放进同一段论证并非错误，只要明确它们是在重建今日证据链，而不是复述当时科学家已经掌握的证据包。

传播范围也不是简单的接受度。翻译、再版、评论与教学介绍，可以使一个假说广为人知，却不表示读者认可其所有部分。反对性评论同样会传播理论，并推动作者回答新的问题。反过来，个别研究者使用某项古地理重建，也可能只接受其工作假说地位。科学思想的公共生命由许多不同程度的采用构成，不能用“流行”或“不被接受”一词涵盖。

因此，研究迟滞首先要问迟滞发生在哪一层。若研究者承认跨洋对应值得解释，却不接受某种运动机制，争论焦点在动力层。若他们怀疑化石相关的精度，问题在材料层。若接受大陆曾位移，却认为现有重建的方向与时间不可靠，问题在历史模型层。只有把这些层次拆开，才能解释为什么某条证据加强后，共识仍不会立即出现。

这种区分也维护了魏格纳贡献的实际分量。他的意义不需要通过让全部反对者失去理性来确立。把此前被分开研究的地质关联置于大尺度位置变化的框架，已经改变了可问的问题；承认这一点，与承认其框架内部仍含未解决甚至错误的成分，并不矛盾。一个有创造性的理论可以同时具有深刻的问题重组和严重的机制缺口。

◆ 22.2 支持位移的证据组合

漂移论证的力量来自组合，但“组合”有两种不同含义。一种是把许多材料排列在一起，造成数量上的充足感；另一种是让不同材料对同一重建提出相互约束的要求。只有后一种能显著提高解释力。如果化石、岩带和气候资料都需要完全不同的大陆位置，材料越多反而越暴露问题。真正的收敛，是它们在各自误差和适用范围内，允许一个相对一致的历史方案。

以南方大陆的古气候问题为例，今天相距遥远的地区保存了冰川作用的地质记录。现代位置不是古代位置，现代气候也不是古代气候，因此不能只说“这里今天温暖，所以过去不可能有冰川”。地球整体气候可能变化，高山也能形成冰川。论证需要更细的条件：记录的时代是否可对比，沉积物是否确由冰川作用形成，擦痕或搬运方向能否约束冰体运动，以及重建后这些方向是否获得更有组织的分布。

这一分析展示了证据的层级。某地存在古冰川沉积，是一个区域判断；不同地区记录属于可比较时段，是年代相关判断；它们可以放进共同冰盖或冰川系统，是空间重建判断；这一系统支持大陆位移，则是更高层的模型比较。每提升一层，都增加了新的前提。若年代相关仍很粗，不能把它叙述为同一瞬间的全球冰盖；若搬运方向受到后来构造旋转影响，也需要先校正再比较。

大陆重组的解释收益，在于有时可以用较少的特殊安排，使这些区域记录形成较连贯的古地理关系。但“较少”并不等于只有一个参数。旋转角度、裂解时间、陆块内部变形以及不同冰期事件仍需规定。最强的漂移论证不会隐瞒这些自由度，而会检查它们是否也得到地质证据支持。若所有古气候矛盾都靠各地独立的高度变化解释，固定位置方案可能承担很大的额外负担；若重组同样必须无限改动各大陆，漂移方案也不会自然胜出。

古生物分布提供另一类约束。陆生与淡水生物的跨洋对应，可以使遥远大陆曾相近的想法更有吸引力。然而，不同生物具有不同传播能力，生物分类也有不同精度。一个广布属名和一种生态要求严格的动物，并不提供等量证据。分布区还取决于保存环境与调查密度。地图上没有化石记录，既可能表示生物从未存在，也可能表示合适岩层没有保存或尚未被发现。把所有空白都当成真实缺席，会使任何重建显得过于精确。

跨洋构造和地层对应则要求同时比较空间方向与时间次序。同类山系的名字不构成连续性，沿相似方向延伸也不够。需要检查形成阶段、变形组合与两侧地层记录。这里最有说服力的情况，是某一重建同时减少多个性质不同的错位：岩带接续更加自然，形成时代相容，古生物分布不再需要多条特殊通路。这样的汇合不是算术意义的证据相加，而是替代模型越来越难同时满足全部条件。

仍须警惕共同前提。若古生物相关被用来判定地层时代，随后又以这些时代对应支持化石分布，两者并非完全独立。如果地质学家先接受漂移重建，才把不确定岩带归入同一系列，那么后来统计的“许多一致岩带”也包含解释依赖。这不意味着相关证据无效，而意味着需要用独立年代、明确鉴定标准或保留样品重新审查依赖链。证据的可信度来自误差如何受到限制，不能来自学科名称的多样性。

可以把竞争解释放在同一张思想上的工作台。陆桥能够回答迁移问题，却需解释桥梁的地壳性质与消失；全球气候变化能够解释某些温度迹象，却未必解决方向性冰川记录；区域升降能够解释海陆交替，却未必重建跨洋构造连续。漂移可以同时处理若干问题，但

要承担移动介质与驱动力的责任。公平比较不是要求每个理论在诞生之日解决全部地球科学，而是防止一方只享受解释收益，另一方却独自承担所有未解问题。

在这一意义上，早期漂移对象具有有条件的身份。它并非纯粹想象，也尚未成为具有完整运动学和动力学约束的全球系统。它能够组织新的比较，提出新的地理关系，并使一些旧问题获得较统一的回答；但它对“如何移动”的说明仍限制了共同体能把这种重建当作怎样的知识。理解这种中间状态，比在正确与错误之间立即二选一，更接近真实科研。

古气候证据还必须区分温度指标与纬度指标。煤层、蒸发岩、冰川沉积等记录形成条件，但形成条件通常由温度、水分、地形与保存共同决定。煤层不能简单译成某个唯一纬度，蒸发岩也不能仅凭出现就给出精确干旱带。大陆重建的收益在于多种指标在空间组织上同时变得较合理，而不是把每一种沉积物变成无误差的纬度计。若同一地点的不同指标指向不一致，应先检查时代和环境，而不是挑选最符合重建的一项。

冰川运动方向尤其提供不同于有无分布的约束。仅知道若干地点存在冰川痕迹，可以由多种冰体安排解释；若知道搬运或擦痕方向，就能检查这些方向在重建后是否具有可能的汇散关系。方向信息具有更强判别力，也对构造校正提出更高要求：岩块后来旋转，记录方向会随之改变。一个看似精确的箭头，必须经过层位、形变与原始姿态判断，才能进入古地理比较。

同样，物种分布并非只比较地图上两个点。生态组合可以提供比单一名称更多限制。若一组具有相近环境需求的生物与相容沉积相共同出现，联系更具体；若只共享一种传播能力很强的生物，解释空间则更大。缺失资料还可能造成不对称：一侧保存丰富，另一侧仅有零星标本。好的重建不以资料丰富一侧的精细分类强迫资料贫乏一侧作出同等精确的回答。

这些差异说明证据综合不是消除各学科方法，而是使方法保留特长后进入共同问题。古生物学负责鉴定与生态，地层学负责先后与相关，构造研究负责形变与方向，古气候研究负责环境解释。大陆组合不能代替任何一项工作，却能把它们的结果放到一个会互相冲突或支持的场所。魏格纳问题的创造性，正在这种跨材料的相互要求之中。

22.3 动力缺口与合理反对

大陆从一个位置移到另一个位置，属于运动学主张；推动它的力、承受运动的介质及其响应，属于动力学与物性问题。两类研究可以暂时分工，却最终要在同一地球上相容。早期漂移论的困难，在于所提出的驱动方式尚不足以满足岩石强度与地球运动的物理约束。USGS 的历史概述也将缺乏令人信服的机制列为争论的重要因素。[B01] 这项困难针对的是过程可行性，与大陆轮廓能否相接属于不同证据层面。

“大陆穿过洋底”与“大陆和部分洋底属于同一运动单元”不是同一种对象。若大陆被设想为相对独立的块体，在较坚固的海洋底部中推进，就必须解释它怎样克服阻力，以及为什么没有形成相应的巨大破坏。后来的板块体系改变了运动单位和边界方式，不能被用来宣布早期这一困难根本不存在。相反，正因为对象改写，某些原先难以处理的问题才获得新的表述。

一个简单的教学估算可以说明这种反对。若平均剪切阻力为 τ，受力接触面积为 A，总阻力的量级约为 τA；以速度 U 维持相对运动，所需功率的量级便约为 τAU。这不是历史参数的复算，而是把机制放进力与能量的共同约束：方向合适还不够，驱动力须与阻力同量级，持续运动又须有相应能量供应。减小接触面积、降低有效阻力或改变运动单位，都会改变估算；每一种改变也必须说明它在地球中的物理依据。

反过来，依据短时间实验观察到的岩石坚硬，直接推断它在地质时间内绝不发生流动，同样可能扩大了结论的适用范围。材料对瞬时载荷与长期载荷可以表现不同；温度、压力、晶体结构和缺陷也会影响变形。严谨反对需要说明采用何种物性模型、对应何种深度与时间尺度。早期争论中存在真实的物理困难，但后来对地球内部变形的认识也确实改变了“什么运动可行”的判据。

这使合理反对具有双重责任。一方面，不能因为一个假说解释了许多地图关联，就豁免它的力学问题；另一方面，不能把尚未完整掌握的深部物性当成永远封闭的禁令。最有生产性的批评，是把“不可接受”具体化为条件：若能提出与物性相容的运动介质、说明能量来源、找到与运动方向一致的独立记录，那么评价可以改变。这样的批评留下了知识增长的接口。

测量争议还具有不同性质。尝试通过重复定位寻找大陆位移，需要位置精度和时间间隔足以分辨预期变化。若测量误差与目标位移相当甚至更大，不能把正值报告当作已测得

漂移，也不能把没有检出当作大陆永不移动。观察能力决定了该项证据的检验力。现代卫星大地测量能够做的事情，不能追授给早期测量体系；现代直接速度也不能替代对古老位移历史的独立重建。

学科评判差异则应落实为任务差异。地质相关研究可能重视一个模型能否统一许多区域关系；物理研究可能要求力学机制与数量级自洽。两者并不是天生一方开放、一方保守，而是在检验不同责任。如果没有几何和地质解释收益，再完备的驱动模型也不知道要解释何种运动；如果没有力学可行性，大尺度重建也可能停留在多种历史图景之一。真正的综合需要让这些责任相互规定。

这里的特征纠缠，可以直接写成地质与力学之间的双向限制。跨洋对应限定了所需位移的方向和规模；介质物性限定这种位移能够通过哪些方式实现；运动单位一旦改变，又会影响哪些地质对应能够共同解释。迟滞便可能来自这些约束尚未相容。例如，一组重建要求保持大陆大体完整，而某个机制所需应力会先使它严重破坏，两者就不能作为同一答案保留。问题在于改写介质或运动单位能否同时保住地质收益与物理可行性。

动力讨论也需要分清力、应力与位移。总力说明作用的总体规模，应力说明力分布在多大面积上，位移则是材料在一定时间内的响应。相同总力作用于不同接触面积，可能导致不同变形；相同应力维持不同时间，也可能产生不同累积位移。只比较一个力的大小而不说明作用位置与时间，或者只引用大陆移动距离而不说明速度，都无法完成物理检验。

介质模型改变时，问题形式也会改变。一个完全弹性的模型把变形与即时载荷联系，黏性模型则把变形速率与应力联系，具有多种行为的材料还会随时间尺度转换。这里不把后来的流变概念直接归到早期争论双方，而用它们解释为什么“岩石很硬”并不是完整结论。对地质时间的运动，最关键的是适当物性与适当时间窗口是否匹配。

与此相对，支持者也不能只说“时间足够长，什么都能发生”。如果某种机制的方向不对，延长时间只会累积错误方向的运动；如果能量来源不足，长时间也不能免除能量收支；如果材料响应会先破坏所假设的完整大陆，模型就需要改写运动单位。长期性是地学的重要条件，却不是取消物理约束的万能钥匙。

由此可见，地质综合与物理批评并不构成创新和保守的天然对立。它们更像两组必须相遇的边界条件：一组规定需要解释的历史关系，另一组规定能够发生的过程类型。科学

转折可能在两组条件冲突时出现，因为冲突迫使研究者改变对象，而不只是调整原有参数。后来从大陆转向板块，就是这种对象重组的一项核心内容。

对共识形成还可以作一项反事实检查：如果后来没有获得海洋磁性、地震与年龄资料，仅有大陆轮廓和区域相关，早期争论是否已经能够解决？不能因为今天知道答案，就假设当时材料具有同样区分力。后来的证据改变了可接受解释的范围，这是实际增长，而不是人们终于愿意承认一件早已完整证明的事。

同样，若某种早期机制被后来的物理分析否定，它的失败不会抹掉大陆位移问题的价值，却会改变应当怎样继承。我们可以继承一项需要解释的现象、一个有效比较或一套调查材料，而不继承全部因果安排。将这些层次拆开，是思想史避免胜利者倒写历史的重要方法。

创新的评价因而包含两个尺度：它改变了哪些问题，又完成了其中多少解释。魏格纳的综合使跨洋地质关系进入位置变化的共同问题；机制缺口则限制了这一综合能够获得的确定性。把二者分开，才看得见后来海洋资料的作用：它们既支持大陆相对位移，又重新划定运动单位，使早期若干难题不再以原来的形式出现。

年代上的迟滞与论证上的停滞因此不能等同。一个问题多年未获共识，仍可能在材料、对象和机制上发生重要改变；只有追查这些改变，才能判断它是在积累条件，还是原地重复。

◆ 22.4 迟滞怎样反向改写假说

一个假说在争论中持续存在，并不等于它原封不动地等待胜利。争论会改变支持者愿意保留的内容、继续搜集的资料和能够接受的解释形式。大陆相对位移可以作为问题延续，而某一种驱动力说明被放弃；某一古大陆组合可以继续作为工作假说，而其精确时间和局部形状被修改。这种分解，使理论的延续具有具体单位，而不是由人物姓名提供同一性。

文本修订能够显示理论如何回应困难，但需要按命题对照。增添一组化石不一定增强动力解释，修改驱动力也不一定改善古地理位置；两者应分别评价。若一版撤去某种机制，却继续保留由该机制导出的移动方向，就要检查这一方向是否另有支持。若新增材料

只进入附录而没有改变原来的结论强度，也不能仅按篇幅增长判断理论已获加强。版本对读的对象，是支持关系如何变化，而非著作是否越来越厚。

后期海洋证据之所以关键，在于它没有只是给大陆轮廓增加几张支持图片。海岭、磁异常、洋壳年龄和震源机制把海洋底部纳入了运动对象。大陆不再是唯一需要移动的实体，洋壳也不再只是大陆移动经过的被动背景。新的边界类型和运动关系改变了早期问题的条件。这既是漂移问题的继承，也是其对象结构的改写。

因此，“魏格纳后来被证明正确”必须限定含义。大陆相对位置随地质时间显著改变，这一方向获得后来的强有力支持；但它不意味着其所有古地理细节、所有机制候选和所有论证都同时正确。科学史中的继承通常是选择性的：一些问题得到保留，一些对象重新划分，一些证据经过重解释，一些机制被放弃。承认选择性，才不会把后来科学发展写成对早期文本的简单注释。

共同体采纳也有自身时间。一个模型可以先在若干研究群体中成为有效工具，再通过更多调查、计算和教学稳定下来。不同地区和学科的采用节奏不会完全一致。若缺少期刊、课程与研究计划的连续材料，就不能给出整个共同体在某一天集体转换的故事。所谓共识，是多个责任逐渐获得可共享答案的过程，而不是接受人数跨过一个神秘门槛。

还可以区分三种支持者。一种把漂移当作已经定论的整体答案，倾向于保留全部原始论述；一种把它作为解释若干地质对应的工作假说，愿意暂缓机制；还有一种主要保留它提出的问题，甚至不接受其具体重建。历史资料中的“支持”若不区分这些层次，就会把不同研究立场混为同一阵营。反对者同样可能只反对某种机制而不否认位置变化的可能性。

这种多层结构也解释了理论为何能够在没有统一共识时继续积累。不同研究者承担不同部分责任，彼此并不必然认可整套体系。一个人的反例可以成为另一人的机制问题，一份区域调查也可能后来被放进完全不同的全球框架。知识延续不是同一结论在更多头脑中复制，而是问题、技能与材料在不同研究路径之间重新组合。

本章的结论因而保留张力。漂移思想的创造性，在于重新组织了古地理问题，并让原本分离的证据进入共同检验；早期未形成广泛共识，则具有机制、资料和判断层面的实际理由。解释迟滞时既不取消创新，也不取消批评。下一章将用同样标准重建固定大陆与地槽解释的强版本：它们究竟解释了什么，又在什么尺度上失去继续承担全球责任的资格。

第 23 章 固定大陆学说
局部有效为何不能保证全球成立

23.1 先重建旧模型的强版本

新理论获胜之后，旧理论常被改写成一个简单错误：人们曾经以为大陆永远不动。这样的句子有助于记忆，却不足以解释地球科学如何发展。大陆位置固定、地球冷却收缩、地槽沉积与造山，是有关联却不相同的主张。一个研究者可以在区域研究中采用地槽概念，而没有给出完整的全球构造模型；另一个人可以承认显著水平缩短，却不接受大陆跨洋位移。把它们合成一个毫无内部差别的固定论，会使真实争论失去形状。

固定大陆首先涉及大陆与洋盆长期位置关系的判断。收缩说试图以地球总体尺寸或体积变化解释地壳变形。地槽解释则围绕长条状沉降区域中厚层沉积、后来的褶皱与隆升建立历史联系。三者可以组合，却不互为逻辑同义词。证明某种收缩机制不足，不等于证明区域剖面中的沉积厚度不存在；证明大陆发生大尺度位移，也不等于所有以旧术语组织的区域调查都应抛弃。

强版本的旧模型，应从它必须回答的实际问题出发。设想一条横穿古老山带的剖面：一侧地层较薄、变形较弱，另一侧沉积序列厚、褶皱复杂，并有变质或侵入作用。研究者要解释为什么沉积曾长期积累在某一区域，后来为什么又变成山地。即使没有全球板块理论，这也是一个真实而困难的问题。以沉降、堆积、变形和隆升构成历史阶段，能够使许多分散观察得到组织。

这一剖面是用于辨析模型的教学重构。关键在于三类观察所约束的内容不同：厚层沉积表明容纳空间与物质供给曾经配合，褶皱记录缩短与变形，变质和侵入则增加温压与热历史信息。把它们排成一个先后序列，可以构成区域历史；要进一步说明这些过程为何发生在此处，还需恢复同时期邻区的关系。单条剖面无法自行给出全球动力来源，也不能据此断言某个历史学派已经完成了这些检验。

旧模型的强处，在于它可以为野外工作提供可执行的问题。哪里应寻找沉积中心，哪里可能出现相变，变形带与较稳定区域怎样过渡，哪些层段应重点测厚？一个模型能提出这些问题，就不只是任意故事。它可能在局部层次上产生有价值的预测，即使其全球机制后来被替换。思想史必须解释这种生产力，否则无法理解为什么训练有素的研究者愿意长期使用它。

但区域生产力不自动赋予全球充分性。一个沉降盆地的纵向演化，可以在多种全球构造背景下发生。相同的厚层沉积并不能独自判断大陆是否跨洋位移，也不能唯一决定盆地的驱动力。旧模型若从“此处曾沉降并积累沉积”跨越到“全球大陆位置长期固定”，就增加了一个需要独立检验的命题。问题并不在于它解释得太少，而在于是否把局部成功当成了无需再检验的全球许可证。

◆ 23.2 局部成功的可保留部分

理论更替时，至少要把四类内容拆开。第一类是可追溯的记录，例如某地点的岩性、层序和标本。第二类是经过解释的几何关系，例如断层重复、褶皱形态与层厚恢复。第三类是历史过程判断，例如沉降先于某次变形。第四类是全局机制，例如由何种大尺度运动造成这些现象。新理论对这四类内容的影响并不相同，不能一键保留，也不能一键删除。

原始记录也不是完全不受理论影响。研究者选择哪一条剖面、如何划分层组、在何处采样，都可能受到预期支配。但理论影响并不意味着记录可以随意改变。保存的标本、测量方法和位置资料，使后来者能够重新检查。所谓旧知识可继承，最可靠的含义是它仍拥有可复核的约束，而不是它曾经被写成“事实”便永远正确。

层厚是一个很好的例子。在构造变形较强地区，露头上看到的累计厚度可能受到断层重复或褶皱影响。若旧研究把重复层段当成连续沉积，就会夸大原始沉积厚度；若已经做了可靠恢复，这一结果则可能被新理论继续使用。板块构造不会凭名称自动改正层厚，仍需具体剖面工作。新理论的优势必须落实为更合理的结构解释与更多独立约束。

同样，地层先后关系与沉积环境可以在更换全球模型后继续有效。某层由浅海转向深水沉积，并不因为解释从地槽转为特定板块环境便重新沉积一次；改变的是这种转变被放进何种空间结构与动力过程。这里同时保留自然事件的稳定性和科学对象的重组。新世界

的发生不是古老岩石被重新创造，而是其属性之间出现了新的、能够指导操作的共同规定。

还有一些所谓事实本身夹带了机制。例如把某地称为“永久稳定大陆边缘”，已经比岩性记录多作了时间与运动判断。新资料可能保留它在某一阶段相对稳定的结论，却否定“永久”这一扩展。继承不是把旧术语换成新术语，而是拆解其承诺：哪些被资料直接限制，哪些属于可撤回的模型部分。这样的分析避免把理论更替理解为无条件累积，也避免把它理解为彻底断裂。

研究训练本身同样能够继承。识别层理、恢复褶皱、区分侵入接触与沉积接触，需要长期形成的辨认能力。新理论常依赖这些能力来获取自己的关键证据。若认为旧框架中的研究者只是在错误中浪费时间，就无法解释新体系凭何拥有高质量材料。理论创造与技能继承并行，构成科学世界发生的重要条件。

但继承必须有适用域。一个区域模型在几十公里的尺度上成功描述沉积与变形，不表示它能解释整个洋盆。将其有效性限制在明确尺度和问题范围内，并不是轻视旧成果，而是给予成果真实的责任边界。知识之所以能够继续使用，是因为它知道自己在哪些条件下仍成立，而不是因为它曾经属于获胜或失败的一方。

区域剖面的恢复还显示了事实与机制之间的细致层级。地表观察到一组倾斜岩层，研究者要先判断它们原先是否近于水平，再判断褶皱是否伴随断层，再估计缩短量。每一步都增加假设，也各有检验。若恢复后的层段长度与相邻区域不相容，可能需要补入遗漏构造；但补入的断层必须留下可查踪迹，不能只是让图形闭合的方便线条。

厚层沉积与后期山地的联系，并不唯一指定其成因。在不同构造条件下，都可能先形成容纳沉积的空间，再经历压缩或抬升。旧模型若成功识别这一次序，成果可以保留；若把次序直接等同某一全球驱动，就需要进一步证据。板块解释所增加的，是把区域阶段与边界运动及邻区关系相联系，而不是否认沉降和褶皱这些过程本身。

沉积中心的位置也会迁移。把整个厚层序列压成一个长期固定槽体，可能遗漏不同阶段的盆地边界变化。重新按时代拆分沉积相和厚度，可以检验所谓单一地槽是否实际上包含多种过程。对象的变化发生在时间分辨率提高之后：原来一个整体名称覆盖的记录，被分为具有不同空间关系的多个阶段。这种改写比简单宣布旧名过时更有解释力。

23.3 海洋证据改变比较尺度

海洋资料扩大了模型必须同时解释的空间关系。陆地山带和沉积盆地之外，海岭、海沟、洋壳年龄及地震分布要求进入同一全球图景。USGS 对板块理论形成的概述，强调海底地形、磁场反转、海底扩张与地震火山分布等方面的汇合。[B02] 这些记录的结合之所以重要，是因为区域模型原可分别设定的条件开始彼此牵制：某处生成多少新海底，会涉及别处怎样容纳面积；某种地形若被解释为活动边界，又要接受地震位置与运动方向的检查。

年轻洋壳的广泛分布提出了年龄结构问题。若海洋底部长期存在，为何可见记录与古老大陆具有显著差异？一个固定洋底模型可以尝试通过再造、覆盖或记录缺失解释，但这些机制必须有范围、速率与物质依据。只说“海底经历了复杂变化”不足以回答年龄随空间分布的结构。一个解释若只处理总体年轻，却不能解释离海岭距离与年龄的关系，就尚未承担全部约束。

海岭的连贯分布又增加了形态与热状态的责任。若某种机制形成一条海底高地，它还需解释为什么磁异常、地震或其他记录与其位置具有系统关系。这里不能把多幅地图视觉叠合便当成全部独立证明，因为不同图件可能共享定位资料，某些边界也由其中一类观测预先画出。即便如此，只要误差和依赖得到控制，模型必须解释的联合结构仍远超过一条区域剖面的要求。

地震带则把过程位置加入约束。一个山带的形态记录过去，地震活动记录某些正在发生的变形。两者的时间窗口不同。旧模型可能成功描述一个已经停止活动的造山带，却难以从其全球框架预测海沟附近震源的空间排列。新理论若能用边界类型同时组织地形、运动与地震分布，就获得了跨尺度解释的优势。但这种优势必须由实际预测与检验展示，不能只凭术语统一。

公平比较要求给予旧模型合理修订的机会。例如，它可以承认局部海底升降、提出洋壳更新机制，或者允许某些水平运动。问题在于这些修订能否互相兼容，是否有独立证据，以及是否还能产生有风险的预测。一个理论并不会因为增加辅助假设就自动失去科学性；任何复杂地学模型都需要边界条件与辅助机制。真正的差别在于辅助假设受到多少约束，是否只在结果不合时被临时增加。

可以用一个教学对照说明。假设两种模型都能解释某地沉积很厚。模型甲需要独立规定沉降，再另外规定海岭出现、地震带分布与洋壳年轻；模型乙用一个运动边界体系联系这些现象，同时仍允许沉积供给的区域变化。若模型乙的速度和边界由部分数据确定后，能够预测未参与拟合的另一类记录，它的统一就具有检验价值。若它只在事后为每种现象附上一个板块名词，优势依然没有证明。

全球尺度并不总比局部尺度更高贵。为了研究一条具体断层或盆地的沉积史，过于粗糙的全球板块划分可能不足以解释关键细节。新体系的力量在于提供可兼容的上层约束，而不是取消区域工作。板块速度可以限制总体边界运动，却不能直接给出每个断层如何分担变形；海洋扩张可以解释基底年龄结构，却不能独自决定一处盆地的沉积厚度。

因此，旧模型失去全球资格的判据不是“后来人不再使用它”，而是它无法以受约束的方式同时承担新增资料。新模型取得资格也不是因为它能够画出统一图景，而是因为统一关系能够被不同记录检查。科学世界的转折发生在比较责任的改变：原本分开的区域现象，被要求进入共同的空间、时间与运动账本。

地球冷却收缩的解释还面对一个总量问题。即使收缩可以产生压缩，也需要比较它在给定时间内能提供多大表面缩短，与不同山带恢复所要求的运动是否相容。若局部山带的缩短可以解释，却无法同时容纳海洋生成和不同方向的运动，全球模型便受到限制。这类比较说明模型不能只在一条剖面上成功，还要与其他地区共享一个可行的地球历史。

对固定洋底方案，也可提出明确的差别预期。若海底更新主要是原地垂向过程，那么年龄与地形、磁性结构的对应应由这些过程具体解释；若海底沿生成中心迁移，则不同位置的历史可以由共同运动关系联系。两者都可能产生年轻岩石，但对年龄序列的空间组织承担不同责任。区分模型的证据往往不是某个孤立现象，而是多个现象之间的排列方式。

这也是为什么局部成功不能机械累加成全球真理。十条剖面都能被某种区域叙事描述，仍可能需要十套彼此无关的条件；一个全球模型则要求它们在同一运动安排下相容。统一性只有在减少独立自由安排并保留预测风险时才有价值。如果只是给十套故事加上同一标题，解释并没有因此变强。

区域细节也能检验全球综合。例如，全球模型给出两侧总体会聚，区域剖面却显示缩短集中在少数带内，两者未必冲突；总体速度还须由不同断层和变形区分担。若恢复出的

各区缩短方向与总运动长期不相容，才需要检查板块划分、时间窗口或剖面恢复。区域知识由此既能在全球框架中保留，也能反过来约束框架；判断的关键是不同尺度之间能否写出相容关系。

旧知识的继承由此可以落实到剖面：哪些层段仍按原次序使用，哪些厚度需要消除构造重复，哪些接触关系应重新鉴定，哪些过程判断只能保留在局部。这样的逐项处理，说明了“被取代”的确切含义。旧模型失去全球解释地位，与它组织的每一条观察是否仍有效，是需要分别回答的问题。

同一套调查成果也可能服务几个竞争模型。材料能够被重新使用，说明它保存了超出原解释的约束；重新使用得越充分，越需要明确哪些原有判断仍然成立。

◆ 23.4 从被取代到适用域重定

理论败北有几种不同结果。有些命题被直接否定，例如某个永久性边界假设与可靠位移记录冲突；有些成为局部近似，例如某区域在限定时间内可视为稳定；有些则从机制解释降为描述性分类。把这几种结果混成“旧理论错误”，会使我们既无法理解知识继承，也无法判断当代模型将来可能怎样变化。

地槽术语的历史命运尤其提醒我们，名称的延续与对象的延续并不一致。一个旧词可能继续出现，但其机制含义已经改变；一个新词可能替代旧词，却保留大部分区域资料。真正需要追踪的是定义、使用条件和操作后果，而不是只统计词频。若同一剖面在新框架中提出了不同采样问题，或者某些原本被视为共同成因的沉积体被重新区分，就发生了更实质的对象重组。

在本章的 SDE 分析里，显露 S 首先是具有明确适用范围的区域模型。差异序列 D 从层序、厚度和构造比较，扩展到与海洋全球资料的对读。特征纠缠 E 具体表现为剖面分辨率与机制尺度互相限制：精细剖面能够排除某些局部结构解释，却未必区分全球模型；全球运动约束又能反过来改变区域剖面中哪些关联值得重点复核。新增世界不是把旧数据统统清零，而是重组其共同成立的条件。

如果新旧解释仅仅调整一个参数，其他对象定义、比较方式和调查任务均不改变，本书就把它称为模型修正，不强行称作新世界。若改变了运动单位、边界身份和可检验关

系，才需要更强的发生性语言。保持这个门槛，能够避免理论术语在每次小改动中被滥用。

判断地槽解释的历史效力，需要区分它提供的区域次序和对该次序的成因归属。沉积、变形与隆升的先后可能被保留，原先共同成因的假设却可能拆开：不同阶段的沉积中心未必重合，后期缩短也未必沿早期盆地边缘发生。将层序按时代重新定位，便能检查一个长期槽体的名称是否遮蔽了多段历史。这是区域资料在新框架下可重新提出的问题；具体学派在何处取得过预测成功，仍应由原始调查记录判断。

旧模型被重新限定以后，教学也应改变叙述方式。学生需要知道哪些观察出自旧框架的长期调查，哪些机制后来被替换，哪些术语仍作为描述保留。这样才能理解理论不是把一堆纯事实外加一个解释，而是长期组织调查与判断的工作系统。新体系的建立既需要突破，也依靠可以继承的测量、档案和技能。

接下来需要解释的，正是使比较尺度发生变化的物质条件：海洋长期难以连续观察，海底资料为什么逐渐能够按位置、时间和精度进入共同争论？答案不能停留在“仪器进步”四个字，而必须进入测深、定位、磁测和航次组织怎样相互配合的实际链条。

第 24 章 海洋观测基础设施
新证据为什么终于能够进入争论

24.1 测深与定位的共同限制

一条海底曲线看上去像山脉的侧面，但仪器最初得到的并不是山脉。声学测深记录声信号发出与回波返回之间的时间差，再依据传播速度换算距离。最简单的关系是水深约等于声速乘往返时间的一半。这个关系已经包含条件：回波来自何处，传播路径是否可近似处理，海水声速如何估计，以及船的位置在这一时刻在哪里。离开这些条件，纸上的起伏尚不能成为可比较的海底地形。

声速并非在所有水体中严格相同。温度、盐度与压力会影响传播，分层还可能使声线路径弯曲。教学中采用一个平均声速，适合解释原理；研究中则要根据目标精度判断这种近似能否接受。若声速系统性估高，换算深度也会偏大。若变化随区域而异，误差甚至可能伪装成空间结构。因此，“测到一个回波”与“获得具有已知误差的深度”之间，仍有校准与解释工作。

定位误差会进一步改变地形身份。假设相邻两次航行都记录到一个深谷，只有知道两条航迹在哪里，才能判断它们是同一结构的重复交叉，还是不同位置的两个凹陷。如果位置误差与谷的宽度相近，连接就会变得不确定。船时、记录仪时间与定位时间如果错配，深度点也会被安放到错误地点。海底形态的成立依赖空间与时间配准，不能把测深仪性能单独视为证据质量。

还要区分沿航线分辨率与航线之间的覆盖。连续记录能够在一条线上显示丰富细节，却不会自动填满两条航线之间的海域。读者看到一幅完整地图，容易把纸面的连续性误认为观测的连续性。实际上，海底制图长期必须从有限剖面推断其间结构。这样的推断可以有充分地质依据，但必须保留测线位置，使后来者知道哪些地方被直接约束，哪些地方依赖插补。

要把两条航线放到一起，首先需恢复它们各自的测量基准。航次时间、定位方式、声速处理和回波筛选决定曲线上的一个点如何取得地理意义。相似曲线若采用不同参考面，

整体深度差可能来自换算；看似不同的曲线若平滑程度不同，小尺度起伏差也可能来自处理。档案保存这些条件，便使后来者能够逐步统一可比部分，同时保留尚不能比较的区段。

这类问题说明基础设施不是科学对象之外的背景。定位精度规定了哪些起伏可以被认作同一地貌；地貌的预期宽度又反过来规定所需定位精度与测线间距。两者相互限定，构成特征纠缠的具体部分。若目标只是识别一个宽阔海盆，较粗资料可能已经足够；若要区分狭窄裂谷与局部断陷，同样的资料便可能失效。仪器好坏必须相对于科学任务判断。

测深的垂直分辨率还与回波识别有关。在崎岖海底，最先返回的回波可能来自侧方较近的高点，而非正下方海底；坡面与反射强度也会影响记录。于是一个尖锐起伏可能包含声束几何的作用。解释曲线时需要知道仪器的观测范围，而不能把每个回波都当成理想垂线。技术改进的意义之一，是逐步限制这种几何歧义，使起伏的位置和形态获得更可靠身份。

水深参考面同样不是无关细节。潮位、船体吃水和换能器位置会影响读数与所报告深度之间的换算。对于宏观海盆，这些误差可能较小；对于精细地貌或重复测量，它们可能不可忽略。同一个误差项是否重要，取决于问题尺度。科学仪器的历史不是所有误差同时归零，而是研究任务不断迫使某些原本可忽略的项进入显式处理。

时间同步可以用一个简单教学例子说明。船以近似恒速前进，若记录时间与定位时间相差一个固定间隔，所有测点就沿航向偏移一段距离。正向与反向重复同一航线时，这种偏移可能表现为同一结构位置不一致。通过交叉与往返比较，研究者能把部分系统误差从地形差异中分离。这样的设计并未增加一个新传感器，却提高了证据的自我检查能力。

◆ 24.2 磁测重测与信号可比性

磁力仪提供另一种曲线。它测量某种磁场量随位置和时间的变化，不能直接拍摄岩石内部的磁化排列。地球主磁场、外部场的时间变化、船体干扰、仪器漂移以及岩石磁性共同进入记录。为了识别海底地壳异常，需要规定参考场并处理其他贡献。若不同航次采用不同基准，其曲线的零点差异便不能直接解释为海底变化。

沿航线的时间变化与空间变化尤其容易混淆。船不断前进，仪器在不同时间经过不同地点。若外部磁场在航行期间变化，记录上就可能出现看似与地理位置有关的起伏。重复

测线、交叉航线及参考观测能够帮助识别这种混合，但不能简单认为所有不重复部分都是仪器误差。海面高度、航向和实际轨迹也可能改变测到的异常形态，需要逐项检查。

交叉点是重要的质量入口。两条测线在近似相同地点相交，其结果若差异明显，研究者必须解释差异来自时间、基准、定位还是处理方式。只有经过适当协调，才可以把多条线拼成一幅异常图。协调也存在风险：如果为了获得平滑地图而强迫所有交叉点一致，可能抹去真实变化或制造长波趋势。因此，改正量应保存，原始记录也应能够回溯。

测线方向影响空间采样。沿一条狭长异常平行航行，可能长时间停留在同一磁性区域；横穿它，才会得到清楚的交替。若研究者只看到一条曲线，无法立即判断异常是孤立斑块还是延伸条带。不同方向的测线不仅增加资料量，还增加了判别形态的能力。由此，采样设计会成为理论检验的一部分，而不再是研究开始之前的中性手续。

磁、重力、地形和地震资料即使都画在地图上，也不具有相同的空间含义。地形点主要限制海底界面的位置；磁异常综合了一定体积内磁化的贡献；重力异常同样整合不同深度密度分布；震源定位则具有方向不均匀的不确定性。把它们放到相同像素网格，可以便利显示，却不能使它们自动拥有相同分辨率。一个细密彩色网格可能只是粗资料的插值。

科学比较因此需要共同尺度，而不是共同颜色。若磁异常的有效分辨率大于地形细节，就应比较较大尺度上的对应；若地震位置误差跨越多个候选断裂，便不能把一个点硬归到其中一条线上。误差椭圆、航迹范围和处理尺度，往往比地图的视觉锐利更重要。图像越漂亮，越需要提醒读者：显示密度并不等于信息密度。

在这里，重测能够产生两种不同成果。第一种是验证仪器与处理的稳定性，增强既有记录的可重复性；第二种是改变对异常结构的判断，例如确认某种信号沿走向连续。前者改善测量可靠性，后者可能参与新对象发生。不能把每一次校准成功都称作理论突破，也不能把校准贬为与思想史无关的技术细节。新的世界需要可靠操作，而可靠操作本身有特定的发展历史。

仪器装在船体何处，也影响测得什么。磁测若受到船体磁性与电气系统干扰，需要通过布置、校正或其他措施减少影响。船改变航向时，干扰可能随几何关系变化，形成与地理结构无关的信号。记录航向不是附属行政信息，而是解释异常的一部分。若原始资料只剩一条脱离航迹的曲线，后来者就失去了排除这些解释的条件。

参考场的减除更应保持版本。一个较新的参考场模型可以改变异常长波部分，但不会改变原始仪器读数。若两篇研究使用不同参考场而报告不同异常，差异可能来自处理，而非海底。把原始观测、改正项和最终异常分别保存，可以让后续研究定位变化发生在哪一层。知识更新因此可能只是处理条件改善，也可能触及地下模型，二者应分别判断。

多仪器调查还有采样位置差。测深、拖曳磁力仪和其他装置未必在同一时刻观测船下同一点。如果拖曳距离、转弯过程或记录延迟没有处理，地图上看似同步的峰谷可能实际错位。这类错位会影响异常与海岭位置的关系，尤其在狭窄结构附近更重要。跨仪器综合首先是一项几何和时间工作，然后才是理论解释。

将数据平滑到共同尺度，是比较的一种办法，却需要保存被舍弃的信息。大尺度关系可能在平滑后更清楚，小尺度差异也可能恰是模型失败的信号。合适做法不是选择最支持理论的滤波尺度，而是比较多个合理尺度下结论是否稳定。尺度敏感性可以告诉研究者，某项对应属于区域结构，还是依赖特定处理才出现的视觉效果。

◆ 24.3 航次制度与数据可获得性

海洋调查还受船期、航路、经费与资料开放条件限制。研究船不能在每个候选地点无限停留，一条测线往往同时服务多项任务。航行安全、海况和设备维护都会改变原定方案。这些条件通过选择哪些地方得到观察、哪些地方留下空白，进入科学证据的空间分布。它们不是可以随意用来解释一切的社会背景，而是能够落实到具体航迹与记录缺口的生产条件。

美国国会图书馆对希曾—萨普档案的介绍，列出参考地图、曲线、地震资料、测深与等深线底片以及船舶航迹，并说明部分资料的保密性质与覆盖不足影响了地貌图的表达形式。[B05] 地图因而同时凝结了测量与编辑判断：船舶航迹限定观测覆盖，曲线限制地形起伏，其他资料帮助判断空白区的连接。馆藏说明揭示的是这些材料如何共同进入制图；某次航行为什么改道、某位研究者何时接受一种理论，则属于另一层历史问题。

资料保密会改变可公开表达的精度和形式。能够展示总体地貌，不等于能够公布每个原始测点；能够看到一幅解释图，不等于能够独立复算其全部细节。历史读者必须区分公开图像所传达的结构与其背后的可复核资料。一个理论可能凭公开图像进入讨论，但严格检验仍取决于更细记录是否可得。这种差别影响知识传播的速度和方式。

档案整理同样参与知识的可用性。原始记录即使保存完好，若没有地点、时间、仪器与处理说明，后来者也很难把它纳入新的比较。资料从船上记录转为可检索档案，经历了命名、分类和关联。它们可能因此获得新的研究价值，也可能在整理中失去部分原始语境。记录版本和保存来源，能够防止后来的研究者把整理后的连续性误认作当时已经完成的综合。

经费与机构合作的作用，应尽量通过可观察的中介说明。增加船时可能扩大航线覆盖，统一格式可能降低跨航次比较成本，开放资料可能让其他团队提出不同解释。这些是可以检验的机制。相反，“国家需要推动了板块理论”之类宽泛句子，如果不说明哪一种资源改变哪一条证据链，就难以区分真实因果与时代气氛。基础设施史的价值正在于把宏大叙述落到具体条件。

同一批资料也可以支持不同理论。海底高地的存在，并不在发现当日自动规定海底扩张；地震集中带也需要解释其运动关系。基础设施增加了可被模型处理的约束，但没有替研究者完成对象重建。某些新解释可能由既有资料的重新组织产生，而不依赖当时刚增加的仪器。因而，本书不把技术增长写成理论进步的单向充分原因。

航路造成的覆盖偏差还会影响对地形普遍性的判断。如果多数测线集中在交通便利或已有研究兴趣的区域，统计上出现频率较高的结构，未必代表整个海洋同样常见。研究者需要把发现分布与调查分布分开。一个区域没有报告某种异常，可能只是没有足够横穿其预期走向的测线。海洋观测的空白既有空间面积，也有方向和分辨率。

资料共享可以改变这一局面。不同航次单独看来都只是短线段，放到共同坐标后，可能显示跨区域结构；但共享前必须统一单位、基准和质量说明。合作的科学作用不是抽象的“团队力量”，而是让原本不能直接比较的记录获得可转换关系。若转换错误，规模扩大反而会扩大系统偏差。因此，协调工作既提高生成能力，也创造新的共同风险。

一份档案在多年后被重新使用，还会遇到仪器性能记录缺失的问题。后来者不能把现代设备精度赋给旧数据，也不能因旧设备较粗便宣布它毫无价值。若目标结构远大于误差，旧记录仍能提供有效约束；若要识别精细差别，就需要新测量。历史资料的价值由它能够承担的问题决定，不由年龄或技术新旧单独决定。

资料处理链还有一个与船舶设备同样重要的条件：不同研究者必须知道一个数值在何种步骤中改变。原始到时换成深度、总磁场减去参考场、航迹转换到统一坐标，每一步都

可能引入系统差异。把这些步骤保存下来，才能区分自然异常与处理异常，也才能在新方法出现时重新利用旧记录。

这使基础设施具有时间延续性。一次航行结束以后，资料仍可能因新的校准、定位或理论问题获得新的用途。但重新处理不应覆盖旧版本，否则科学知识怎样变化便无法追查。保留版本不仅方便管理，还让研究者能够检验某次新结论究竟来自新增观察，还是来自既有记录的重新解释。

新增观察与重新处理还可以互相检验。若新定位使旧曲线移到另一地点，解释变化首先来自空间配准；若定位与处理保持不变而新测线否定原有连接，受到限制的则是地貌模型。两种变化都能改写海底图，却把下一步调查引向不同环节。

一次航次的证据贡献，应按它弥补的空白来评价。沿向空白限制结构能追踪多远，横向空白限制宽度与两侧关系，时间空白限制重复观测是否稳定。新记录若只增加已有剖面的细节，未必能消除另两类不确定性；航次设计须针对尚未解决的那一项，而非以记录总量代替判别力。

◆ 24.4 补测回路与基础设施增量

当一个候选结构已经出现，观测设计便可能改变。研究者不再只是沿方便航路记录，而会选择横穿预测边界的测线，或者沿走向追踪连续性。这样的改变使科学对象反过来规定差异序列。一个海岭候选若能提出“下一条剖面应在何处出现何种形态”的预期，就比地图上的一条连线具有更强的操作身份。但要写成具体历史回写，仍需航次计划与测量结果的连续证据。

预测性补测包含清楚的时间关系：候选结构先规定位置范围，航线据此安排，新增记录再检查其连续性。若一条原定为别的任务服务的航线后来恰好穿过候选海岭，它仍可提供有价值的独立检查，但不能据此说海岭模型曾改变航行计划。两类情况在证据作用上可以接近，在思想如何影响调查的历史解释上却不同。保存计划、航迹与解释图的版本，才能区分资料发现了什么与问题选择了什么。

教学上可以设想两种测线方案。方案甲在海域中继续增加与原航路平行的线，资料数量提高，却可能仍不能区分条带与斑块。方案乙增加垂直候选走向的线，并设置若干交叉

点，数量未必更多，却能检验形态连续和曲线可比性。这个例子说明基础设施增量不等于仪器数量增量；同样的设备，通过问题驱动的重新组织，可以产生不同判别能力。

补测失败也有几种含义。预测结构没有出现，可能否定连续性；位置偏移超出允许误差，可能要求改画边界；曲线形态改变，却与磁场方向或测线几何一致，可能支持更精细的物理解释。不能把所有失败都归为“海洋复杂”，也不能遇到一次偏差便宣布整个模型崩溃。研究者需要事先说明哪些偏差属于测量和近似，哪些会触及对象身份。

S=F(D,E)在这里提出一个具体问题：怎样的记录序列，在怎样的位置与精度关系中，才能显露一个共同海底对象？D=G(S,E)追问候选对象怎样改变航线与重测；E=H(S,D)则追问重测结果是否改写校准、配准和跨仪器比较规则。这三式不替代声学、磁学或导航方程，而是帮助我们检查证据链是否遗漏了决定对象成立的环节。

设备扩充与理论形成之间，需要通过资料的实际用途建立联系。若某个解释主要来自旧测线的重新配准，新增船只就不是这一步发现的必要条件；它们仍可能为跨海区检验提供关键覆盖。单条航线支持的解释往往只在附近可查，多航次共同资料则使另一团队能够在不同海域检查同一关系。可迁移性依赖具体条件：源层深度改变时调整磁学正演，航向改变时处理几何投影，定位精度不同则扩大相应比较范围。海洋知识由地方记录走向共同体系，并非所有地点获得同一种曲线，而是各地差异能够按公开条件转换并接受检验。

海洋观测基础设施的真正贡献，是让过去难以相互校验的记录逐渐具有共同位置、时间与误差语言。它使某些问题从不可操作变为可以规划，也使某些解释承担更明确的失败风险。接下来，我们需要理解尚未获得共识的思想如何在这些新条件到来之前保存问题，又如何避免将每一种长期坚持都误认成等待胜利的先驱。

第 25 章 未被接受的思想如何存活

延续、修订与真正的失败

25.1 把拒绝理由写成可检验问题

一个思想被拒绝之后仍有人研究，可能表示它保存了尚未解决的重要问题，也可能只是表示支持者不愿面对失败。两种情况从外表看都具有坚持、少数人与长期争论，不能凭这些特征区分。思想史若只选后来成功的案例，就容易把顽强等同正确，把所有早期反对都变成成功叙事的陪衬。本章以漂移问题为入口，但判断标准必须同时适用于最终失败的路线。

第一步是把拒绝拆成能够回应的具体问题。力学反驳要求说明驱动与阻力是否相容，地质不吻合要求检查重建和年代相关，证据质量批评要求改进测量、鉴定或误差估计。这些问题不能由同一种新增材料解决。再找到一组相似化石，未必补上动力缺口；提出一个可能的内部流动，也不能消除岩带年代冲突。生产性延续取决于后续工作是否真正触及原来的困难。

拒绝也可以有不同强度。“当前证据不足”允许未来新增证据改变判断；“这个特定机制与已知物性冲突”否定的是机制版本；“任何大陆相对位移都不可能”则承担更广泛的排除责任。研究者必须识别自己受到哪一种批评，才能选择回应。若把有限批评改写成对方否定整个问题，就可能用一个局部成功宣布全面胜利，而真正困难仍未解决。

可以把争论改写成一组条件句。如果跨洋相关建立在可靠时代与相似构造史上，那么重建得到加强；如果所需力量远小于阻力而无其他机制，当前动力说明不能成立；如果新的记录在预先指定位置出现，则某项运动预期获得支持。条件句的好处在于使双方都承担改变判断的义务。拒绝者说明什么会使自己重新考虑，支持者说明什么会使自己放弃一个版本。

这种做法并不意味着每个历史问题都能立即设计实验。深时事件无法重新发生，深部过程也难以直接观察，但可以通过不同记录之间的约束提高区分力。所谓可检验，不限于

实验室重复同一事件，而在于某些可能获得的证据会使竞争解释受到不同影响。若任何结果都能被同样轻易地纳入，研究路线就缺少实际风险。

在漂移争论中，未解决的核心并不是一个抽象的“世界会不会变化”，而是大陆位置变化、海洋底部性质与可行运动方式能否共同成立。只有继续保存这组相互限制的问题，才谈得上生产性延续。把大陆漂移变成一句不能被批评的信念，反而会阻断它获得新对象和新机制的机会。

把拒绝转成任务时，还应区分决定性困难与外围困难。一个模型可能在某个小区域表现不佳，却不触及其主要对象；也可能总体图形很漂亮，却在一项关键物理条件上不成立。研究优先级应由困难对模型身份的影响决定。若支持者不断解决容易的外围问题，却始终绕开核心机制，资料增长就可能掩盖理论停滞。

同样，批评者提出的检验也要可实现。要求在无法达到的精度下直接观察古老大陆运动，不能成为当前讨论的唯一准入门槛。可以寻找间接但具有差别预测的证据，并明确其限度。合理严格性不是只接受一种理想测量，而是要求任何证据都说明误差、依赖和排除能力。这样的标准既防止轻信，也避免用不可能完成的任务冻结研究。

◆ 25.2 文本传播是否带来新增内容

魏格纳著作的不同版本提供了追踪问题延续的入口。[B20] 但文本传播不能仅按数量评价。一本书被引用很多次，可能因为它提出了难题，也可能因为它成为反面教材；一种说法反复出现，可能保留了关键问题，也可能把条件限制越删越少。思想的生命不等于名字的可见度，而在于它是否继续组织能够改正自身的工作。

阅读修订版时，可以先检查三类变化。新增证据是否真的独立于旧证据，还是把同一材料换一种表达？机制修改是否减少了原来的困难，还是只更换术语？失败例子是否被认真处理，还是在新版中悄悄消失？这三类检查不预设作者必然进步。理论在传播中可能变得更精确，也可能为了说服力而变得更模糊。

判断文本是否真正吸收批评，要看受影响的支持关系是否改变。某条跨洋岩带原本支持特定重建，后来的年代判断若使两侧不再同期，修订便有几种可能：撤回这条对应，改变旋转位置，或另找能够解释时代差的地质过程。三者对整体理论的影响不同。仅在结尾

承认“问题复杂”，同时继续把该岩带列作确定支持，并未解决冲突；真正的修订会改变论证中哪些结论仍能由这项材料推出。

传播还会改变读者的研究能力。有人学会以漂移问题重新阅读区域资料，可能发现以前未被注意的对应；有人因此寻找反例，也可能提高模型质量。思想的延续可以经由支持和批评两条路径发生。不过，只有出现新增资料、方法或明确问题，才算知识生产。单纯争夺先驱地位，或者重复赞美勇气，不构成同样意义的延续。

需要区分历史问题与原始答案。大陆为什么呈现某些跨洋关联，可以作为问题持续存在；某位作者给出的确切驱动机制则可能被淘汰。若把问题与答案绑定得过紧，支持者容易认为放弃机制就是背叛整个研究方向。事实上，可生产的研究常依赖这种拆分：保留真正受到资料约束的困难，释放缺少支持的解释，让新的对象和操作进入。

本书使用“问题链”一词，指的是这种可追踪的承接。一个问题引出某项比较，比较暴露某种不足，不足又规定下一项工作。链条可以中断、分叉，也可以由不同团队接续。它不要求思想始终拥有同一个名称，更不要求所有参与者共享同一哲学。只要后来的工作明确承担了早先留下的责任，就存在可研究的连续性。

这也解释了为什么某些看似边缘的思想值得保留，另一些则不值得无限投入。前者能够说明目前缺什么、如何获得以及结果会怎样改变判断；后者只有“未来总会证明我”的承诺。时间本身不会把一个命题变真。长期未被接受，既不是错误的证明，也不是创造性的凭证，必须回到实际问题推进。

◆ 25.3 动力替代方案的真实地位

霍姆斯的动力方案涉及三个不同的文献节点。1931 年《自然》短通信主要澄清有关表述，其论证范围不同于完整长文。[B12] 格拉斯哥地质学会的官方材料保存了 1928 年 1 月 12 日相关讲演的会议记录摘要，并记载长文后来刊于该会 1931 年的《会刊》第 18 卷。[R25a] 讲演记录说明问题曾如何提出，长文承载展开的论证，短通信则属于公开讨论的另一个环节。区分三者，才能判断某种机制构想何时出现、以何种形式进入讨论。

会议摘要显示，他把深部缓慢流动与不均匀放射性加热联系起来，并尝试用上升、横向作用与下沉的安排讨论地壳运动。[R25a] 这是一项重要方向，但不应直接翻译成今日

完整板块动力学。早期的物质分类、地壳结构和过程图景包含自身时代的前提。把现代海岭、俯冲和岩石圈概念全部填回早期文字，会让真正的创新与真正的差异都消失。

对流候选为什么有吸引力？它把能量来源、材料运动和地表变形放进同一机制问题。与只给大陆附加一个外部推力相比，内部热状态可能改变介质本身的运动能力。但“有热”不等于“一定以所需方式对流”。温度差、材料黏性、尺度和边界条件共同决定流动是否发生及其形态。一个热源可以存在，却不足以支持某个具体对流格局。

教学上可以用加热流体说明浮力与黏性之间的竞争，但必须指出地球内部不是一锅普通液体。岩石在长期高温高压条件下的变形，需要物性与时间尺度分析。以厨房经验代替深部地球，只会把早期机制困难换成另一种直观误导。对流概念打开研究方向，具体动力学仍须通过理论、实验和间接观测逐步约束。

还要区分流动存在与表面运动由何种作用主导。即使内部存在对流，表面刚性单元如何与之耦合、哪些边界受牵引、哪些地区发生伸展或挤压，仍是独立问题。早期支持者若提出一种机制候选，可以使漂移问题不再只有旧的驱动力选项，却不能仅凭候选存在就宣布运动已获充分说明。候选的地位是提高可研究性，而不是取消检验。

这一节点的重要变化，是反对意见被转化为新的研究对象。原来问“哪种力能推动大陆”，后来可以更具体地问“内部热与材料流变是否允许某种长期运动”“这种运动如何与地表证据一致”。新的问题需要不同资料：物性、热状态、深部结构及边界关系。若后来的研究确实沿这些方向形成可追踪任务，就能说机制争议参与了研究改道；具体到某一团队，则仍需工作记录支持。

会议记录支持将霍姆斯的构想定位为深部热、缓慢流动与地壳运动相联系的机制候选。[R25a] 这一定位还没有给出流速、温度与应力的定量相容性，也未说明表面运动究竟由何种作用主导。候选方案的思想价值，在于它改变了动力问题的研究方向：所需检验从寻找一个附加推力，扩展到热量怎样输运、岩石怎样变形及深部运动怎样传递到地表。它与后来的板块动力学之间存在可研究的联系，也保留着这些尚需填补的环节。

热源与循环之间还有一个重要差别：热量可以通过传导传递，也可以随物质运动传递。若只证明内部存在放射性产热，还没有证明对流是主要输运方式。必须考察温度梯

度、材料响应与空间尺度是否允许流动。这个区分使霍姆斯问题具有进一步可研究性，因为能量来源、传热方式与表面变形可以分别接受约束，再检查它们能否组成同一机制。

对流图中的上升和下沉也不应被当成固定管道。材料运动与热分布可能互相改变：流动带走热量，温度改变黏性，黏性又影响流动。即使采用简化模型，也需要知道哪些量被固定，哪些量允许反馈。早期机制图的意义在于提出这种内部活动的可能性；完整预测则要求确定参数和边界条件。把概念图当作已经计算出的地球内部，会夸大它的完成程度。

物质密度进一步连接浮力与结构。温度差可以改变密度，成分和相变也能改变密度，几种作用未必同向。某种材料在高温下倾向上浮，不代表其在不同压力与成分条件下始终如此。深部运动的可行性因而不能从单一性质读出。后来的实验与地震研究之所以重要，正因为它们逐步限制了这些互相牵连的属性。

机制替代的真正价值，是使原来不可处理的反对变成一组可以分工的问题。测年与放射性研究限制热源，物性研究限制材料变形，地震资料限制内部结构，地表构造则限制需要解释的运动。它们不必一开始就形成完整答案，却可以使争论不再停留在“是否相信漂移”。这就是未决问题获得生产性的一个具体机制。

动力候选的研究价值还取决于不确定性能否被约束。若黏性范围变化会使预期流动时长跨越许多个数量级，物性研究就可能直接改变机制可行性的判断；若某项局部参数怎么改都不影响预测方向，测准它却未必能解决争论。优先调查哪些未知量，应由它们与核心结论的联系决定。这样，有限资源才有可能集中到真正使支持者与反对者改变判断的地方。

未决问题由此可以形成不同工作路线。年代冲突要求重查样品与相关，运动方向冲突要求检查重建和边界关系，力学不相容则要求比较介质与载荷模型。三种工作可以并行，却不能相互充数。只有明确每条路线解决哪一种困难，新的化石、新的测线或新的物性结果才会进入论证，问题的延续也才具有可辨认的内容。

对漂移问题而言，最值得保留的不是一种不可动摇的立场，而是不断要求轮廓、年代、物性与运动彼此负责的研究关系。这个关系允许旧答案退出，也允许新答案接受同样严格的检查。

因此，同一漂移问题再次进入讨论时，可以保留跨洋关联，却不必保留原先的驱动力；可以沿用旧样品，却重新规定它在新的运动单元中的位置。延续发生在可说明的承接与撤回之间。

◆ 25.4 设置失败而非永远延期

研究路线需要停止条件，否则任何理论都能靠“证据还没来”永远存活。停止不一定意味着全部问题毫无价值，而可能意味着某个机制版本、某种测量路线或某项重建不值得继续投入。关键是把判断落到能够检查的变化，而不是支持者的热情。一个假说若不断生成新工作，却始终没有任何结果能够改变它，也许只是在生产忙碌。

第一种停止条件，是关键预测在足够精度和适用条件下反复失败。这里“足够”必须事先说明，否则失败可以永远归咎于仪器。第二种，是每次修补都增加不受独立限制的自由参数，使模型失去区分力。第三种，是核心对象的定义持续漂移，以至于不同版本已不能回答同一个问题。第四种，是可行检验消失而剩下的主张又没有明确解释收益。它们都需要具体判断，不能化为一张自动淘汰表。

资源限制也影响研究选择，但应与真假判断分开。一项问题可能仍有意义，只是当前测量成本过高；另一项问题可能容易测量，却缺少值得追问的差异。暂停工作不等于宣布假说错误，继续工作也不等于宣布假说正确。科学共同体需要在多种不确定路线间分配精力，合理决策可以包含少量高风险研究，但不能让风险本身成为免检资格。

SDE 框架在这里提出的是条件化延续。显露 S 是一条仍可组织工作的研究问题；差异序列 D 包括批评、修订、测量与反例；特征纠缠 E 则由机制可信度与解释收益相互限制。解释收益使某个未解决机制值得继续研究，机制困难又限制研究者能够对外作出多强的结论。若这两者不再互相约束，只剩收益宣传而不承担失败，所谓存活便退化为维护信念。

一个更深的区分是，知识可以以问题形式存活，而以答案形式死亡。某个驱动力被否定以后，它所试图解释的大尺度位移并未因此消失；某一重建被修改以后，跨洋对应仍要求说明。知识更新的单位不是整本书或整个人物，而是可以分解的对象、关系和操作。这样的分析既避免把旧知识全部丢弃，也避免为保留历史贡献而拒绝修正。研究路线的衰退

也不意味着所有产物随之消失。失败模型可能留下高质量地图、岩石收藏、测量技术或一个后来仍重要的反例。评价时应把结论效力与研究产物分开。反过来，产物有用也不能替错误结论恢复真值。科学史若只在英雄与废墟之间选择，就看不到这种复杂继承。

真正的停止条件应当允许重新启动，但启动需要新理由。过去因测量能力不足而暂停的问题，在新仪器出现后可以重新研究；过去因关键预测失败而放弃的模型，则需要解释旧失败为何不再适用。仅仅隔了几十年，或换了新的术语，不构成重启依据。知识的再生需要条件变化，而不是遗忘失败记录。

大陆漂移问题的后续发展之所以值得研究，是因为海洋资料改变了问题的可检验结构，而不只是因为时间终于站到支持者一边。新的海底骨架、磁性时间记录和运动边界，使某些长期问题获得新的操作入口。下一章由此回到地图：有限测深剖面怎样生成连续海岭体系，又怎样在新的观测中承担被推翻的风险。

第 6 编

海底扩张与板块世界的发生

第 26 章 海底地形：测深剖面怎样生成新的地球骨架

第 27 章 海底磁条带 从异常信号到时间与空间的对应

第 28 章 海底扩张：一个假说怎样承担多项独立检验

第 29 章 转换断层 新运动关系怎样改变边界的含义

第 30 章 板块世界的成立：运动学闭合、共同检验与新任务

第 26 章 海底地形
测深剖面怎样生成新的地球骨架

26.1 剖面不是一幅现成照片

海底地图让不可直接连续观看的世界呈现在眼前。读者看到山脊、谷地和海沟，容易以为科学家只是把它们照下来再描到纸上。实际上，早期大洋制图的重要困难，正是没有这样一张完整照片。测深沿船舶航线展开，航线之间留下大片空白；一幅连续图景必须包含由剖面走向、地形形态与其他资料共同限制的重构。理解这一点，才能同时欣赏地图的创造性和检验其可靠性。

希曾与玛丽·萨普 1957 年的大西洋海底地貌图，是这一转变的著名节点。国会图书馆的馆藏介绍说明，该图采用从上方略带侧视的地貌表达，并在资料不连续的情况下根据剖面推断结构延伸。[B05] 这样的图像可以迅速传达总体形态，但不是每个笔画都具有测点同等的证据身份。直接约束、解释连接与表现性绘制应当分层阅读。

侧视表现具有强烈认知作用。一条等深线需要读者在脑中恢复起伏，而山脊式绘法让起伏关系直接显露。它降低了理解门槛，也可能增加视觉确定感。画得连续的山岭看上去比一组离散测点更像真实对象，然而图像说服力不能替代连续性证据。制图的美学价值与科学责任并不冲突，只要它们不被混为同一种可靠性。

剖面本身也经过表达选择。纵向比例若相对横向夸大，缓坡会显得陡峭；不同剖面采用不同夸张比例，形态比较便可能失真。平滑程度决定哪些小尺度起伏可见，基准面决定整体高度关系。研究者不必拒绝这些表达工具，但必须说明它们的作用。教学图为了辨认结构而夸张起伏，与定量测量真实坡度，是不同任务。

一条航线穿过一个凹槽，并不能单独说明存在纵向延伸的裂谷。凹槽可能是局部洼地，也可能是斜穿复杂地形的结果。若相邻剖面在相应位置也出现类似形态，连续结构便成为更强候选。此时最重要的信息不是每条曲线都完全相同，而是它们的相似与差异能否由共同几何解释。过分要求相同会忽略自然变化，任意允许不同又会使连续性无法检验。

地图因此是一种有责任的综合：它把分散资料组织成可以继续操作的结构，同时把自己的薄弱处暴露给新测线。若一幅图只展示完成后的美感而抹去资料空白，它就不利于后续检验；若只保留原始点而拒绝任何连接，又可能无法提出更大尺度问题。科学制图的困难，在于创造性推断必须与可追溯限制共同存在。

地图中的透视还会造成前后遮挡。为了让海岭显眼，绘制者可能突出坡面而弱化某些细小结构，这有利于整体阅读，却不适合直接测量距离和坡度。不同地图形式因此服务不同目的：地貌图帮助识别组织，等深线图帮助读取高程关系，测线图帮助检查数据覆盖。将三者并读，能够避免让一种图承担全部任务。

空白也可以成为主动设计的信息。若图上标出约束较弱的区域，后来调查便知道在哪里最有可能改变解释。没有空白的图看似完成度高，却可能降低后续研究的针对性。科学地图的成熟不在于隐藏未知，而在于使已知和未知具有可操作的边界。

◆ 26.2 从局部凹槽到连续裂谷

把多个剖面上的凹槽认作同一裂谷，需要建立一种跨空间同一性。这里的“同一”不是每个点形状完全一样，而是这些局部部分属于可连贯追踪的结构。研究者要比较它们相对于海岭轴部的位置、走向、邻近地形及可接受的偏移。一个没有任何限制的连线，可以穿过所有想连接的点，却不具有地质意义。

测线方向是首要变量。垂直穿过狭长谷地时，横向宽度比较接近其真实宽度；斜穿时，沿航迹观察到的宽度会扩大。如果不做几何分析，相邻剖面的宽度差可能被错误地解释为结构突然变宽。类似地，一条沿谷地延伸的航线可能长时间保持较深水深，却难以表现两侧谷壁。不同方向的剖面共同提供结构约束，而非简单重复同一信息。

连续性还需要处理位置误差。若每条航迹的不确定范围较大，可以允许相应的连接带；但这不等于允许任意弯曲。误差应独立估计，不能为了保住连续裂谷而逐段扩大。若连接要求的弯折与相邻地形完全不相容，或需要在无资料区绕过所有反例，模型就开始失去判别力。空白区不是无限自由空间，而是明确尚未受约束的区域。

可以用一个教学设计理解可检验插补。已有三条横断面显示近似对应的轴部凹槽，研究者提出一条候选连线。随后在两条旧线之间增加一条新线。若新线在预期范围内也显示

凹槽，连续性获得新增支持；若只有外围丘陵，便需要检查是对象中断、位置预测错误还是测量不足。关键在于预测范围必须在新线测量之前规定。否则，任何结果都能通过重新画线被纳入。

海岭也不是一条没有分段的理想直线。局部方向变化与错开，可以成为其真实结构的一部分。问题在于这些分段是否遵循可说明的连接关系。后来的转换断层概念将为某些错开提供运动学身份；在早期地图阶段，不能提前把所有错开都按成熟理论解释。地形观察先提供候选结构，运动关系需要另外的证据。这一时序区别保护了发现过程中的真实不确定性。

跨剖面同一性一旦稳定，海岭不再只是很多高地的集合。研究者能够沿其走向提出比较，能区分轴部、两翼和连接区，能设计在不同位置重复的测量。新增对象具有了内部组织和可操作坐标。它不是凭名称变成整体，而是因为不同局部特征开始互相限定：某处被判为轴部，会影响邻近部分的身份；新的偏移又会要求重新检查连续性。

剖面形态与跨线连续性之间，形成了本章所说的特征纠缠。一个凹槽若位于海岭轴部，邻近剖面便有理由在相应位置寻找延伸；邻近剖面若显示轴部已分叉，原凹槽又可能需要重新归类。局部形态限制连线，整体结构限制局部成员的身份。这里的往返有明确的自然约束：连线必须经过位置误差允许的区域，轴部归属必须与周围起伏相容，新增剖面也可以切断原先认为连续的结构。

跨剖面对比可以进一步分解为位置、形态和邻接关系。位置问凹槽是否落在可连接范围，形态问两侧起伏是否相容，邻接问它与更大尺度海岭及周边盆地怎样排列。三项一致比任何单项更有力，但并不要求细节完全重复。一个自然结构可以在延伸中改变宽度和深度，正如河谷不会每一段都相同。研究者需要寻找稳定组织，而不是复制图形。

假设一个候选裂谷在四条剖面上出现，其中一条不明显。至少存在几种可区分解释：该段真实中断，剖面方向不利，沉积覆盖了形态，或测线位置偏离。每一种解释都有后续操作。真实中断需要检查邻近连续性，方向问题需要增加横断面，覆盖问题需要其他探测，位置问题需要重定位。将反例分解为不同任务，比直接删去异常剖面更能促进对象稳定。

结构命名还应具有层级。一个局部凹槽可以是较大裂谷体系的一部分，却不必与整个体系共享全部性质。整体名称用于组织比较，局部描述保留差异。若把整体特征强加给每一段，就会把真实异质性看成噪声；若只保留局部差异，又难以提出跨区域规律。海底骨架的发生正是在这两个尺度之间建立可往返关系。

◆ 26.3 地震叠图的限制与力量

地震资料提供了不同性质的空间记录。测深限制界面形态，震源位置则与发生破裂的区域有关。将两者放在同一坐标中，可以检查某些地形结构是否与活动分布相联系。国会图书馆对相关制图过程的说明提到地震与地形的叠加。[B05] 这一事实说明地图形成已经包含多源综合，也立即带来独立性问题：参与绘图的地震资料，不能随后原封不动地再被当作独立验证。

独立性并非要求不同仪器绝无共同条件，而是要求模型没有提前利用检验信息来取得成功。若研究者根据地震带位置填补了地形空白，再指出海岭与地震带一致，这种一致属于构造模型的一部分。它可以显示多类资料能够相容，却不能作为额外一票。真正新增的检查，可以来自后来获得的测深剖面、未参与绘图的震源资料，或者另一种具有不同误差的观测。

震源定位有自身不确定性。台站分布不均、到时识别和速度模型会影响位置，误差常不是各方向相同。把震源画成一个小圆点，容易掩盖可能覆盖较大区域的定位范围。若地形线条与震源点之间的距离小于定位误差，只能说两者相容，不能宣称震源精确落在某条狭窄断裂上。图上重合的锐利程度不等于实际证据的锐利程度。

地震活动还存在时间抽样问题。某段结构在有限观测期内没有被记录到地震，并不自动意味着它不活动；某个震群密集出现，也不一定代表长期平均活动最强。不同目录采用不同检出阈值，海洋地区的小震可能更容易漏记。比较之前要问记录窗口和观测能力是否相当。否则，地震空白既可能是物理现象，也可能是观测空白。

尽管如此，多源相容仍具有实质力量。一个根据测深提出的狭长结构，若与独立定位的活动带在合理误差范围内持续对应，纯粹偶然地形解释便受到限制。更进一步，震源机制可以检验运动方式，而不仅是位置。于是证据从“这里有一条地形”推进到“这里以何

种方式变形”。这种推进需要保留各项资料的具体贡献，不能把它们压成一个模糊的综合图像。

叠图还会暴露冲突。地形连续而地震活动发生偏移，可能表示现今运动边界已改变；地震带延伸到没有明显地形的区域，可能提出新的调查目标；地形错开而运动机制不同，则可能要求不同边界解释。冲突不是综合失败的唯一标志，也可能是对象需要分层的信号。重要的是不能为了维持一条整齐的线，把不同时间尺度的结构强行并为同一实体。

自然地形保存过去的累积，地震记录有限时段的活动，磁异常保存另一个时间窗口的过程。它们相遇时，不只是空间坐标重合，还需要历史时间的相容。新地球骨架的形成，正是这些不同记录逐步取得可对读关系的过程。每项记录都必须保持自己的时间身份，才能避免把古老遗迹与当前边界混为一谈。

地震与地形相容还有一个方向性问题：某些事件可以发生在海岭轴部附近，另一些位于横切结构。若只统计它们都离海岭系统不远，便会丢失区分边界类型的重要信息。空间关系应具体到沿向、横向和分段，而不是只有接近或远离。后来运动学之所以能提出更强判断，正是因为这些细节不再被总体相关吞掉。

同一条地形线还可能包含不同年代的段落。今天活动的一段，与两端已经成为遗迹的延伸，在地图上都可能清楚可见。若仅凭地形连续认定活动连续，就会产生错误预期。第 29 章的转换断层将把这一点具体化：可见断裂带的全部长度，不等于当前相对运动边界的全部长度。地貌对象与运动对象可以相关，却不必同一。

地形连续而地震只集中于一段，便要求把“结构延伸”和“活动范围”分别表达。前者可以由海底留下的年龄与热状态差异维持，后者取决于当前相对运动。强迫两张图完全重合，反而会丢失这个区分。新的对象分类有时由此产生：同一条地貌带内部，可能只有部分区段承担今天的边界运动，其余区段保存的是继承结构。

新地图还有一种跨学科作用：它为不同记录提供共同的位置语言。热流点、磁测线和岩石样品原本可能分别归档，海岭及其分段让它们可以按相同结构位置组织。但共同语言不能消除尺度差异，某个热流点仍只代表有限范围，某条磁曲线仍综合一定体积的信号。结构坐标使比较成为可能，具体方法继续规定比较的权限。

这种协调也会产生新的不一致。例如轴部样品显示的性质与预期不同，可能提示地形轴不等于当前生成中心，也可能反映局部过程。地图因而不是最终裁判，而是提出问题的工作平台。它使异常具有位置和关系，异常又迫使地图的对象分类更加精细。

从地方剖面到连续骨架，新增的是局部记录之间的约束。一处轴部位置改变，会影响邻近测线的对应；某段走向修订，会改变下一条横断面的安排。整体结构因此既能指导局部观察，也会承受局部观察的反驳。

地图中尚未确定的连接，应随覆盖与误差一同保留。对一个长空白区，增加中间横断面可能比继续细描两端更重要；对覆盖充分而几何仍冲突的区段，则应重查对象是否被错误连接。未知部分具有不同类型，地图的作用正是把它们定位成不同调查任务。

这些任务使海底骨架超出了展示功能：它成为安排测线、比较样品和解释活动位置的共同坐标。

◆ 26.4 新骨架如何改变研究任务

海岭作为可追踪对象出现后，研究问题会发生方向性变化。热流测量可以比较轴部与两侧，岩石采样可以围绕结构位置安排，磁测可以选择横穿和沿走向的航线。这样的设计把原本难以定位的问题变成有空间坐标的比较。科学对象的稳定价值，不仅是允许人们说“这里有海岭”，还在于允许不同团队问可相互连接的问题。

但对象引导采样也会造成选择偏差。若研究者只在预期最典型的位置取样，可能忽略破坏模型的边缘地区。为了检验整体性，需要同时设置对照和异常区。轴部与两翼的比较有意义，离开候选结构的区域同样有意义。研究路线被对象选择，并不意味着对象有权排除所有不利观察；相反，成熟对象应当能够规定最有风险的检验位置。

资料增长后，地图也可能改变。某段被画成连续的山岭可能实际分为几段，某处凹槽可能与邻区结构无关。若修改有清楚原因和版本记录，这是对象稳定化的一部分，而不是制图失败的耻辱。无法修改才更危险，因为它意味着地图已从可检验模型变为必须维护的图像。一个可靠地貌体系，应能够吸收局部修正而不靠无限自由插补保护核心。

若用 P1—P3 标记分析相位，P1 对应跨剖面地貌对象的形成，P2 对应依据该对象安排后续测线，P3 对应新测量促使连线与分类修订。三者有不同证据要求：解释图能显示

第一项，航次计划与实际航迹之间的关系才可能显示第二项，前后图件及修订理由则用于判断第三项。现有馆藏说明确认多源制图与空白区推断；具体测线是否由图像引导，仍需要上述连续记录。这些相位用于区分改变发生在哪一层，并不预设每次制图都会经历完整循环。

如果后来测线反复否定关键连续性，研究者只能靠越来越曲折的插补维护海岭，那么这个对象的稳定身份就应受到质疑。如果新测线总体维持结构而修正局部细节，则连续体系获得更强约束。这样的失败门槛使“地球骨架”不只是一种修辞。它指的是能够组织操作、承受新证据并参与下一轮推断的空间系统。新骨架还改变样品之间的比较方式。过去两块海底岩石可能只按采样海域并列，现在可以按轴部、侧翼或远离生成区的位置比较。空间身份为岩石性质增加了过程问题，但不能替代成分与年代分析。只有当位置分类与独立性质形成稳定关系，地图对象才真正进入其他学科的工作，而非停留为图版上的名称。

空间系统成立还只是一步。海岭与两翼可以被定位，不等于人们已知道两翼何时形成、怎样运动。下一章将进入磁异常：同一空间骨架怎样获得时间信息，又为什么船上曲线的正负绝不能直接等同海底岩石的黑白条纹。

第 27 章 海底磁条带
从异常信号到时间与空间的对应

◆ 27.1 先把测线信号与磁化分开

海底磁条带不是海底铺着的黑白地毯。船上的仪器记录磁场变化，研究者经过处理和解释，才把某些变化与海底磁性结构联系起来。教科书把两侧画成相间的颜色，能迅速表达概念，却也可能掩盖最重要的工作：曲线怎样对应地下材料，地下材料怎样对应形成历史。

Vine 与 Matthews 的 1963 年论文，把海岭附近磁异常与洋底扩张、地磁倒转联系起来。原文还提供了一个值得保留的警告：在所讨论的低磁纬条件下，沿当时地磁方向磁化的物体，也可以在其上方产生负异常。因此，“负异常就是反向磁化”不是可以脱离位置与几何直接套用的规则。[B03]

异常的正负，是相对于某一参照磁场而言的差值；磁化方向，则描述材料的磁性状态。两者不是同一个量。仪器在一定位置记录的结果，还受磁性体的深度、形状、方向与观测几何影响。若把这层关系抹掉，磁条带似乎只需要看图认颜色；恢复这层关系，才会出现真正的科学问题。

为理解这一点，可作一个观测推演。设地下磁化分布为 m，观测位置、场方向及几何条件合记为 g，信号写成 a＝L_g[m]＋ε。L_g 表示从磁化分布计算观测信号的正演操作，ε 包括测量误差与模型尚未表达的贡献。在固定几何、采用线性磁场叠加等近似时，这一写法便于区分信号来源；若几何或磁性响应也需估计，相应反演会更复杂。该式是教学表达，用于说明观测层与材料层的关系，并非 1963 年原论文公式的转录。

这个写法的价值，在于显示两种不同工作。正演问：假设地下如此，应当观测到什么？反演问：已经观测如此，地下可能是什么？从正演转到反演，通常不能不加条件地得到唯一答案。不同分布可能产生相近信号，尤其当资料有限、几何不确定或误差较大时。

因此，发现一组模型能拟合曲线，只说明它暂时具备解释资格。要提高可信度，还应问：模型是否用过多自由参数，是否与地形或其他资料冲突，是否能解释未参与调整的新测线，以及是否存在同样有依据的替代模型。曲线越复杂，越不能把“能画出相似形状”当作已经证明真实地史。

本书所说的特征纠缠在这里有具体落点。观测几何规定同一磁化状态会怎样显露；由信号推断的磁性结构，又必须与几何和岩石条件相容。条带宽度、磁化方向和地形不是三张互不相关的清单。改变其中一项，其他项的证据意义可能随之改变。

例如，增大磁化层厚度可能增强异常振幅，减小源体到测线的距离也可能增强信号；若只有一条曲线，两种变化可能部分补偿。海底地形限制了距离的一部分，却未必确定磁化层底界；岩石磁性资料则可以限制磁化强度，却未必直接给出整个源体形状。所谓特征纠缠，在这里要求辨明哪些量能够相互补偿、哪些外部资料能够拆开补偿，才有助于确定地下结构。

磁异常从测量结果进入科学对象，并不是信号被主观赋予任何意义。它意味着研究者形成了一套受到材料约束的辨认能力，能够区分哪些差别可能来自磁化，哪些可能来自地形或观测方式。新世界的发生，首先表现为错误不再只有一种笼统名称，而能被定位到不同环节。

为了把这一点落实，可以把同一条曲线分成三个处理层级来阅读。第一层是仪器及定位记录：它回答何时、何处记录了什么。第二层是异常表达：它说明从哪些参照和处理得到曲线。第三层是地下磁性与历史解释：它提出什么模型能够说明这些变化。第三层改变时，第一层记录可以保留；第二层的参照改变时，第三层也可能需要重做。三层之间存在联系，却不能用一层的确定性替另一层担保。

三个处理层级还要求保留不同记录。仪器输出可用于重查漂移与异常点，航迹和时间标签可用于重做配准，参考场与改正项可用于恢复异常定义。如果只保存最后一条曲线，就难以判断一次新解释来自磁化模型改变，还是来自处理基准改变。资料的可复核性因而依赖这些层级间的连接，而非仅依赖某一条记录被称为“原始数据”。

假定两条航线看到相近波形，仍有几种可能。它们穿过同一磁性结构；它们分别穿过不同结构，却因观测几何产生相似信号；或者处理中的共同步骤制造了相似性。要说“同

一条带延伸到这里”，必须把地理连接与信号对应一起检查。曲线相似不是空间同一性的自动证明。

反过来，同一结构在不同观测条件下也可能显得不同。若研究者只接受外观近似的曲线，可能把真正可连接的资料拆开。因此，对比的任务不是寻找表面一样，而是找出在明确变换和条件下哪些差异仍能共同成立。这里“同一”本身需要工作，而不是所有比较天然拥有的起点。

曲线对比由此产生了一项新的辨认任务：在不同观测位置与条件下，找出能够由同一磁化结构解释的部分。研究者既不能只按峰谷外形配对，也不能任意改变磁化分布来迁就每条曲线。共同模型必须同时接受多条测线、地形及物性限制。海底磁性对象的稳定，体现于这种跨记录的相容性，并为下一步赋予它时间含义提供前提。

◆ 27.2 把反转年代作为外部约束

即使磁化分布已经获得支持，时间意义仍需另一步。正常与反向磁化的排列说明方向差异，却不自动给每一条带标出形成日期。要从空间顺序进入时间顺序，需要把材料记录、地磁极性变化与海底形成过程连接起来。1963 年的提出阶段与后来更充分的标定和检验，必须保留在不同证据时点。[B03]

这里有一个容易隐藏的循环。研究者先用某种扩张速率把空间条带换成时间，再宣布条带与据此编排的时间表吻合，最后用这种吻合证明速率正确。如果时间表与速率都由同一组条带自由调整，就没有获得独立约束。参数之间相互协调，不等于模型已经接受外部检验。

这不意味着所有相互校准都无效。许多科学工作本来就需要联立估计多个未知量。关键在于有多少独立信息，有多少自由调整，以及剩余约束是否足以排除替代方案。若一个模型有很多参数，只接受很少资料的限制，它可以看起来很合适，却仍没有清楚的判别力。

可用一个教学例子说明。假定某测线辨认出三段交替磁化区域，宽度依次为 2、5、3 个距离单位。仅从这一组宽度，不能唯一判断每段持续多少时间，因为速率未知，也可能变化。若另有独立约束指出三段持续时间比例为 2∶5∶3，恒定速率模型便获得一种支

持；若时间比例为 1∶1∶1，则至少需要变化速率或其他解释。这里所有数字均为逻辑演示，不是历史测线资料。

进一步说，即使比例符合，也还不能结束。条带边界是否辨认准确？测线是否斜交条带？所谓独立时间比例是否真的没有用过这条测线？每一项都会改变“符合”的含义。科学检验不是寻找一幅让人满意的重合图，而是确定重合受到哪些不可随意移动的约束。

在思想史层面，必须防止把成熟阶段的约束提前交给早期作者。后来能够使用的年代标尺，不能自动成为 1963 年提出模型的依据。早期研究可能只证明一种机制有希望解释现象，随后研究才逐步增强或限制它。若把这两个阶段合并，便会让理论看起来从出现时就已经拥有全部证据，也无法解释后续检验为何重要。

从 SDE 的不同入口看，研究可以先从测线差异 D 检验几何与磁性条件 E，再判断时间序列 S 是否成立；也可以先提出候选时间序列 S，回查其所需条件 E，并据此选择下一组比较 D。前一路径容易暴露曲线解释的歧义，后一路径容易暴露时标与速率之间的补偿。两条路径最终必须汇合到同一组资料限制。这是对研究组织的分析，不据此指定 1963 年作者实际采用的心理或操作顺序。

两种入口若给出不同结论，应追查是否悄悄改变了约束。例如，从曲线出发时仍待确定的某个条带边界，不能在从时间序列出发时突然成为固定锚点；反过来，已有独立年代范围也不应只因拟合不便就被放宽。路径改变能够帮助发现这种前提移动，却不能增加资料本来没有提供的信息。

当外部约束逐步增加，条带才可能取得一种新的时间身份。它不再只是空间中磁信号不同的区域，而成为可以在一定条件下比较生成先后的记录。这个对象包含的能力比图形识别更多，也承担更多失败方式：一旦关键对应关系失效，时间解释就必须缩小或修订。

时间序列的识别还涉及序列长度的问题。一个短序列常常可以在许多候选序列中找到相似片段。假定记录只有“正、反、正”三个符号，这样的排列本身区分力很低；增加若干区段，保留相对宽度及排列关系，可能提高约束，但也可能同时增加边界识别错误。更多信息不是无条件更好，必须看新增信息是否可靠而且能排除候选。

还要区分“极性类别相同”与“属于同一个极性时期”。地磁方向可以再次进入同一类别，所以两个正常磁化区域并不因此同龄。它们在序列中的位置、相邻区段和外部时间

锚点共同决定是否可以相关。这与化石相关中的问题相通：一个孤立的相似特征，通常不足以完成时间身份。

由此可提出一个本书的约束层级。没有外部时间信息时，只讨论某种空间序列与生成模型是否相容；只有相对时间约束时，主要讨论先后与对应；取得可靠绝对时间锚点后，才讨论相应区间的速率和持续时间。若资料只到第一层，文字却已经进入第三层，论证便发生了越级。

当然，研究不会总是按整齐层级线性前进。一次新的外部测年可能迫使重新辨认条带，条带重辨又可能改变对测年样品相关位置的理解。这样的往返并不天然是循环论证。判断关键在于每次修订是否增加独立限制，还是仅仅让两项未知彼此迁就。前者可以使解逐渐收敛，后者可能只维持表面一致。

一个可以实施的检查是暂时固定部分约束，观察其他结果的敏感性。把某一时间锚点在合理误差范围内移动，条带对应是否保持？把一处边界识别改动到另一个合理位置，主要结论是否崩溃？这些问题不是要求所有结果完全不变，而是要知道结论依赖在哪里。敏感性如果集中在一个薄弱环节，后续研究就有明确的补证方向。

这里的“改道”具有可操作含义。若主要不确定性来自时间锚点，继续加密同一磁测线可能收益有限；若主要来自空间定位，则再增加一份相同时间标定也未必解决问题。研究资源应当沿证据的薄弱处重新分配。这是 SDE 路径分析可以提出的候选增量：它不仅整理已有材料，还应能定位下一步最可能改变判断的操作。

这种诊断与常规反演的敏感性分析相通。它是否带来额外作用，可以用后续工作检查：依据诊断增加的年代点或测线，是否排除了原先并存的候选；若没有，原因是误差仍大，还是所选资料对候选本就缺少差别？SDE 在此承担的是连接对象身份与调查选择的组织任务，额外解释力应由这种可追查的结果体现。

◆ 27.3 条带宽度与扩张速率

把时间与空间接上以后，才出现速度问题。最简教学模型把海岭一侧某段新生地壳离开生成位置的速率记为 u，持续时间记为 Δt，沿运动方向的距离记为 w。在速率近似恒

定、几何适用的条件下，$w = u\Delta t$。如果 u 随时间改变，应使用相应时间区间上的积分，而不是把一切变化塞进一个平均值而不说明损失了什么。

这里 u 是一侧的速率。在两侧对称并且参照明确的简化情况下，两侧分离速率才是 $2u$。若把半扩张速率与全扩张速率混用，即使距离与时间都正确，也会出现成倍差异。速度不是只带一个单位的数字，它还包含谁相对于谁、沿哪个方向、在哪个时间区间内运动。

再看测线方向。若条带近似平行，船沿斜交方向穿过，则沿航迹量到的宽度通常大于法向宽度。设航迹与条带法线夹角为 θ，平直近似下法向宽度为航迹宽度乘以 $\cos\theta$。这一几何修正并不证明扩张模型，却说明资料进入模型以前，必须处理观察路径。把航迹长度直接当成生成方向的距离，会让操作差别伪装成自然速率差别。

以上两式都是本书为解释条件而给出的简化推导，不是对真实海域的计算结果。实际结构弯曲、扩张不对称、边界变化和资料分辨率都可能使近似失效。一个好用的公式，不应通过隐藏适用条件来获得普遍性。

可以再作一个不含实际年代的比较实验。固定同一段时间间隔，如果一侧条带明显更宽，可能是两侧速率不同，也可能是几何投影没有处理，或条带对应本身不正确。仅凭宽度差，不足以指定唯一原因。研究者应寻找能区分候选解释的材料，而不是选择最适合既有叙事的一种。

区分这些解释，可让同一对条带接受不同检查。若差异来自斜交几何，按测线方向改正后宽度应趋于相容；若来自错误相关，延长剖面后邻接序列可能无法对应；若几何与相关均稳定，独立年代仍支持两侧宽度不同，不对称扩张才成为更强候选。增加条件的意义，是产生这些不同后果，而非使任何宽度都能被原模型容纳。

1963 年原文对计算图作了有限声明：展示基本异常形态，并未逐点复现观测剖面。这一限制应保留在思想史中。它告诉我们，当时的工作是在建立可行解释，而不是已经完成后来教科书式的全部确认。[B03]

有限声明不削弱科学价值。相反，它使下一步任务变得清楚：需要什么样的独立时间信息，需要哪些新测线，需要何种几何或磁性检验。理论如果能把模糊疑问转成这些可执行问题，已经改变了研究能力；但这种改变与完整验证仍是不同阶段。

条带一旦成为时间记录，下一轮比较就会改变重点。已有磁曲线若允许两种年代对应，加密相同曲线未必解决歧义；位于两种对应年龄差最大的区段中的独立样品，可能更有判别力。新年代若否定某种对应，速率、磁化边界与邻近序列也要随之重查。以三式表述，这是 S 对 D 的选择以及新 D 对 E 的约束；具体历史中是否发生了这一回写，还需论文与取样记录相接。

距离与时间关系还有一个更深的对象问题。同样的平均速率，可以对应完全不同的运动历史。设在一段总时长内，模型甲始终均匀运动；模型乙前半段较快、后半段较慢，但总距离相同。只看两个端点，两者无法区分。若内部条带提供额外时间位置，就可能增加判别。细分记录的价值因此不在于图形更复杂，而在于它能否排除原先等效的历史。

这可以写成一个简单的逆问题。已知总距离 w，只能约束速率函数 u(t)的时间积分；许多不同函数都可能有相同积分。增加内部锚点，等于增加若干子区间积分约束。仍未必得到唯一历史，但允许的历史集合可以缩小。这个数学事实说明，海底空间记录为何能够改变地史提问的粒度，同时也说明为什么有限条带不能恢复任意精细的运动史。

若在资料不允许时追求过细时间分辨率，就可能把噪声拟合成速率起伏。一个参数较多的模型通常更容易追随曲线，却不必更接近实际过程。评价时应同时考察拟合改善、参数增加与独立检验。不能用更多自由度换来的图形贴合，直接宣布发现了新的地球动力节律。

不对称也应采用相同纪律。只有在参照、几何、边界对应和时间约束得到适当检查后，宽度差才更有资格参与不对称扩张的判断。把所有不对称预先归给观察误差，会使模型无法发现真实变化；把所有不对称都解释成自然机制，又会把操作问题写进地史。两种预设都提前关闭了研究。

本书称这种状态为有边界的开放：允许多个解释，但要求每个解释承担下一步证据责任。几何问题应当提出可以检查的测线或方向条件；速率问题应提出与内部时间约束相容的预期；对应错误应说明另一种匹配如何接受新资料。不能只列出一串可能性，然后把无法判断称为复杂。

磁条带的编号还承担一种容易被忽略的同一性工作。两条剖面上各有一个突出的峰，并不意味着它们代表同一极性时段。真正的对应要比较前后邻接序列及其相对宽度，而不

是只按峰高配对。一个强峰可能由较厚磁化层形成，另一个较弱峰却代表同一年代。若只抓最醒目的形态，便会把振幅相似误当成时间相同。

可以用文字序列作教学类比。单个字在许多句子中出现，长一些的字序列则更有辨识力；但缺字、重复和不同间距会使配对复杂。磁性序列也需要利用上下文，不过它不是自然已经写好的文字，而是地磁历史与地壳形成过程共同留下的空间记录。类比的价值在于强调序列约束，不能把地磁反转想成有意编码。

空间分辨率还会把很短的极性时段合并或削弱。某条窄磁化带即使真实存在，其远场信号也可能不足以在测线上形成清楚峰谷。因此，观测中缺少一个小特征，不必立即判为年代序列不存在；但允许缺失的范围要由测量高度、源层深度和信号强度限制。若任意允许所有不合的条带消失，时间对应便失去检验力。

另一项重要检查，是速度调整是否具有独立理由。允许每条带分别选择不同速度，几乎可以把许多宽度序列拉伸到同一时标；限制速度只在有证据的阶段改变，则保留了模型风险。较好的对应不是调节后看起来最像，而是在较少、受约束的历史参数下仍能解释长序列，并对邻近测线提出新的预期。

因此，条带对应应保存多个候选，而不只展示最后选中的方案。若几个候选在现有资料下同样可行，就应报告这一歧义，并寻找最能区分它们的新测线或年代点。唯一性不是绘图时选定一组编号便获得的，而是替代对应在新增约束下逐渐受到排除的结果。

一旦这些条件公开，条带便不只是展示海底历史的图案，而成为能够由不同团队重新计算、比较并修订的时间对象。它的稳定性来自共同操作，而非图像自身的直观力量。

◆ 27.4 时间编码成立的证据门槛

把海底理解为时间记录，不是看到一组重复图案就足够。重复可以来自多种过程，时间对应需要额外约束。本章至少要求区分三个判断：某种磁性模型能够产生相似信号；某组条带能够与时间序列相容；这个对应在新资料下仍具有排除竞争解释的力量。三个判断逐步增强，不能从第一个直接跳到第三个。

以一个现代方法论推演说明独立检验。先用部分测线建立模型并固定关键参数，再考察未参与调整的邻近测线。如果预测失败，应先检查模型预期的适用范围，而不是直接把

新资料排除。如果预测成功，也要检查训练与检验是否共享了未说明的调整，例如两者的条带边界是否都按同一预期人工选定。形式上分成两组数据，不保证推理上相互独立。

还可以比较不同类型的资料。磁异常、地形、年龄和地震分别提供约束，但它们的结果可能共享同一地理定位或同一年代框架。证据之间有联系并不使其无效，只要求评价支持强度时不重复计数。四张图若都由同一假设导出，就不能当成四份相互独立的验证。

在这里，最强的竞争解释不应被写成“旧人什么也不懂”。局部磁化、岩性与地形因素本来就能够解释部分异常。新的扩张与倒转解释，需要说明它新增了哪些跨区域与跨时间的约束，而不是否定一切局部因素。新理论往往通过重新安排旧知识的有效位置而成长，未必通过宣布旧知识全错。

本书提出的发生学判断因而是有条件的。当磁性、空间几何与时间约束共同形成可重复辨认的序列，研究者得到了一种此前缺少的对象能力：海底的某些空间差异能够被用于追问形成先后与运动。这个新世界具有物质约束，也具有新的视觉与数学操作；它的成立不是只在脑中换一幅图。

这里改变的是海底记录的可用方式。自然中的磁化与地壳运动早已发生，研究者通过磁学正演、空间配准和年代相关，将其保留的痕迹组织成可比较的历史序列。序列能够支持的结论受分辨率限制：长波异常可能保存较长时段的轮廓，却不足以恢复每一次短暂变化。科学对象的形成因而包含新操作和新推断资格，也包含对这些资格的限定。

常规模型检验已经能够处理参数、独立资料和预测失败。发生学分析在此继续追问：一条磁异常在何种操作后获得“同一极性时期”的身份，这种身份又怎样改变测线与样品的选择。前者可由对应规则与时标约束分析，后者需要接续调查的历史记录。两项证据不应混用；仅凭磁性模型拟合成功，还不足以完成关于研究实践如何改变的解释。

本章的教学推演还允许形成一个清晰的失败树。若曲线本身不稳定，先检查记录与处理；若曲线稳定而地下解释多解，增加几何与磁性约束；若地下结构较稳而时间对应不稳，检查外部时标与序列身份；若时间对应成立而速度冲突，检查运动方向与模型范围。这不是一种保证成功的算法，而是避免把不同失败混成一个“理论不对”的分析次序。

失败定位也改变了保存资料的方式。只有最后一张漂亮示意图，无法让后来者分清上述问题；保留测线、处理参照、候选对应及被拒绝模型，才可能让后来的检验找到真正需要修订的环节。知识传播因此不只是传播结论，还包括传播足以重新发生判断的条件。

这一要求与本书全书命题相接。一个稳定的科学世界，不是再也不能修改的成品，而是在修改时仍能指出哪些关系继续有效、哪些结论必须撤回的系统。磁条带解释的价值，不需要靠删去全部不确定性来维护。恰恰因为它能把不确定性分到不同操作与对象上，研究才获得继续前进的方向。

海底磁条带让空间获得了受约束的时间含义，却尚未单独给出整个运动网络。下一章先把磁性记录与基底年龄、沉积和地壳消耗相联系，检验海底扩张能否承担多项共同责任；随后再由转换断层讨论具有方向判别力的边界证据。

第 28 章 海底扩张
一个假说怎样承担多项独立检验

28.1 从概念示意到约束命题

一幅海底扩张示意图通常只需几笔：中央有新物质上来，两侧向外移动，远处某些地方向下进入地球内部。图很容易理解，也容易使人误以为问题已经解决。实际上，每一笔都承担不同责任。中央为什么能够持续生成，离开中央的物质怎样保存历史，两侧运动怎样量化，新增面积到哪里去，都是需要独立约束的问题。示意图的价值在于把责任连接起来，而不是用连接代替证据。

Hess 的《海洋盆地的历史》发表于 1962 年，是海底扩张论证的重要原典。[R28a] 本章把它作为提出综合问题的历史节点，不把后来的深海钻探结果倒装为作者当时已经拥有的证明。扩张思想的新意之一，是把海洋底部从长期被动容器改写为可能持续更新的系统。大陆的相对分离由此不再只能理解为大陆穿越一个固定洋底，而可以与洋底本身的生成和迁移联系起来。

这一改变立即产生物质与面积账本。在近似恒定大小的球面上，某些边界不断增加表面面积，另一些边界就需要消耗或重新安排面积。否则，系统必须承担地球半径长期变化的后果。把海沟纳入消耗机制，意味着它不只是海底最深的一类形态，而可能是物质循环中的特殊位置。形态因而获得运动意义，但这种意义需要地震、地壳结构及其他记录支持。

扩张并不要求每一处海岭都完全对称，也不意味着海岭位置永远不动。最简单模型常把海岭视为固定中央线，两侧以相同速率离开，这有利于建立初始预期；真实重建还可能涉及不对称增生、海岭迁移、边界调整和局部变形。复杂化是否合理，要看它增加了多少有独立约束的过程。不能用最简单示意图直接否定一切偏离，也不能用“真实世界复杂”保护所有偏离。

思想史上还必须区分漂移与扩张。漂移关注大陆相对位置改变；扩张关注洋壳如何形成并离开生成区；板块体系则进一步规定含洋壳与大陆的运动单元及其边界关系。这些概念彼此支持，但并非同一个词的不同版本。扩张能够解释大陆分离的一部分几何，却仍需说明全球运动如何相容，以及边界如何连接。

从对象发生看，洋壳不再只是某处海水下面的岩石，而获得了生成位置、离岭历史、年龄和去向之间的联合身份。这样的身份可以提出测量任务：沿离岭方向取样，寻找年龄变化；比较基底与覆盖沉积的时间关系；检查消耗区域的运动记录。一个理论如果不能从图示走到这些任务，便仍停留在有启发性的构想。

海洋地壳与岩石圈也要分开。地壳主要按成分和地震结构区分，岩石圈则与较冷、较强的力学层有关。海底形成后冷却，力学性质可以随年龄变化，而不意味着地壳成分边界以同样方式改变。扩张示意图常把它们画成同一层带，便于入门，却会掩盖后续热演化和运动单元的问题。板块不是只有薄薄一层海底玄武岩。

冷却还提供与年龄相关的物理预期。热的物质冷却后会改变密度和厚度，海底深度及热流因而可能与离开生成区的时间有关。这里的关系不能单凭年龄决定，因为地壳厚度、火山活动、沉积和动力支撑都会影响地形。它的意义在于增加另一类受物理约束的比较：年龄模型不仅要与磁性相容，还应与合理热历史不冲突。

◆ 28.2 年龄梯度作为预测

最直接的局部预期，是在简单持续扩张条件下，离开海岭越远，洋壳形成时间越早。这里的距离应沿运动方向或经过适当几何处理，不能直接用地图上到当前海岭的最短距离替代所有历史路径。海岭若迁移或跳跃，当前位置与过去生成位置不同；板块若旋转，直线测距也需要转换。因此，年龄梯度既提供检验，也要求明确重建模型。

设一侧海底以恒定速度 u 离开海岭，形成后经过时间 t，则距离 x 约为 ut，年龄可写成 t=x/u。这是本书的教学关系，不是某次航次的原始计算。它显示，只要由外部资料约束速度，便可以预测取样点的年龄；反过来，独立年龄又能估计平均扩张速度。若同一组年龄既用来决定速度，再被称为速度的独立证明，就重复使用了信息。

速度变化时，距离应理解为不同时段速度的累积，而不是一个固定比例。若年轻部分扩张较快，年龄—距离图的斜率会改变；若某阶段停止或明显减缓，空间条带会变窄。不同历史能够产生相近的总距离，因此一个单点年龄通常不足以恢复完整速度史。多个点的价值在于限制变化的形状，而不只是增加平均值的精度。

取样方式决定年龄命题的身份。直接研究基底岩石，可能更接近其形成过程，但蚀变与后期作用会影响某些测年体系；研究最靠近基底的沉积，则提供与基底形成有关的时间限制，却不一定等于基底精确年龄。沉积可能在基底形成后延迟开始，也可能存在间断。把最老沉积年龄直接命名为洋壳形成时间，会把需要说明的时间差隐藏起来。

深海钻探把遥测曲线与可以保存、重新鉴定和多方法分析的样品联系起来。DSDP 第三航次发生于 1968 年末至 1969 年初，正式报告体系包含站位报告、岸上研究和年代框架。[R28b] 钻探由此加入了海底扩张提出之后的实物证据层。磁曲线描述的是一定体积内磁性的综合响应，岩芯则把局部岩性、接触和层序置于可反复检查的材料上；两者需要通过站位与地质模型连接，不能仅因来自同一海域就直接相互替代。

岩芯也不是自然送来的一根完整年表。钻进可能漏取，回收率不足可能造成空段，软沉积扰动可能影响层位，硬基底进入深度可能有限。样品编号、孔深与岩芯内部位置必须对应。若只保留一串化石年龄而丢失取样位置，就无法判断年龄关系是否连续。物理样品的可复核性依赖档案链，并不因为“摸得到”就比遥测自动更可靠。

最强的预测性检验可以这样组织：先依据磁异常或运动模型给出一组站位年龄范围，再对钻探样品进行独立生物地层或其他年代判断，最后比较差异。需要记录预测在取样前是否已经存在，以及鉴定者是否利用了预测。完全盲测在历史研究中未必做到，但信息依赖至少应被说明。否则，模型与年龄互相调整后得到的一致，只能称为联合拟合。

现实海盆的年龄分布不必处处严格单调。海山火山活动可能在古老基底上形成较年轻岩石；边界迁移可能改变离岭关系；俯冲会删除部分记录；被截留的旧洋壳可能处于复杂位置。这些情况不自动否定扩张，却要求区分“样品年龄”“当地基底年龄”和“简单扩张路径年龄”。解释异常时，必须寻找独立地质依据，而不能只因其不合模型就重新命名。

年龄梯度的检验力由此清楚起来：它不是一句“远处较老”的宽泛印象，而是在规定的运动单元、形成过程与取样条件下，年龄随距离出现何种可预期关系。若多项独立年龄长期与预测相反，且无法由有证据的边界历史解释，模型必须修改。把失败条件写清，才能理解钻探为什么为扩张思想增加实质支持。

年龄误差需要传播到速度，而不能只放在图表脚注。若距离较可靠而年龄区间较宽，速度也只能给出范围；若年龄接近零，简单比值会对小误差非常敏感。多个点拟合能够减少部分随机误差，但不能消除共同的时标偏差。研究者必须区分点与点之间的散布和所有点一起移动的系统误差，因为它们对速度历史的影响不同。

两侧同时取样还可以检查对称性。相同年龄在海岭两侧到生成轴的距离若不同，可能说明不对称扩张，也可能说明海岭位置变化或边界识别问题。判断需要结合磁性序列和区域几何。对称是简单模型的预期，不应变成整理资料时必须制造的结果。两侧年龄各自可靠而距离不对称，正可能成为改进运动模型的重要入口。

站位选择也决定检验强度。若所有钻孔都靠近同一年龄段，即使结果一致，也主要检查局部关系；沿足够宽的离岭范围布置站位，才能限制梯度和变化。复杂区域可以提供高风险检验，却需要更详细重建；简单区域便于隔离基本机制，却不能代表所有海盆。两类取样具有互补作用，不能只选择最容易成功的位置来声称普遍成立。

把预测年龄与实测年龄画在同一图上时，还应注意误差方向。点落在对角线附近不一定意味着完美一致，若误差很大，许多模型都能通过；误差很小的系统偏离反而可能更有信息。检验的力量来自模型之间可区分的范围，而非图上点数多或线条整齐。一个少量但高精度的关键站位，有时比大量低区分力数据更能改变判断。

◆ 28.3 沉积厚度与保存偏差

旧海底往往拥有较长时间积累沉积，因此沉积厚度可以参与年龄讨论。但厚度不是时钟。沉积供给多的年轻区域可以很厚，供给少的古老区域可以很薄；底流侵蚀、溶解和不沉积会造成缺失；地形低处还会汇集来自周围的物质。若把厚沉积直接等同老洋壳，就把累积时间、供给速率与保存条件压成了一个变量。

最简单的教学关系是厚度约等于沉积速率乘时间，但真实剖面需要考虑压实。下部沉积受上覆重量压缩，当前厚度不等于最初沉积时的体积厚度。粒径、含水量和成分不同，压实程度也不同。即使两处沉积经历相同时间，最终厚度仍可能不同。因而，厚度梯度最好与年代层位和沉积环境共同解释，而不是独自承担扩张速度。

沉积还具有横向搬运。浊流可以把较老物质搬到较年轻沉积层中，化石也可能被再搬运。一个样品里出现老化石，不必意味着整个层段同样古老。鉴定需要结合保存状态、组合和层序，判断化石是否原地沉积。再搬运并非罕见到可以忽略的逻辑可能，而是海洋沉积研究必须面对的过程类型。理论检验的质量取决于这些具体判断。

基底与上覆沉积的接触具有特殊意义，因为它限制了海底形成后记录开始保存的时点。但接触也可能复杂：岩浆侵入沉积可以形成较年轻岩体，后期熔岩覆盖旧海底可以造成新的接触，侵蚀可以剥去早期沉积。因此，识别基底不能只凭“钻到硬岩”。岩石结构、接触关系和区域背景共同决定样品究竟代表哪一次事件。

这种细分让多源证据关系变得更准确。磁异常年龄可能依赖某套反转时标，生物地层年龄可能依赖另一些定年与相关，基底同位素年龄又有不同误差。它们并非绝对独立，也不是完全同一。研究者应追查共享校准在哪里，并特别重视不同误差来源仍得到相容结果的部分。独立性是程度和结构问题，不能以“用了三种方法”一票解决。

俯冲造成的保存偏差则影响全球洋壳年龄分布。今天可见的洋壳并不是历史上所有洋壳的完整样本。若旧洋壳被持续消耗，幸存记录就偏向较年轻部分。于是，缺少极古老洋壳不必意味着古代没有海洋，也不能直接用来给海洋本身定一个年轻诞生时间。生成与消耗共同规定了今天能够取样的总体。

从统计上说，这是观察总体经过选择的情形。只有幸存的部分进入样本，选择机制又与年龄、位置和运动历史相关。不能把样本分布当作无偏的全球历史。思想上的重要改变在于，记录缺失不再只是调查不足，也可能是理论需要解释的循环结果。海洋世界的身份因此包含不可见部分的受约束推断，但这一推断必须与消耗证据相容。

沉积资料的价值最终在于丰富时间与环境限制。它能够告诉我们某段海底何时开始积累何种物质、经历怎样的保存过程，并与基底和磁性记录相互检查。把它降为“厚就是

老”的辅助图，会浪费它的独特信息；把每个厚度差都当作扩张反例，也会忽略沉积过程本身。成熟检验需要同时保持运动模型和沉积学的责任边界。

沉积速率还可能随海底所处环境改变。洋壳从海岭附近移向其他纬度、深度或陆源物质影响区，其上沉积不必维持同一种速率。若把整个岩芯用一个平均速率换算时间，就会抹去环境迁移。年代控制点能够把厚度分成若干阶段，进而判断厚度变化来自时间长度还是供给与保存变化。运动史与沉积史因此相互限制，却不能互相替代。

化学溶解会改变某些生物沉积的保存。在不同水深和水体条件下，同一种碳酸盐组分可能保存较好或更易溶解。一个层段缺少某类化石，可能是生态缺席，也可能是沉积后损失。用于年代鉴定时，需要结合其他化石或材料检查。海洋时间记录的建立，不是从岩芯中直接读出年份，而是让生物、环境和保存特征共同支持某种年龄身份。

不整合与沉积间断尤其影响基底上方的时间解释。设距今年龄以正数表示，确认属于原位覆盖层的最老沉积年龄为 t_s，那么在接触关系可靠、且无后期侵入等干扰时，基底形成年龄至少应不小于 t_s。它提供的是基底年龄的下限，而非自动给出形成时刻；早期沉积缺失越多，这一下限离实际年龄就可能越远。若接触附近的年代与独立基底定年相近，并且沉积环境支持较小间断，才可进一步将其用作近似形成时间。

时间分辨率还限制了扩张速度的解释。由相隔很久的两个年代点得到的速度，是这段时间的平均量，可能掩盖其中的加速、减速或短暂停顿。更密集年代点可以提高分辨率，但也可能放大局部误差。研究者要在历史细节与可靠性之间作出有证据的选择，不能以曲线越曲折就代表历史越真实。

取样年龄还需与速度定义相匹配。一侧样品的年龄—距离斜率限制该侧相对生成轴的离开速率；两侧同龄带的间距则限制它们在该时段的累计分离。两侧同龄带的间距除以共同年龄，给出该时段的平均全扩张速率；在对称且海岭参照适用时，单侧平均离开速率为这一全速率的一半。不对称情况下，应分别估计两侧再组合。把单侧距离直接除以全扩张速率，会把同一年龄估小；这种定义错误不会因增加钻孔而自行消失。

海底扩张的认识因此经历了从可能图景到可计算系统的变化。最初示意给出生成和消耗的方向，磁性与年代提供时间尺度，钻探与沉积检查历史身份，全球运动约束则限制各

部分能否同时成立。没有哪一项单独完成全部工作，但每一项都改变了其他证据可以支持什么。

由此，一组年龄异常应分层判断。若只有某个岩芯层位不合，可先查取样与保存；若同一海域所有年龄共同偏移，应查时标与基底识别；若偏离随离岭距离系统变化，则速率历史或生成轴位置可能需要修改。这样的区分保留了沉积学的独立责任，也使扩张模型面对与其运动假设直接相关的检验。

这一过程还使海洋底部取得了不同于大陆地表的历史身份。它的可见记录受到持续生成与消耗共同筛选，因而不能用大陆古老岩层保存的方式直接类比。理解这种记录机制，是解释海洋年龄结构的必要一步。

缺失的早期记录也会影响不同海盆之间的比较。一个海盆保存的最老年龄较小，可能反映海盆本身较晚打开，也可能反映较老部分已被消耗。要区分两者，必须将现存年龄序列与边界历史合看。

◆ 28.4 竞争模型与总账闭合

海底扩张不能只与一个毫无修补能力的固定洋底作比较。固定模型可以提出广泛更新，地球膨胀模型可以用面积增加容纳新生海底，其他方案也可以引入不同的内部过程。比较时，必须让每个模型承担同一组问题：年龄分布、磁性序列、海岭生成、地震与消耗位置，以及全球几何关系。只挑自己最擅长的一项，不能证明整体优势。

地球膨胀在逻辑上可以减少对表面消耗的需要，但会增加半径变化及其物理机制的责任。它还必须解释那些显示会聚与向下运动的资料。即使某些局部扩张现象与膨胀相容，也不能因此跳过全球收支和地震约束。相反，近似固定半径的板块体系必须说明新生面积在哪里被抵消，而不能只展示海岭附近的成功。

面积收支可以写成一个简化的教学关系。沿海岭，把两侧相对于生成轴的法向离开速率相加，得到局部张开速率，再乘对应边界长度并沿边界累加，便估计新增面积速率。沿消耗边界，则须计算待消耗板块相对于该边界的法向进入速率；海沟迁移时，它不必等于两板块的全部相对会聚速率。两类量单位都是面积除以时间。球面几何、分布式变形和边界变化会增加计算环节，但比较生成与消耗的总体责任保持不变。

这里的闭合不是要求每个短时间窗口都精确相等。地球表面存在复杂变形，观测与重建也有误差。重要的是误差范围内能否形成受约束的整体，而非为每次不平衡临时增加不可观测区域。总账的力量在于把局部成功互相绑定：某海岭扩张速率调高，会影响其他边界需要承担的消耗；一个局部修订因而不能只在本地获益而不付全球代价。

SDE 的分析增量在于突出这种对象间的互相规定。洋壳年龄限制运动历史，运动历史规定取样位置，新增样品又改变年龄相关与沉积解释。E 不是“海洋环境”这个笼统名称，而是年龄、保存、运动和消耗特征之间的约束结构。新的洋壳世界能够运行，是因为这些关系让不同团队的工作接得上，也让局部错误能够被别处发现。全球收支还涉及边界长度与速度方向。沿一条边界的全部相对速度，并不都用于生成或消耗面积；只有垂直于边界的分量承担这一作用，平行分量主要表示侧向运动。因此，同一总速度在不同边界方向下，会对应不同面积变化。这个几何事实把边界分类与速度测量绑在一起，也说明仅凭箭头长度无法完成全球账本。

若边界本身迁移，还要明确速度相对于什么。海沟后退、海岭迁移或其他边界变化，会影响物质穿过边界的速率与地理位置变化。一个固定在地图上的观测者，与跟随某板块的参考系，会报告不同的绝对速度，但相对运动关系应能转换。把参考系写清，可以避免把坐标选择造成的差别误认为物理矛盾。

海底扩张获得强支持，并不要求每一项辅助现象都只有唯一解释。局部热流高、局部岩石年轻或局部沉积薄，都可能由别的过程造成。真正的优势在于这些现象在共同空间与时间安排下相容，并受到独立记录检查。理论强度来自联合结构的约束，而不是把每个单项现象宣布为专属证据。

这一体系仍未解决全部驱动力问题。运动学与收支自洽，说明某种运动安排能够同时容纳证据；它并不自动说明深部哪种力贡献最大。海底扩张的成功应按它实际解决的责任评价，不必为了显得完整而吞并未解的动力学。下一章将转向边界类型：横切海岭的断裂线怎样成为转换断层，为什么运动方向的检验能在两种形态相似的解释之间作出区分。

第 29 章 转换断层
新运动关系怎样改变边界的含义

29.1 为什么形态容易误导

在一张地图上，两段海岭相互错开，中间被一条横向结构连接。最直接的阅读方式，是设想一条原本连续的海岭遭到切断，两侧沿断层错动，于是留下今天的偏移。这种想象有其熟悉的地质类比，但“能够想象出一个过程”与“材料证明发生了这个过程”不是同一回事。

地图展示当前位置，运动判断则要说明位置如何改变。两者之间需要模型。一条线的形状可以相同，线两侧材料的相对速度却可能不同；同一个表观错距，也可能由不同生成与运动历史形成。若只从形态直接读出方向，就把待证明的历史放进了最初观察。

先用一个平面教学模型辨明方向。设北段海岭轴位于较西位置，南段海岭轴位于较东位置，两轴之间由一条近东西向线连接。假定两段海岭近于南北走向，在共同参照下两侧对称生成并向东西分离。那么，连接段以北处在北段海岭的东侧，材料向东运动；连接段以南处在南段海岭的西侧，材料向西运动。连接区两侧的相反方向由它们相对于各自海岭的位置确定。这一设置用于教学推演，并非威尔逊原图的复现。

而在连接区向西的延伸部分，上下两侧可能都处于各自海岭的西侧，具有同向速度；向东的延伸部分也可能同样如此。相同的地形线因此可以分成两种运动地位：中间一段分隔相对运动的材料，两端延伸记录不同历史，却未必承担同样的现时相对运动。

此推演使用刚性、平直和对称等简化假设。它不是说真实海底外延段绝无地震或变形，而是说明一种理想运动学区分：地貌连续性不必等于活动边界连续性。任何把现实偏离简单归为“异常”的做法，都需要重新检查理想模型是否适用。

思想转折的关键由此出现。过去我们可能问这条断层把海岭错开多少；现在首先要问它连接哪些生成边界，两侧分别属于怎样的运动单元，哪里存在相对滑动。问题改变之后，原来的一条地形线获得内部区分，所需证据也随之改变。

这正是对象身份重构的一个候选情形。新对象不只是旧对象的新名字，因为它规定了不同的活动区段与运动预期。如果两个名称导出的预测完全相同，名称更新当然不足以构成强创新；如果相同地形在两种模型中导出不同方向，就有可能通过新资料进行判别。

还可以把前述模型中的速度写成符号，减少仅凭图形直觉的歧义。令向东为正，在海岭之间的连接段，北侧速度为＋u、南侧为－u，则北侧相对南侧的速度为＋2u。在西侧外延段，若上下两侧都为－u，则相对速度为零；东侧外延段同理。这里的零只属于设定的理想模型，不能无条件推广为现实中绝不变形。

这个推演揭示一项重要差别：地形落差或材料年龄差可以在相对速度为零时仍然存在。因此，识别不同材料与识别运动边界是两种任务。把它们分开以后，研究者才能解释为何某条连续地貌线的不同区段，应当接受不同的地震与运动检验。

再看“原来连续、后来错断”的解释。在同一平面图上，北段现位于较西处，南段位于较东处；若把两段视为原先对齐的标志，并将现有错距全部归于随后沿连接线的错动，便推得北侧相对南侧向西，正好与扩张模型所给的向东相反。这个对照依赖标志先连续、错距由滑动形成的假设，说明两种模型为何能对同一形态作出相反方向预期。Sykes 的后续论文正以这类相反预期进行判别。[K06]

◆ 29.2 威尔逊命题的结构

威尔逊 1965 年论文《一种新的断层类别及其对大陆漂移的意义》，提出转换断层这一类别。[B04] 其概念要点可以通过边界连接与相对运动来理解：两侧材料怎样运动，取决于断层所连接的生成或消耗边界及其所属单元。后续震源研究则把这种运动关系转成方向检验。[K06] 以下继续采用教学重构，区分概念提出的责任与后来检验所增加的证据。

一个新断层类别若只是几何分类，新增内容或许有限。转换关系更重要的地方，是把局部断层置于边界网络中理解。连接处的运动，不仅由该线自己的走向决定，还受它连接的两端边界及两侧相对速度限定。这使“断层是什么”与“整体怎样运动”成为相互限制的问题。

在局部平面近似下，设边界两侧速度为 $v_{甲}$ 与 $v_{乙}$，则相对速度 $\Delta v = v_{甲} - v_{乙}$。若 t 是沿边界的单位切向，n 是在地表内垂直边界的单位法向，理想转换关系要求

$\Delta v \cdot n \approx 0$，而 $\Delta v \cdot t$ 可以非零。它规定的是两侧主要沿边界相对滑动；局部弯折与非理想变形仍可能产生张开或挤压。这个分解把方向预期写清，也说明一条横向延伸的地貌线未必就是转换边界：还须判断其两侧是否具有相应的相对速度。

这个判据还有一个容易忽略的后果。整体一起平移，不改变两侧相对运动。地图上的绝对箭头可以随参照系变化，相对滑动的某些判断仍然保留。因此，研究者必须说明谈的是绝对位置、相对于某参照的速度，还是边界两侧相对速度。参照不清，箭头画得再整齐也不能完成论证。

回到前节的教学模型，两海岭之间的运动由两侧生成方向共同限定。外延部分即使具有地形差别，也可能属于同一理想运动单元内部。它可以保存年龄或热状态差异，但这些差异不能直接等同于正在发生的边界滑动。“历史留下的痕迹”与“当前运动的界面”从此被拆开。

在 SDE 分析中，S 是具有明确连接与运动条件的边界身份；D 是对几何、方向和活动分布进行的比较；E 是生成边界与相对速度之间的相互规定。若任意改变一段海岭的位置或两侧速度，连接线的运动意义就可能改变。因此 E 不是在已经完成的断层旁边补上一个大环境，而是参与构成断层的具体身份。

这也解释了为什么新概念可能改变下一步调查。若目标只是确认地形线，增加测深也许足够；若目标是判定滑动方向，就需要能约束运动的资料。模型把“看得更清楚”变成“检验哪一个方向”。这是一种研究路线的转变，不能靠重复同一种观察自动完成。

不过，逻辑上需要新的资料，不意味着历史中所有调查都由这个模型直接引发。某些地震记录可能先已存在，后来被重新分析；某些新观测可能有独立目的。要写清思想如何回写实践，必须区分新采集、旧资料重用与事后解释。三种路径都可能重要，却具有不同的历史因果。

运动边界还包含尺度条件。在很小的范围内，真实断裂可能分叉、弯折或形成复杂变形带；较大尺度模型则可能把它近似成一条线。这种简化是否合理，要看正在回答的问题。用于比较整体相对方向的模型，不必包含每一处小尺度破裂；用于解释局部危险或复杂断裂过程时，同样的简化可能不够。

同一边界的尺度改变时，检验命题也需随之改变。区域尺度的走滑方向可以稳定，局部弯折处仍可能出现挤压或张开；只要这些局部变形未推翻两侧总体相对运动，便不构成对转换关系的直接否定。反过来，若沿主要区段持续出现不能由弯折或误差解释的法向分离，就需要重新检查边界分类。被省略的细节是否重要，取决于它是否改变所检验的速度分量。

边界分类因而不应被写成给自然贴上永久标签。分类的科学价值在于它使某一尺度上的相对运动与可检验预期稳定下来。当尺度改变或新的变形证据出现，分类的适用范围可能需要调整。新对象并非没有历史的固定实体，它是在特定约束下能够维持的研究单位。

边界活动区与地形遗迹的差别，也可以从两侧物质的运动历史理解。在两段错开的海岭之间，断裂两侧可能属于不同运动单元，因此存在相对滑动；超过海岭端点以后，断裂带两侧可能随同一板块共同运动，保留的是年龄与地形差异。这样，一条视觉上连续的线可以包含不同运动身份。对象定义由关系决定，不能仅由线条是否延长决定。

这也改变了对海岭错距的理解。普通错断图像容易让人想象一条原本连续的海岭后来被断层切开并搬走；转换关系则允许海岭段在不同位置持续生成，而连接区的相对运动由两侧扩张规定。两种图像可以产生相似地形，却对滑动方向与活动范围作出不同预期。科学判别正发生在这个形态相似而过程不同的位置。

年龄差进一步提供检查。断裂带延伸区两侧可能具有不同形成时间，因此冷却程度和地形高度不同，即使那里不再承担两板块当前的相对滑动。这说明地形陡坎不必全部解释为正在活动的断层位移。时间记录、热演化和运动关系结合后，某些看似直接的地貌解释需要改写。转换断层的新增对象能力，包含区分活动边界与继承结构。

由此，预测地震分布时不应只问断裂线上有没有事件，而应问事件主要位于哪一段、机制方向如何、与海岭端点是什么关系。空间范围和运动方向共同提供判别，任何一项单独看都较弱。后来的震源研究之所以关键，正是它把这些几何预期变成能够接受独立检查的命题。

◆ 29.3 1967 年的震源机制检验

Sykes 在 1967 年的原始研究中分析了 17 次地震，其中 10 次位于与洋中脊相交的断裂带。论文摘要报告，这 10 次事件的走滑方向符合转换断层的预期，与把海岭简单错断的预期相反；地震的空间分布也参与判别。[K06]

这一结果的力量在于方向差别。若两个模型都只说附近会发生地震，发现地震就难以区分它们；若两者对相对滑动方向提出不同预期，震源机制便能够承担更强检验。科学证据的重要性因此不能仅用记录数量衡量，还要看它能否让竞争解释作出不同回答。

但震源机制也不是拍摄地下断层的照片。它从台站记录与波的传播关系中推断震源的运动特征。记录分布、极性识别和模型条件会影响约束强度。我们不能因为结果支持后来成功的理论，就把推断过程写成无条件直观事实。

震源机制解还包含断层面识别的问题。在双力偶的现代教学表示中，两个节面可对应同一个辐射型式，实际破裂面需由地震分布、地质走向或其他资料进一步判断。因此，机制图与地图线条的对应包含两步：先检查观测支持怎样的运动型式，再判断其中哪个面与候选断层相容。本文使用的是 Sykes 报告的方向性结果，不将其扩展为新的逐事件反演、拟合比例或置信区间。

对思想史而言，1965 年与 1967 年是两个不能合并的证据时点。前者的模型需要在当时已有材料下评价，后者的研究提供后来取得的检验。若把 1967 年的结果写成 1965 年提出模型时已经掌握的充分理由，就抹掉了预测承担风险的阶段，也无法说明后来的研究新增了什么。

反过来，后续结果也不能被描述成某个理论瞬间被全体科学家接受的证明。原始论文可以支持某项科学检验，却未必支持共同体接受史。要论证接受过程，还需要同期评论、会议争论、教材变化与具体研究计划。把科学证据的有效性、传播速度和人数变化混成一个故事，会让思想史失去层次。

构造模型的胜负与解释框架的评价需要分开。1967 年的方向结果支持转换关系，是因为它在竞争运动模型之间作出了区分。要进一步说明科学对象如何改变，还需追踪这项结果使研究者重画了哪些边界、重新使用了哪些资料，或者改变了何种后续调查。SDE

将这些操作变化与边界身份联系起来，但不能仅凭构造模型得到支持，就认定这条历史解释已经获得同等验证。

常规的科学史解释也可以完整叙述模型预期、地震检验与证据汇合。两种解释的比较，应落到它们是否识别同样的历史变化。例如，一份旧地震目录在新模型下被按活动段与外延段重新分组，数据未必增加，比较对象却已经改变；若新框架能够揭示这项变化的具体规则与后果，才提供超出一般“证据增多”的说明。这是需要档案支持的候选机制，不能替代对实际研究过程的追查。

目前可以确定的论证结构，是地貌错距经两种过程模型导出不同方向，再由震源机制提供判别。[B04][K06] 这使转换断层成为讨论对象重构的有力案例：被解释的仍是同一片海底，边界的身份却改由连接、区段与相对运动共同规定。原图、逐事件资料与后续工作记录将进一步限定这一判断的历史细节，尤其是预期如何形成以及结果怎样被不同研究采用。

方向证据与空间分布证据还可以在逻辑上互补。若某模型预测中间连接段承担主要相对滑动，而外延段在理想条件下不承担同类运动，那么事件发生在哪里与事件怎样运动，分别限制模型的不同方面。两者一致，比只满足其中之一更有力量；但如果两项判断来自同一定位或筛选假设，也应说明它们共享的条件。

“这里没有地震”则需要格外谨慎。没有记录可以表示活动较少，也可以表示观测时间短或检出能力有限。因而空间分布的支持通常应结合观测条件来读，不能把一个空白区域直接转译成绝对不活动。对地震目录的这种审查，与第 15 章关于未发现化石的分析具有结构相似性，但两者的具体误差机制并不相同。

化石记录与地震目录中的“没有”，因而只具有有限的结构类比。化石缺失可能来自生态、保存或露头条件，地震缺失则可能来自活动间歇、台站覆盖与检出阈值。共同的要求是将事件分布与记录过程分开；实际检验却必须分别进入地层保存和地震观测。跨学科比较只有保留这种差异，才有助于识别外延段地震空白究竟支持多强的运动判断。

还应追踪被拒绝的解释如何退出。一个模型与某项方向证据不符，不意味着它的所有局部知识都失效。普通断层研究中的许多观察与方法仍然有效；被限制的是把某种表观错

距直接转换为该类边界运动的解释。保留这种局部性，可以避免把理论转折写成旧世界全部消失、新世界从零开始。

对人物评价也应保持同样尺度。提出一个后来成功的模型，不意味着提出者同时掌握全部驱动力；对早期证据提出疑问，也不意味着批评者没有科学能力。思想史应当让读者在不同时间截点看见实际可用的资料，然后理解为何一种解释值得尝试、为何仍需要检验。后见之明最容易夺走的，正是当时的不确定性。

转换边界还提醒我们，相对运动不需要为每一侧指定唯一绝对速度。两侧共同增加同一个参考运动，相对滑动关系仍不变。真正需要由地震机制检验的，是两侧速度之差及其方向。明确这一点，可以避免把参考系选择误当成对断层性质的不同解释。

这种关系性的定义也使不同海区能够比较。只要边界两侧的运动与连接条件相当，即使地形外观不同，也可以提出同类检验；反之，外观相近并不保证对象相同。

边界的身份必须由可检验的相对关系维持，任何名称都不能替代这一责任。

◆ 29.4 从断层名称到板块关系

海岭之间的一条连接线，看似局部问题，却把科学对象的单位推向更大范围。只要判断这条线的运动需要知道两侧材料属于哪些运动单元，研究就不能停留在孤立断层。边界之间必须相容，局部速度也必须能够进入整体关系。

可以用一个简单的运动学恒等式说明整体约束。三个单元甲、乙、丙，在同一参照系中具有速度 $v_{甲}$、$v_{乙}$、$v_{丙}$，则它们的相对速度满足 $v_{甲乙}+v_{乙丙}+v_{丙甲}=0$。这个局部平面表达只是下一章球面运动问题的入口；它不是对全球板块运动的完整模型，也没有提供驱动力解释。

恒等式的意义在于，三段边界不能各自任意选择最方便的速度。如果逐段拟合都很好，组合起来却无法闭合，就需要检查对象划分、测量、参照或刚性近似。整体并非把局部成果装订在一起，而是对局部结论增加共同约束。

因此，新科学世界的发生包含一种新的失败能力。早先彼此无关的解释，现在可能因不能共同成立而受到质疑。能够指出这种不相容，是整体研究能力的重要部分。一个理论若只是把所有现象收进宏大名称，却不增加任何共同限制，它的整体性就缺少科学内容。

转换边界的思想价值，可以沿这条线理解：地形、相对方向与活动区段获得共同规定，研究者因而能够提出更精确的检验。海底世界不再只是许多高低起伏和裂线的集合，而成为能够通过运动关系区分对象的系统。这是本书所说的新特征纠缠聚合体系统的一种具体候选。

这一发生也包含主体能力的变化。研究者学习辨认的，不仅是地图上的形状，还包括哪些形状不能决定运动、哪些资料具有方向判别力，以及哪些局部结果不能组成一致整体。主体、互动操作与客体规定共同改变，才能形成可继续研究的世界。仅把一个新名词加入词典，并没有完成这个过程。

三式在这里分别指向不同材料：S 式关注地貌线何以获得转换边界的身份，D 式关注这一身份为何要求方向与区段检验，E 式关注检验后边界划分与后续比较是否改变。前两项可以从模型和方向性研究中讨论；第三项还需连接到图件修订、事件分组或后续调查。这样区分，既能追踪共同约束的形成，也避免用一次成功检验代替完整的历史回路。

按分析相位看，表观错距与运动方向之间的张力、针对这种张力提出新的比较、比较后重新划分活动区段，是三类不同变化。某项研究可能只完成其中一类，另一类由后来工作承担。相位划分的作用是辨别改变发生在哪里，而非要求一个案例具备整齐的三段过程；它也不替代边界运动学本身。

本章最后保留一个未解决问题：相对运动能够被描述，并不等于运动为何持续已经解释。运动学相容性与深部动力机制承担不同责任。下一章需要把边界纳入全球网络，后续深部地球章节还要面对物性、结构与驱动的问题。地球科学的新世界就在这些已成立与尚待成立的连接之间继续发生。

第 30 章 板块世界的成立
运动学闭合、共同检验与新任务

30.1 从大陆到板块的身份更替

大陆漂移与板块构造之间最重要的差别，不是后者为前者增加了更多证据，而是运动对象发生了改变。大陆由地壳性质和海陆形态辨认，板块则由一定尺度上相对稳定的内部运动与边界关系辨认。一块板块可以同时包含大陆和洋壳，一个大陆也可能跨越不同运动单元。若仍把板块画成大陆轮廓的同义词，就无法理解海洋证据为何能够重写整个问题。

板块中的“刚性”是一项近似，不是材料绝对不变形的宣言。它表示在特定空间和时间尺度上，板块内部两点间的相对位置变化较小，可以用共同运动描述，而主要相对运动集中在边界附近。内部地震、局部断裂与长期变形并不因定义而消失。问题在于它们是否小到不破坏该尺度的近似，还是已经要求划出新的单元或采用分布式变形模型。

这一对象定义改变了地图上的边界。海岸是海水与陆地的交线，板块边界是相对运动发生显著变化的区域，两者没有必须重合的理由。某些大陆边缘可能位于板块内部，某些边界则穿越大陆。研究者不再只看地表颜色，而要比较地震、断裂、扩张、会聚及其他运动证据。板块地图因此是一幅受多类资料约束的运动学几何图，而不是地理分区图。

边界也不总是一条无限细线。大洋某些边界可以较清楚地集中，广阔大陆变形带则可能把运动分散在多个断层与块体中。若为了画图方便，将宽阔变形区压成一条线，模型可能仍适合全球尺度计算，却不足以解释区域细节。需要明确地图的分辨率与任务，不能把简化边界直接用于所有地质判断。

由此，板块身份同时依赖内部和外部。内部近似共同运动支持把区域合为一块，边界两侧运动差异支持把它与邻区分开。只看边界线条而不检查内部一致性，可能把任意闭合多边形称为板块；只看内部若干点相似运动而不检查外部联系，也可能漏掉边界。对象成立需要这两种约束相互配合。

1967 年 McKenzie 与 Parker 的《北太平洋：球面构造的一个例子》，以及 1968 年 Morgan 的《海岭、海沟、大断层与地壳块体》，是球面运动学综合的重要原典。[R30a] [R30b] 从这些研究进入问题，关键在于看到运动单位与计算方式的共同改变：边界上的局部方向和速率，要由同一组球面旋转解释。论文的发表年份标记了公开论证的节点；思想交流与独立构思的先后则还涉及其他史料，不能仅用刊年排序代替。

板块名称的稳定也不能代替边界稳定。一个名称可以在不同模型中覆盖略有差别的区域，或者因分辨率提高而分成几个单元。比较研究结果时，应查看实际边界与参考时段，而不是只看名称相同。类似地，地质历史中的某一板块与今天同名板块，可能经历增生、分裂或边界调整。对象身份包含历史承接，不等于几何范围永远不变。

刚性近似可以通过内部相对运动检查。若同一板块上不同位置的运动能够由同一旋转在误差内描述，近似得到支持；若出现有组织的系统偏离，就可能存在内部变形或遗漏边界。这比先画好板块再强迫所有资料归属更严格。现代观测能够提供更直接检查，而古代重建则需要由地质记录和相对运动间接约束，二者的证据条件不同。

◆ 30.2 球面运动与网络闭合

在一张平面纸上推动一块图形，最直观的是平移。地球表面却是球面，一个近似刚性的表面块体相对另一个块体的运动，可以用绕穿过球心某轴的旋转描述。轴与球面的交点提供旋转极，转动快慢给出角速度。对不熟悉数学的读者，可以想象把地球仪固定，再让表面的一片薄壳绕某轴缓慢转动：薄壳上不同位置的线速度并不相同。

距离旋转极越近，转动一小角度所走的路越短；在与旋转轴垂直的最大圆附近，同一角度对应的路程较长。于是，同一对板块并不要求沿全部边界具有相同线速度。速度方向也随位置改变。若把距相对旋转极具有不同角距的地点所测扩张速度，不加球面几何处理地比较，可能把共同旋转本来允许的变化误认作板块内部不一致。

数学上，以地心为原点，角速度向量 ω 与位置向量 r 给出局部相对速度 $v=\omega\times r$。设地球近似半径为 R，该点与所取旋转极之间的球面角距为 ψ，则速度大小为 $|v|=|\omega|R\sin\psi$；角速度用弧度每单位时间、R 用长度表示时，结果就是长度每单位时间。两极附近

线速度小，距旋转轴最远的大圆上最大。速度方向垂直于 ω 与 r 张成的平面。这些都是刚体运动学关系：它们把相同角速度转换为各地不同线速度，并未指定维持运动的力。

转换断层提供方向约束。若两侧主要沿断层相对滑动，断层走向应与相对速度相容。在球面上，这意味着相关走向可以帮助限制相对旋转极。海岭扩张资料提供速度大小的约束，震源机制又提供另一些方向信息。不同资料不是重复说明“板块会动”，而是分别限制同一个旋转模型的不同部分。它们是否相容，可以成为模型的检验。

Morgan 的原文将转换关系推广到球面，并以相对旋转组织边界与速度，还讨论了通过角速度向量闭合估计其他相对运动。[R30b] 这里最深的改变，是局部边界不再各自拥有完全自由的运动。若它们属于同一对近似刚性板块，就必须服从共同旋转关系。某处断层走向或扩张速率一旦确定，会限制远处同一边界上的预期。这使全球综合具有了真正的约束，而不只是把很多图放在一起。

三个板块之间的闭合，可用同一地心坐标中的瞬时角速度说明。记 ω_AB 为 A 相对 B 的角速度，采用一致符号约定，则 ω_AC＝ω_AB＋ω_BC，从而 ω_AB＋ω_BC＋ω_CA＝0。这些量须对应同一时刻或相容的平均时段。实际估计若留下非零剩余，应检查测量误差、时间窗口、符号与刚性近似。它的含义很具体：A 相对 B 与 B 相对 C 一旦确定，A 相对 C 便不再是一组可自由另选的参数。

这个零和条件十分有力，因为它允许不直接测量某一段边界，也能由其他关系给出预期。随后取得的新资料便可以检查闭合。若每一段边界都独立拟合，几乎总能获得局部不错的结果；要求整个回路相容，会显著减少自由度。一个局部模型为了提高拟合而改变速度，可能破坏另一个区域的关系。理论因而获得了互相追责的网络结构。

必须区分瞬时角速度与长时间有限旋转。瞬时向量在同一参考系中可按相对运动关系相加；有限旋转的顺序通常不可交换。先绕一个轴转动，再绕另一轴转动，与相反顺序未必得到同一最终位置。古板块重建若跨越很长时间，需要按阶段组合旋转，不能把所有角度像普通数一样相加。教学上忽略这一点，会使现代速度描述与古地理重建混淆。

时间窗口也是闭合条件的一部分。磁异常可能给出某段地质时间的平均速度，地震机制反映现代或较短时段的运动，长期地质位移又平均了更复杂历史。若速度确实发生变

化，不同窗口的结果不必完全一致。把这种差异称为模型失败之前，应先检查比较对象是否同一时段；但也不能用时间变化无限免责，变化本身需有可检验记录。

球面运动学建立的是全球运动的相容性：边界方向与速率可以接受共同旋转模型的检查，局部结果又可据此推向未测区段。它对动力学提出了更明确的目标，却未自动确定物质强度、热对流或板片作用各占多少。多个动力模型若能产生相近表面速度，单凭运动学还难以区分；深部结构与物性资料因而仍有独立任务。

一个三板块教学例子可以进一步显示闭合的含义。设 A 相对 B 的角速度已经由某段海岭估计，B 相对 C 的关系由另一组边界资料限定，那么 A 相对 C 的角速度可由向量关系得到。这个推算不要求 A 和 C 拥有一段方便测量的共同海岭；它依赖的是三者属于同一参考和时间安排。若后来在 A—C 边界获得可靠运动资料，就能检验整个回路，而非只检验一个局部参数。

闭合失败后，不能立即指定某一类数据错误。应比较误差大小和方向，检查是否存在系统偏差，确认不同数据对应的时间窗口，最后再考虑单元划分。若所有误差都被独立估计后仍出现显著剩余，新增微板块或分布式变形便成为候选。它们要改善的不只是一个局部拟合，还应减少整个回路的矛盾，并产生新的可查边界。

速度向量在球面上的比较还要求统一表达。不同地点的东、北方向并不是全球固定的一组平行轴，不能把局部东向数值直接当作同一笛卡尔分量相加。角速度向量提供了一种统一方式，然后再投影到各地点得到局部速度。这项数学转换看似技术性，却使全球不同地点的观测能够进入同一账本，是板块世界形成的重要操作条件。

旋转极附近也存在识别困难。离极较近的点线速度小，方向对误差可能敏感；仅用这些点估计旋转，约束可能较弱。分布在较大空间范围的资料更有机会区分不同极位置。由此，数据数量与几何覆盖再次分开：很多集中在小区域的观测，未必比少量分布良好的观测更能决定全球旋转。

古重建中的参考板块选择，不改变相对关系，却会改变图像中的绝对路径。把非洲固定与把另一板块固定，可以使大陆在动画中走出不同轨迹，但只要相对旋转一致，某些地质比较仍相同。研究者必须说明要回答的是相对位置问题，还是相对于地幔等参照的运动问题。参考系不是可以随意混用的绘图习惯，而是命题内容的一部分。

三联点把边界连接问题集中呈现。三种边界在一个区域相遇时，其速度关系必须允许该交会处以相容方式演化。不是任何海岭、海沟和转换段的任意拼接都能长期维持。即使不展开全部分类，读者也能看到网络理论的新增责任：不仅每条边界要合理，边界之间的交会也要合理。这是从孤立对象转向系统对象的关键一步。

◆ 30.3 独立检验如何形成共同体

板块体系的证据力量来自共同约束，而不是简单计票。磁异常限制扩张历史，转换断层限制相对方向，震源机制限制局部运动方式，钻探与年代资料检查海底历史。若这些资料共享某些校准或已被用于模型拟合，就要把依赖关系画清。三类不同名称的资料，可能仍通过同一时标或同一边界选择相连；相反，同一类资料中的新增区域也可能提供有价值的外部检验。

可以把一项研究分为输入、推断和留出检验。输入用于估计板块划分与旋转参数；推断据此计算其他边界的方向与速度；留出检验则用未参与估计的资料比较。这个程序不要求所有历史研究都按现代统计规范开展，而是帮助我们识别哪一种成功最具区分力。参数已经根据某条磁曲线调整，再报告曲线拟合良好，说明模型能容纳资料；预测另一区域的运动并得到支持，承担了更强责任。

震源机制特别展示了方法互补。它不需要从海底条带宽度直接推导滑动方向，因此能够对某些边界关系提供不同来源的约束。但机制解也有自身歧义，不能把图上的双力偶符号当成无需解释的直接照片。断层面与辅助面、台站覆盖及定位误差都需要与区域结构共同判断。独立方法的价值是改变误差结构，不是消灭误差。

钻探的实物记录同样如此。它让年龄模型可以接受生物地层、岩石与沉积资料检查，但不同年代标尺仍可能共享外部校准。可信度提高来自多种误差不会以同样方式同时制造一致结果，而不是来自“实物一定比模型可靠”的直觉。模型指导取样，样品反过来限制模型，是科学正常的往返；只要依赖关系公开，就不必把这种往返一概称为循环论证。

共同体的形成，也可以从共同任务理解。不同研究者不必采用完全相同的理论语言，只要他们能够围绕同一板块边界交换可比较的量：方向、速率、年龄、位置及误差。一个

地震学家的结果会影响海洋地质学家的重建，一个磁测模型的速度又会提出新的钻探问题。学科联系不再只是专题会议中的并列报告，而能够进入彼此的计算与采样。

共同标准同时带来约束和风险。统一坐标、符号与时标降低了合作成本，却可能传播共同错误。如果一套时标存在系统偏差，许多看似独立的速度估计会一起受影响。若教学中过度固定板块名称，也可能让研究者忽视分布式变形。成熟共同体不仅共享成功模型，也要保存版本、替代方案与异常资料，让共同错误有机会被识别。

教学普及是稳定化的一部分，却不是理论成立的独立证据。一张世界板块图进入教材，说明某套表达具有传播能力；它是否可靠，仍取决于资料与推断。课堂图常将边界画得整齐、箭头画得均匀，适合初步学习，却会隐藏不确定性与尺度差异。高等教育需要进一步恢复这些条件，让学生理解板块不是永不改变的世界分块，而是具有强大解释力的近似体系。

某些区域的困难不会因为全球体系成功而消失。宽阔大陆变形带、复杂微板块与变化边界可能要求更细模型。局部困难也不自动推翻全球运动学，因为近似理论本来有尺度边界。关键在于能否明确指出哪里需要细分、细分增加什么参数、哪些资料支持这些参数。若所有困难都靠任意增加小板块消除，模型会重新失去检验力。

共同检验还需要区分参数误差与模型误差。参数误差是在既定板块划分和运动形式下，对旋转极或速度估计的不确定性；模型误差则涉及刚性假设、边界位置或过程遗漏。增加同类数据可以缩小参数误差，却不一定修正模型错误。若所有资料都以同样方式偏离，继续提高小数位数可能只会更精确地描述一个不合适的模型。

残差的空间分布可以帮助区分两者。随机散布可能与测量误差相容，有组织的带状偏离则可能提示遗漏变形区。时间上持续变化的残差，可能要求速度分阶段处理。诊断不应依赖单一指标，而要结合资料类型与区域地质。板块体系的成熟不在于所有残差都很小，而在于能够把残差转成具体研究问题。

另一种风险是成功资料更容易发表和进入教材。与模型相合的磁剖面和地震机制具有清楚图像，复杂或难以解释的资料可能留在技术报告中。若思想史只追踪经典图版，就会高估早期证据的整齐程度。保留失败拟合、异常区域与选择标准，有助于恢复理论实际面对的困难，也使后来的研究可以区分真正突破与展示方式改变。

共同体因此不仅由同意某个结论的人组成，也由共享检查程序的人组成。他们可以对局部边界、参数或机制持不同意见，却知道用什么资料推进争论。争论的可协调性比意见完全一致更重要。一个稳定世界允许分歧在共同尺度上表达，而不是要求所有研究者复述同一图景。

板块网络的统一性并不排斥多样性。不同边界承担不同过程，同一类型边界也可以在速率、几何和材料条件上不同；统一来自这些差异能够在共同运动规则中相容，而不是把所有边界变成同一种结构。和谐性因此是一种受约束的配合，允许局部特点保持，同时要求它们对整体关系负责。

这一点也可以用于评价科学综合。若一种总理论只有把地方差异删除才能维持统一，它可能过度简化；若它允许每个地方任意特殊化，又失去共同规则。板块体系的力量，在于用较少的运动学关系联系多种现象，同时保留岩石、沉积、热与地震的专门过程。它的局限，也发生在这些近似无法继续兼容的地方。

对象、规则与能力三者的共同改变，解释了为什么板块理论被视为地球科学的重要转折。新的名称能够传播，是因为背后已有一套能够使不同工作彼此检查的结构；新的结构能够延续，是因为它允许异常转成研究任务。理论的生命不只在于某个发表年份，而在于此后不断发生的测量、修订与再组织。

深部地球研究将继续检验这一体系的动力条件。表面运动越被精确约束，内部模型需要解释的责任就越明确；内部物性与结构越清楚，表面运动的可能机制也越受限制。两条研究路径并非前者结束后才轮到后者，而是在不同证据条件下相互推进。

新世界的形成因此具有开放结局。板块理论提供稳定而强大的研究结构，却仍需要新的资料、局部模型与动力解释持续参与。它所创造的不是没有问题的地球，而是能够以更明确方式提出和解决问题的共同地球。

由此，区域差异不只被容纳在统一体系中，也成为检验其边界的位置。一个异常区段若始终无法纳入共同旋转，研究就要判断它记录了局部变形、时段变化，还是原先运动单元的划分有误。

◆ 30.4 板块世界新增了什么

回看本编开端，研究者面对的是若干轮廓相似、海底起伏和异常曲线。到板块体系形成时，这些不再只是孤立材料，而成为能够共同规定运动单元的关系网络。海岭是生成边界，海沟与某些会聚结构承担消耗关系，转换断层限制横向相对运动，磁性和年龄记录则为历史速度提供时间尺度。它们的身份相互关联，任何一项变化都可能要求别项重算。

同一区域的研究任务因而发生实质改变。原先可以问两侧大陆是否拼得上，后来能够问某一时段相对旋转怎样、某段边界是否应当伸展或会聚、某处海底应具有什么年龄，以及新增震源机制是否符合预期。新问题不是旧问题的简单精细化，因为可操作对象与排除条件已改变。大陆轮廓仍有价值，却不再独自承担全球运动。

在 SDE 框架中，显露 S 是具有共同运动规则的板块网络，差异序列 D 由磁测、地震、年代与几何闭合的交叉检查构成，E 则表现为边界类型与相对速度互相规定。一个边界被判为转换类型，会限制速度方向；速度闭合不容许该方向，又会要求检查边界身份或板块划分。特征纠缠不是图上连线的总称，而是这些不能任意独立选择的成立条件。

三式可把这套关系分为三个相连问题。S＝F(D,E)询问网络如何受联合资料限制；D＝G(S,E)询问已确定的旋转与边界怎样选择新测线或站位；E＝H(S,D)则询问新增结果是否要求改变参数共享、边界划分或误差处理。例如，一个回路的剩余若通过改正共同时间窗而消失，变化发生在比较条件；若只有新增运动单元才能消除，变化才触及板块对象。二者都可以推进知识，却不是同一类重构。具体历史仍需由修订前后的材料来区分。

所谓新的科学世界，既不是科学家创造了古老海洋，也不只是他们在头脑里换了一幅图。研究者形成了新的辨认与推断能力，测量操作有了新的组织方式，海底属性在共同检验中取得新的身份。这一主体、互动与对象共同形成的系统，能够继续产生知识和纠正错误。世界的科学发生因此具有现实操作后果，而不局限于表达层面。

板块体系的综合首先依赖地学中已有的运动学、磁学、地震学与年代学工作。SDE 的组织作用，是让这些工作的对象规定和相互回写更容易被逐项检查：一项边界身份怎样限制可用速度，一组新速度又怎样要求重查边界。若只是将成熟解释改写为三个字母，并未增加说明。教学上的能力变化则更容易辨认：掌握板块名称还不够，读者还需能由边界

类型推出相对运动，由运动提出年龄或地震预期，再判断异常究竟要求修订参数还是对象。这个推断过程，才是全球图景得以继续用于新资料的条件。

这种能力也包含知道何时停止使用简化模型。在局部工程或灾害研究中，全球板块速度不能代替具体断层的结构与活动史；在深部动力研究中，表面运动不能唯一反演内部力；在古地理研究中，现代边界不能直接投影到久远过去。承认这些边界，能够让板块理论与其他尺度知识相互配合，而不是成为解释一切的万能名词。

本编也把尚未解决的问题交还给下一编。板块能够如何运动，与地球内部是什么、怎样变形和传递能量，是相连但不同的责任。海底世界的成立并没有让深部地球变得透明。接下来，地震波、实验物性与反演将提供另一条发生路径：从地表记录出发，怎样让无法直接到达的内部结构获得可检验身份。

第 7 编

深部地球的间接显露

第 31 章 地震波与分层地球：莫霍、古登堡、莱曼的三个节点

第 32 章 高压岩石实验：实验室材料如何获得地幔意义

第 33 章 重力与磁场：多解性如何约束地下世界的生成

第 34 章 内部热与地幔动力：能量来源怎样成为运动解释

第 35 章 深部地球的公共事实：参考模型怎样稳定又怎样开放

第 31 章 地震波与分层地球

莫霍、古登堡、莱曼的三个节点

◆ 31.1 莫霍：同一次地震的不同到时

我们今天熟悉的地球内部剖面，并不是有人把地球切开以后画成的。它来自地表一批记录之间不肯彼此重合的差异。同一场地震从震源向外传播，在不同距离的仪器上留下不同的起跳、振幅和到达时间。研究者必须先回答一个看似琐碎的问题：远处记录上的某一段摆动，与近处记录上的哪一段摆动属于同一种传播过程？只有这个问题得到有条件的回答，时间差才可能成为深度差的证据。

莫霍洛维奇讨论的地震发生于 1909 年 10 月 8 日，研究论文发表于 1910 年；1992 年的英文译刊是后来重新开放这份论证的一个入口。这三个时间不能折叠成一个“发现年”。地震制造了可被记录的传播过程，台站留下了可比较材料，论文才把材料组织成关于地下结构的公开主张。[B06] 当我们说莫霍面在这一研究中被发现时，所指并非地下界面在这时才形成，而是一个具有可识别波相、可预测走时与可复查位置的科学对象获得了明确身份。

为什么一张到时表能约束地下结构？可以用一个不冒充历史原图的教学模型说明。设地表以下有一层厚度为 h 的较慢介质，速度为 v_1；其下是速度为 v_2 的较快介质，且 v_2 大于 v_1。若震源近似位于地表，沿上层传播的直达波走时近似为 x/v_1，x 是地表水平距离。另一路波先向下传播，在满足临界折射的条件下沿界面附近的高速层传播，再返回地表。它绕了远路，却在较长一段路上走得更快。距离较小时，绕路的代价较大；距离增大后，高速段节约的时间可以超过上下往返的额外时间。

在平面、均匀、各向同性的二层近似中，折射支的走时可写成 $t_r = x/v_2 + 2h\cos i_c/v_1$，其中 $\sin i_c = v_1/v_2$，i_c 从界面法线量起。该头波路径要求水平距离足以容纳上下行斜段，在此模型中即 $x \geqslant 2h\tan i_c$；把直线外推至零距离得到的截距不是实际零距离头波。这里重要的不是要求读者记住公式，而是认出两种可观察差异：一条走时支的斜率较大，却几乎没有截距；另一条的斜率较小，却有额外截距。斜率联系传播速度，截距

联系向下和向上的路径。两条支在某一距离相交之后，先到达的波便可能换了身份。因此，最早到达者不总是沿最近路线传播者，给第一到时标一个统一符号也不等于完成了波相解释。

这段教学推导揭示了一项经常被地球剖面图遮蔽的工作：界面不是从“地震波变快”这一句话直接推出的。研究者还必须检查，同一组速度与层厚能否同时说明不同距离的到时，能否容纳后来到达的相关波支，能否在更换地震后保持大体一致。如果只解释一个测站，震源位置的误差、台站时钟的偏差或者某次起跳的误读都可能扮演地下异常。跨距离的系统性结构，使个别误差解释逐渐变得昂贵。

真实地球的震源具有深度，速度可以连续变化，界面会倾斜，地表也并不平坦。二层平板只抽取了速度、路径与走时相互限制的关系，并非莫霍原计算的复现。推广到实际观测时，必须重新计算这些条件对走时曲线的影响。例如，倾斜界面会使相反方向测得的视速度不同，连续速度梯度也会使直线分支弯曲。由此，层厚与速度的估计要依靠整组记录，而不能只凭图上两条直线的交点。

一个反对者可以提出：较深处速度连续增大，是否也能造成远处先到的波？这不是故意抬杠，而是检验界面主张必需的竞争模型。若观测只能说明传播速度随深度增加，便只能支持这一较弱结论；若不同走时支、折射特征与其他波相共同要求较急的变化，界面的解释才得到更强支持。地震学中的“不连续面”还必须联系观测分辨率理解。对于波长很长的信号，一个较薄的过渡带可能表现为突变；对于更短波长的观测，这个突变又可能展开为内部结构。因此，科学上的界面身份既不任意，也不意味着自然必须服从一条无限薄的数学平面。

SDE 分析在这个节点上应当落到具体关系。S 是能够在记录与模型之间被追踪的壳幔分界；D 包含不同距离、不同波支和不同地震的到时差异；E 不是台站、人员、仪器的清单，而是速度规定允许路径、路径规定应有到时、到时再限制速度与层厚的特征纠缠。界面成立以后，研究者可以有目标地追问某种折射波应在哪些距离出现。这是对象身份对下一轮比较的回写。只有当这样的新检验实际发生，我们才把它与一般的事后解释区分开来。

还可以从反射波看同一界面的另一种约束。在上述平层教学模型中，若波在水平界面反射再回到地面，其走时随距离并不是一条简单直线，而近似满足 $t^2 = t_0^2 + x^2/v_1^2$，零偏

移走时 t_0 与上下往返的层厚有关。折射支主要通过斜率和截距限制结构，反射支则通过曲线形态提供不同联系。若两类资料指向相容的厚度，界面解释得到加强；若二者冲突，就要检查是否识别了不同界面、是否存在速度梯度或几何倾斜。这个推导揭示，所谓独立支持可以来自同一物理理论下不同传播路径，但二者共享速度假设，仍不是完全无依赖的证据。

震源时刻的不确定也可以通过差分处理部分削弱。同一台站上两个相位的到时差，不含共同的发震时刻；同一事件在不同台站的相对到时，也可以帮助检查时钟与路径问题。然而，做差只消除共同项，不会自动消除相位误读或不同路径的模型误差。一个数学处理究竟降低哪种不确定性，应从误差进入公式的位置判断。地震学的精细性，常常体现在这些不起眼的相减、归组和残差检查中。

◆ 31.2 古登堡：核幔边界的约束

壳幔边界的出现，没有把整个地球内部一次照亮。较浅的速度突变与深部地核边界，涉及不同距离、不同波路和不同的证据组合。古登堡 1914 年的工作是核幔边界约束中的重要节点；后来常用的边界深度约为地表以下两千九百公里，但现代参考数值不应不加区别地倒填为早期论文的全部结论。[B07] 更要紧的是，把地核存在、地核边界位置、外核物态与内部成分分开。它们可以互相支持，却不是同一个命题。

在一个球形、速度随半径变化的模型中，射线会弯曲。地震波从地表附近出发，向更深处进入，最后又返回地表。不同出射方向对应不同的最深到达位置和地表震中距。若速度结构连续而平缓，地表到时随震中距的变化也受到一组连续条件约束。加入一个速度变化明显的深部边界，某些射线会在边界上反射，另一些会折射进入内部，其返回地表的位置也会发生改变。于是，地表上的一段信号稀少区域，可以成为地下传播结构的约束。

“阴影区”容易令人联想到黑暗房间中的遮挡，但地震波的阴影不是绝对无声。不同波型、绕射、散射以及仪器灵敏度都会影响某一距离上能看见什么。严谨的说法是：在指定波相和观测条件下，某种直接传播的射线分支不再按原模型出现，或者其振幅显著降低。若不注明波型，把某类纵波的阴影区与横波无法直接穿越液态外核混为一谈，就会把不同物理机制合成一个含糊图像。

纵波以压缩和膨胀方式传播，其速度取决于体积弹性、剪切弹性与密度。横波要求介质承受相应的剪切响应。这里所说液态外核不能传播通常意义的体横波，是关于地震频率尺度上剪切响应的物理判断；它不表示所有以横波起始的能量都不可能抵达远方。波可以在界面转换类型，以纵波形式经过某段路径，再转换回另一种类型。一个字母标记所代表的，不只是仪器上一条线，而是一条受到传播物理限制的路径历史。

边界深度也不是由阴影区的起点直接读取出来的。到达哪个震中距，取决于核幔边界两侧的速度、上覆地幔的速度梯度和震源深度。若改变这些参数，边界位置的估计可以随之改变。研究者因此需要许多距离上的走时，而非一条“这里没有波”的记录。边界反射与穿核波的到时若提供一致约束，才能压缩深度与速度之间的补偿空间。这与上一节的二层教学模型相通，但球形几何、弯曲射线及更广距离使问题具有新的结构。

从知识发生的角度看，缺失并不天然是证据。只有在仪器能够记录、台站位置合适、震源激发足够、且某模型明确预期信号出现时，缺失才可能反对该模型。仪器没有开机，不能证明地核存在；某一方向辐射较弱，也不能直接证明地下介质不能传播。阴影区成为有价值的对象，依赖于对“本来应看见什么”的受约束预期。它把没有读数的空白转化为可以比较的缺失模式。

空白沿震中距、波型和震源条件形成可重复结构以后，才成为差异序列 D 的一部分。速度结构、可行路径与波相识别的相互约束构成 E：速度结构限定哪些路径能够到达台站，路径又限制某个到时可以归入哪种波相。这组预期与实际缺失相比较，才逐步限定边界 S。若改变一个尚未校准的站点增益便让“阴影”消失，差异首先属于观测系统；若多台站、多事件仍保留相近的缺失模式，单站故障解释就难以维持。阴影区的证据力量来自这种排除过程。

核幔边界研究还留下一个重要未决：同一速度跃变究竟对应什么物质差别？仅靠走时，不能唯一回答温度、成分和相态。地震学在这里获得的是动力传播意义的分层，不是一次完成的化学剖面。这条限制不是其成就的减损，而是把后续矿物物理与高压实验引入深部研究的明确接口。

球形几何还赋予走时曲线一种可利用的联系。射线参数把局部入射角、半径与速度联系起来，在球对称近似下沿路径保持相应不变量；走时对震中距的变化也与该参数有关。研究者因此不只观察某点到时，还观察曲线的斜率与分支怎样变化。曲线弯曲不是绘图风

格，而是在把不同深度的路径组织到同一结构中。若台站只覆盖很窄距离，就可能难以区分一条真正的新分支和旧曲线的局部弯曲。

此外，振幅证据与到时证据承担不同责任。到时主要限制传播积分，振幅还受几何扩散、聚焦、衰减、散射与震源方向性影响。把振幅加入分析可以增加辨别力，却同时引入新的条件。最有说服力的组合不是要求每种资料都独立给出同一个深度数字，而是让它们在一组共同参数下解释各自能够约束的部分。深部对象的增厚，正是这种不同证据责任的结合。

◆ 31.3 莱曼：异常 P′与内核假说

一个对象一旦稳定，就会使此前难以辨认的偏离获得新的意义。已有地核模型能够解释大量走时，但也会对穿核波的到达位置、相对振幅与分支作出预期。某些记录若持续越出这些预期，问题便不再只是“地震波为什么有许多摆动”，而是“既有传播结构为什么允许不了这些摆动”。莱曼 1936 年的 P′论文，正处在这样的分歧中。[B08]

莱曼讨论了 1929 年新西兰地震的记录，注意到某些穿核波的振幅和组成不宜被视为无足轻重。公开的原论文节选及后来的学术整理保留了她的推理：在地核内部加入速度更高的区域，可以解释原有模型难以容纳的信号；但她同时指出资料不足，假说尚不能由当时材料完全证明。[R31a] 新对象在这里以候选结构进入研究，支持它的是已辨认的传播差异，尚待确定的则包括这一结构的具体边界和更多物理性质。

先看为什么同一异常会存在竞争解释。地震波是波，不是只有一条几何细线的粒子。射线理论近似下的阴影区中，仍可能有绕射能量；复杂震源可以产生几个相邻脉冲；仪器方向不同，又会让同一运动在记录上的显著程度不同。于是，阴影区中出现微弱信号，并不自动要求内核。若新假说只做到“这里也能画一条射线”，其优势还很有限。它还应当说明信号的到时、在不同距离上的趋势，以及为何某些分量明显而另一些分量较弱。

莱曼所面对的对象因此不是一条孤立的异常线，而是一组必须联合解释的特征。振幅大小会受传播聚焦、震源辐射和仪器响应影响；分支之间的时间间隔又受路径与速度控制。这些特征不完全独立，却也不能任意相互替代。一个模型若为了修补到时而把速度改得很高，同时造成与观测相反的振幅趋势，就只是把困难从一个栏目移到了另一个栏目。新内部区域的价值，在于它有可能以同一结构同时改善几个有关联的约束。

“内核”在这里先是一种传播结构假说。它表示，在已经区分出的地核内部，存在一个需要另作规定的区域，使穿核路径和相应波相发生变化。不能因为今天知道内核具有固态性质，就写成莱曼仅凭当时的 P′记录已经同时完成了内核固态、成分、温度以及增长历史的证明。一个对象可以先以某种差异获得身份，随后才逐步取得更多性质。科学对象的成长，不是把后来所有属性偷偷装进最早的名字。

还必须区分“突变界面”与“强烈而连续的速度增加”。在有限分辨率下，两者可能给出相近的走时。莱曼节选中保留了对这种可能性的讨论。[R31a] 这意味着，内核候选的证据与边界无限尖锐的证据并不等强。后来剖面图把内核边界画成一条整齐圆圈，是表达一种模型的便利；它不应反过来抹掉原论证中对过渡形式的开放。承认这种开放，反而能准确说明后续研究还需要解决什么。

我们可以设计一个教学性复审，不声称它是莱曼当年的工作顺序。首先建立不含内核的候选模型，计算若干距离处允许的穿核波到时；其次加入一个速度较高的内部区域，只允许有限参数变化；再把某一部分记录留出，比较两类模型对留出记录的预测。若新模型只是用更多参数追随已知数据，它未必获得支持；若它在未参与调参的距离上仍能预测分支出现和相对位置，证据就更强。最后还要改变震源假设，检查此前归给内核的脉冲是否可能来自前震或震源复杂性。

这个教学程序显示，所谓异常并不只是旧理论的失败标记。它能够成为新测量的选择器。若两个候选模型在某一震中距的预测几乎重合，再多收集同类记录也未必有用；若在另一距离上产生明确不同的分支，那里才是具有区分力的观察位置。SDE 中的对象回写，不是研究者心中对新概念更有信心，而是候选对象改变了应寻找的记录、应比较的分量和应保留的误差说明。这样的改变可以在台站计划、相位目录和论文检验设计中留下证据。

莱曼对资料不足的说明，使内核假说保留了清楚的后续任务。新记录既可以加强内部高速区的解释，也可以迫使人们修改它的尺度与过渡形式。候选结构的提出与全部性质的确定之间因此存在一个研究区间：在这个区间中，模型已经能够组织比较，却尚不能排除所有竞争解释。把早期论文写成毫无保留的正确宣言，会遮住这一从异常到新检验的过程。

对于弱信号，还可以比较“相同震中距”与“相同方位”的不同作用。距离相近的台站有助于检查某条波相分支是否一致，方位不同则可能暴露震源辐射或地球横向结构的影响。若信号只在一个方位明显，不能立刻归给全球同心内层；若在多方位保持相关走时趋势，全球结构解释才更有基础。重复不是把同一种条件复制很多次，而是在保留关键联系的同时改变潜在干扰条件。

三分量记录也提供这样的机会。地表运动是向量，单一水平分量的小振幅可能只是运动投影较小，不能直接当作总能量小。联合垂直与水平分量，会改变哪些波被认为值得解释。这里不是仪器被动地多画一条线，而是新增可观察方向重排了异常的地位。一个曾被边缘化的细小起跳，可能因分量关系进入核心论证；反之，分量不相容也可能暴露误识别。

地震学中的命名稳定还使跨代复核成为可能，但需要同时保存原波形。一个相位在早期被归入某条分支，后来可能因更好的震源定位而改名；若只有整理后的到时表，后人难以判断改变来自新的识别还是重新计算。原始波形、仪器响应与旧解释一起保留，才能让对象的历史身份具有可追溯性。知识的传播不是只传播最终结论，也包括传播别人有权重新判断的材料。

◆ 31.4 层界怎样成为可复查对象

把莫霍、古登堡和莱曼并排放置，可以看出“地震波揭示内部”并不是一个一次完成的动作。壳幔界面、核幔边界和内核候选，分别依赖不同尺度的记录差异。它们共有传播物理，却不共有完全相同的观测条件；它们都受走时约束，却不能用同一张走时图替代彼此。统一性应当来自可比较的问题结构，而不是抹平每个节点的材料特征。

第一个共同问题是波相身份。某个起跳被标为 P，意味着研究者认为它属于指定类型的纵波路径；被标为一种穿核相，又附加了经过内部介质的路径规定。命名使交流成为可能，但名字本身不能证明身份。可靠的相位识别需要看它是否随距离连续变化，是否与其他分量和相邻台站相容，是否能够排除表面波、噪声或复杂震源的混入。若旧相位目录没有保留这些条件，后人不能只凭一个字母恢复全部证据。

第二个问题是共同误差。台站时钟偏差可能同时影响许多事件，震源定位误差可能同时影响一场地震的所有台站，速度模型偏差又可能影响不同事件的推算。看似有一百个数

据点，并不等于有一百个相互独立的限制。若所有残差朝同一方向移动，研究者应检查共享条件，而不是机械提高某个界面的位置精度。对象的稳定，不只是数据点越来越多，也是依赖结构越来越清楚。

第三个问题是分辨率。反演输出可以显示到小数点后很多位，但实际资料可能只区分几十公里尺度的变化。所谓界面深度，应附带在什么参数化、频率范围与空间覆盖下能够区分多少变化。这里的边界不是知识外面的附注，而是对象规定的一部分。没有分辨率条件，“地球在某处有一条界面”容易变成一项超过证据的几何承诺。

还有一个更难的问题：稳定究竟意味着什么？同一数据可能容纳不同的细节模型，但这些模型都要求某个显著速度转换区域。此时稳定的对象是这一较粗尺度结构，而不是每一条内部细纹。把模型集合的共同部分与彼此分歧的部分分开，比挑选一幅最顺眼的图更接近科学的实际进展。地下对象不必等到全部性质唯一确定以后才成立，但它也不能用少数稳定性质替所有细节担保。

三式可以把这些不同层次的结果联系起来。S＝F(D,E)在这里容许资料支持一组内部结构候选：它们共享某些可辨认层界，却在速度梯度和细节位置上有差别。D＝G(S,E)指向能够区分候选的事件、距离和台站组合。E＝H(S,D)则追踪波相识别、速度假设和误差关系怎样因新记录而调整。例如，若新增短周期信号揭示过渡带内部结构，旧的单界面表示便可能需要展开。相应改变发生在结构分辨与计算规定中，自然传播定律仍是共同约束。

分层地球由此获得了可接续的科学身份。研究者能够辨认新的波相，仪器记录进入新的归组方式，内部层界则被赋予深度、速度对比和可检验路径。三者共同产生进一步计算与复查的能力：后来的资料可以针对既有层界提问，也可以揭示它原先隐藏的细节。古老内部结构的物理历史与现代研究对象的形成历史，正通过这一持续的证据链相接。

现在我们已经能够说，在何处需要一个传播性质不同的区域。但为什么这些速度对应某种矿物组合，温度与成分怎样彼此补偿，地下压力又会改变哪些性质？只在射线图中继续加线，不能回答这些问题。地震学需要实验室提供新的约束，同时必须接受实验室关于迁移条件的追问。这正是下一章的起点。

第 32 章 高压岩石实验
实验室材料如何获得地幔意义

32.1 从岩样到受控物性

地震波把地球内部区分成传播性质不同的区域，却没有把这些区域的化学名称直接送到地表。速度较高可以与压力、成分和相态有关，速度较低也可能来自温度、部分熔融、流体或者织构差别。若要从一条速度曲线走向深部材料判断，研究者需要另一种互动：把有明确来源和组成的样品放进受控条件，考察其响应怎样改变。实验室不是地下世界的缩小照片，而是建立可检验对应关系的场所。

Birch 于 1952 年发表的《地球内部的弹性与构成》，为材料物性与内部模型之间的联系提供了重要入口。[B09] 这一问题要求把弹性、密度与压力放在同一个状态判断中：速度拟合若依赖某个压力条件，该条件就应与候选深度相容；密度若同时受整体地球约束，也不能再自由调节。具体结果还须区分直接测量、理论推算与测量范围外的外推，因为三者对地下材料判断提供的约束不同。

把一块岩石放到桌面上，我们可以测得一个波速，但这个数值没有脱离条件的权利。微裂隙在低压下可能张开，加载后会闭合；温度变化可能改变弹性响应；样品中的矿物颗粒有各自方向，合起来可能产生各向异性。样品经历过怎样的加热、卸载和反应，也会影响此刻读数。因此，更完整的物性陈述是：在指定组成、结构、温度、压力、频率与加载历史下，获得某种响应及其误差范围。

条件说明使重复实验有了明确的比较单位。另一实验室在不同温度下测得不同速度，首先提供的是温度响应信息；在组成、结构、温压与频率均可比时仍不能重复，才需要追查样品或测量环节。随着条件逐项对齐，哪些差异属于材料、哪些属于装置便更容易分开。物性知识的普遍性由这些可迁移的响应关系建立，而不是由一个脱离条件的数值建立。

在简化的均匀、各向同性弹性介质中，纵波速度满足 $v_p^2=(K+4\mu/3)/\rho$，横波速度满足 $v_s^2=\mu/\rho$，K 是体积模量，μ 是剪切模量，ρ 是密度。这两式是本章的教学关系，不是对某一历史样品的复算。它们立即说明，波速既不是密度的另一种叫法，也不是刚性的单独读数。相同速度可以由不同分子与分母组合得到；两个波型一起使用，才有机会分别限制剪切与压缩响应。

实验设备因而不仅负责把数字测得更精细，还决定什么差异能够被产生和分辨。若压力装置能升压，却不能可靠测温，温度与压力的效应仍可能纠缠不清；若只能测得纵波，便缺少对剪切响应的一条独立约束。新实验能力的意义，是把过去不能分开的变化拆开，让一个材料候选承担更严格的条件责任。深部材料作为科学对象，正是在这种压力、温度、结构与响应的共同规定中取得性质。

◆ 32.2 速度匹配为什么仍不够

假设地震模型给出某一深度范围的纵波速度，一组实验样品也能在某些条件下达到相近速度。可以立即宣布地下由这种材料组成吗？不能。匹配一项观测，只说明候选尚未被这一项淘汰。若另一种矿物组合通过不同温度也得到同样速度，两个候选仍然并存。实验与地震的联系因此不是“找到一个相同数字”，而是让尽可能多的独立性质在同一条件下同时相容。

为说明这种逻辑，可以构造一个不使用实测数值的教学情境。候选甲较致密，也较难压缩；候选乙较轻，但剪切响应不同。通过调整温度，二者的纵波速度都落入地震允许范围。此时再引入横波速度，可能发现甲必须取一个过高温度才能符合；若这个温度又与相态或热约束冲突，甲的可行范围便缩小。进一步加入密度约束，或许还能削减两者的补偿空间。这一过程不是投票，多一项测量也不自动胜过另一项，而是检查同一材料状态能否同时满足多个关系。

特别需要避免把几个共享假设的量当作完全独立证据。某个密度剖面若本来就由同一套地震资料及经验速度—密度关系推得，再拿它来“独立确认”速度解释，就有循环危险。研究者应追问密度约束来自质量与转动惯量、自由振荡、实验，还是先前的速度转换。来源不同，所能增加的信息也不同。把数据种类列成三栏，不等于已经获得三条相互独立的证据链。

组成还涉及混合问题。地下不是每个位置都由一种纯矿物均匀填满。不同矿物比例、颗粒排列和连通方式会影响总体响应；同样的平均化学组成，也可能因结构不同而产生不同的有效弹性。若存在少量熔体，熔体如何分布尤其重要：零散包裹与沿晶界连通，可能对波速产生不同影响。以纯样品结果解释聚合岩石，必须说明混合模型怎样把各组分连接成总体性质。

这使“组成”成为一个层级化的对象。化学元素比例、矿物相比例与颗粒织构并不是同一层次。某套数据可能较好地约束平均弹性，却只能宽泛限制矿物比例；另一套实验能够识别相变，但不能确定地球中该相的实际体积分数。科学写作应让结论停留在它所达到的层级，不能从“与橄榄岩类组合相容”跃迁成“已经唯一知道每种矿物的含量”。清楚地保留候选范围，胜过伪装成完整的地下清单。

地震波速度还具有尺度与频率。实验中的超声测量与地球中的长周期传播，可能受到不同松弛过程的影响；在某一频率下近似弹性的材料，在另一频率下会出现明显耗散。若不处理这种差别，实验与地震数值的偏离可能被误认成组成改变。反过来，借助任意的频率修正强行使二者吻合，也会让模型失去可反对性。修正必须有独立物理依据和可审查范围。

因此，深部材料候选的稳定来自交叉限制，而不是“实验更直接”这一抽象权威。实验直接接触的是实验样品，地震直接记录的是地表运动；两者都需要经过模型才能指向深部。它们各有强处：实验能主动改变条件，地震能采样巨大自然尺度。它们也各有限制：实验迁移有外推风险，地震解释有多解风险。把两种限制相接，恰好能够让一种方法暴露另一种方法藏起来的自由度。

多指标联合还可以通过一个局部敏感性表来理解。把纵波、横波和密度的变化分别写成温度、组成比例和相态参数变化的近似线性组合，系数表示各变量的响应程度。若两个未知因素在所有观测上的响应方向几乎相同，即使每个观测都很精确，也难以把二者分开。若新增一种测量的响应方向不同，参数才可能变得更可识别。这说明选择测什么，比单纯缩小已有误差条有时更关键。

这种教学分析也解释了为什么独立温度约束有价值。温度若可以在很宽范围内自由调节，许多组成候选都可能获得速度匹配；温度范围被其他证据限制后，某些候选便不能再

借温度补偿。反过来，若矿物组合已有较强依据，速度差异可以更多地用于限制热状态。资料的解释用途不是固定的，它随其他条件被约束的程度而改变。

◆ 32.3 实验外推的可撤回条件

实验室材料获得地幔意义，需要跨过几种不同的距离。第一种是温压距离。装置测到某一范围，不等于更高压力下必然沿同一曲线延伸。若跨越相变，原先的状态方程形式或参数可能不再适用；即使没有明显相变，参数之间的协方差也会让远距离外推迅速失去精度。论文中一条平滑曲线的延伸段，应与实测点所在区间明确区分。

第二种是时间距离。实验可能持续分钟或小时，地下过程则可能跨越漫长年代。平衡相是否来得及形成，取决于反应动力学；实验中为了加速过程采取的高温，又可能改变本来要研究的状态。另一方面，地球中的材料也并非时时达到平衡，快速俯冲或温度变化可能保留亚稳状态。不能简单以“地下时间很长”消除动力学问题，也不能把短时实验里未出现的相直接宣布为自然界不可能出现。

第三种是应力距离。理想静水压力与装置内的非静水应力有差别，后者可能改变晶格、裂隙和测得的方向响应。地幔中也存在偏应力与变形织构。研究者需要判断，实验偏应力究竟是需要消除的装置误差，还是可用于研究自然变形的一项变量。相同现象在不同问题中可能有不同身份，关键在于它是否被明示、控制和纳入解释。

第四种是样品尺度。小样品有利于达到高压，却可能不足以代表多相岩石的大尺度结构；大样品更接近某些非均匀特征，却更难保持温压均一。这里不存在一个永远最好的尺度。研究问题决定需要保存哪些差异，又允许平均掉哪些差异。若要研究晶体弹性，单晶有其优势；若要解释岩石总体的各向异性，晶体方向的分布就不能被删除。

可撤回的迁移主张应当写成条件链。例如：在某一温压范围、指定组成及混合模型下，该候选同时符合纵横波速度和密度的允许范围；若新的实验发现其相态不稳定，或独立约束要求温度超出允许区间，则该候选失去当前支持。这种写法没有假装已经掌握地下全部状态，却让失败在哪里发生一目了然。相反，“深部就是这种岩石”的无条件句，很难让读者知道哪一种反证会促使作者修订。

处理不一致时也应有顺序。先检查测量单位、温标、压力标定与样品描述，因为这些环节可以制造伪冲突；再检查频率转换、各向异性和混合关系；最后才讨论是否需要改变材料候选。这个顺序不是为保护某种组成而设置无限缓冲。每一次修补都应增加可验证的内容，且不能凭空扩大误差直到任何材料都合格。若所需修正远超独立证据允许范围，候选就应退出。

在发生学上，组成—相态—弹性—温压构成 E 中的具体联系：压力改变稳定相，稳定相与组成共同限制弹性，温度又使这些关系移动。加载、加热和取向改变形成可比较的 D，S 则是在这些序列中被限定的材料状态。若异常来自样品发生反应，应修订材料身份；若样品状态保持而换用频率后速度改变，则要检查由实验到地震的映射。由此可以区分迁移链上不同位置的失配，避免每次不一致都被解释成地下组成变化。

压力标定本身也是迁移链的一环。装置施加的外部负载不必等于样品内部的均匀压力，研究者往往需要压力标尺或其他校准关系。若标尺版本改变，一批实验的压力位置可能共同移动。地震速度突变与实验相变压力的对应也会随之改变。因此，物性数据库除了保存结果，还应保存使用的标尺、温度估计与不确定性来源；否则，后来所谓实验间冲突可能只是校准体系不一致。

同样，实验中观察到一种相，并不自动知道它在地下形成的速度。若反应需要成核和扩散，样品在目标条件停留多久、是否有缺陷和流体，都可能影响转变。用升温实验加速反应时，应判断改变的是到达平衡的速度，还是最终平衡状态本身。这些细节决定实验现象能否成为地幔事件的代理，而不是只在装置里可重复。

这种可追溯性还应延伸到样品本身。化学分析、含水情况、孔隙、颗粒尺度与制备过程若缺失，后来所谓同一种岩石的重复测量就可能没有相同对象。样品名称提供分类起点，完整描述才使物性比较有实际依据。

◆ 32.4 深部对象回写实验目标

实验与地球模型之间并非只有一个方向：先测一堆物性，再从中寻找与地下相近的数值。深部模型一旦提出某个速度跃变、异常区域或方向差异，就会改变实验应当制造和测量的条件。若地震问题关注各向异性，只测随机方向平均速度就不够；若问题关注某一相

变可能造成的波速突变，只提高常温测量精度也可能答非所问。对象规定因此能够回写实验设计。

查阅已有物性表，可以检验一个深部候选是否落在已知响应范围内；若候选之间的分歧超出表中变量，实验设计就需要改变。比如，数据库只有平均速度而地震异常表现出方向差异，研究者需要测取向依赖，而非继续扩大同类样品数量。回写的程度可由新增变量、温压路径和被排除候选来辨认。查表与改变实验路线因此承担不同工作，二者通过明确的材料问题衔接。

可以继续前面的教学情境。若两种材料在纵横波速度上仍难区分，但对温度变化的敏感程度不同，那么新的实验目标不是再测一次同温度速度，而是在几个受控温度下测响应斜率。若两种候选对密度相似，却在剪切衰减上不同，则耗散测量可能更有区分力。实验的方向来自候选之间的分歧，而不是来自装置恰好能测什么。这样，地球模型为实验提出问题，实验又迫使模型缩小或重排自己的材料集合。

这也提示我们，深部材料的公共知识需要保存失败实验。未能复现某一异常、某种组分在目标条件下发生了不可忽略的反应、某条外推曲线与新的测量冲突，都是对象边界的有效组成。如果只保留成功匹配，读者会误以为实验室到地幔之间存在没有歧义的直通道。失败记录让后人看见候选曾经如何变化，避免一再把被淘汰的自由度重新包装成新解释。

主体在这种研究中也会改变。地震学家逐渐学会区分传播异常与材料异常，实验者则逐渐把地下尺度的限制转成装置中的可检验条件。新的深部世界不是一方把结论交给另一方，而是两类能力在具体问题中相接：什么能够被测量，什么能够被外推，什么仍然不可识别，都在往返中重新规定。这一复合体生成新的对象与行动能力，但仍由独立实验和自然记录决定其成败。

高压研究因此没有消除反演问题，而是使反演获得更多结构。即使材料候选变少，地下几何、温度分布和混合比例仍可能互相补偿。下一章转向重力与磁场，把这种多解性直接放到台前：当几个地下世界都能画出相似的地表信号，研究怎样继续，而不以一张漂亮图像冒充最终答案?

第 33 章 重力与磁场
多解性如何约束地下世界的生成

◆ 33.1 先做两个同样好看的模型

一幅地下剖面常有一种超过证据的说服力：地表之下填上不同颜色，边界连成曲线，未知似乎就被转换成可见物。但在重力和磁场解释中，地表信号与地下结构之间往往不存在一一对应。不同深度、形状、密度或磁化分布，可以产生相同或极为相近的观测。Saltus 与 Blakely 在 2011 年的论文中强调，这种数学上的非唯一性，不等于所有地质解释在实践上都同样好。[B11] 问题不是是否承认多解，而是怎样使保留下来的解承担足够具体的地质与观测责任。

先做一个教学模型。把一个有限质量的地下异常体近似成点质量，其正上方的重力扰动与异常质量成正比，与埋深的平方成反比。若把埋深加倍，再把异常质量增为四倍，在这一个地表位置就能保持同一扰动。于是，一个较浅而质量较小的异常体，与一个较深而质量较大的异常体，可以给出相同读数。这里的“质量”是相对于周围介质的异常质量，不是地下物体的全部质量。这个例子不使用实际测区数据，也不声称点质量足以描述真实地层。

如果沿地表测量一条完整剖面，情况会改善。较深异常体产生的信号通常更宽缓，较浅者变化更集中，所以峰值相同并不意味着整条曲线相同。可是一旦允许改变异常体的横向形状、厚度与密度，补偿空间又会扩大。测线之外也可能有贡献。一个二维剖面模型若把沿走向无限延伸当作默认条件，得到的形状就可能与实际三维物体不同。增加数据能够削减某些自由度，却不自动废除所有非唯一性。

磁场解释还增加了方向问题。磁异常不仅取决于物体在哪里、磁性多强，也取决于磁化方向与观测场的几何关系。感应磁化和剩余磁化可能并存；若把总磁化方向强制设为现今地磁场方向，某些异常的位置和形状就可能被误解。地表一个正峰不必正对磁性体中心，负值也不等于地下不存在磁性材料。重力场中的直觉不能不加修改地搬到磁场图上。

这类多解还应分成两层。第一层是数学结构本身留下的非唯一：即使没有噪声，也可能存在不同的源分布产生相同外部势场。第二层来自实际观测有限：测点稀少、噪声存在、测区不完整，使本来可以区分的候选在当前资料下难以区分。前者需要额外类型的约束，后者有时可以通过更好的观测改善。把它们都叫作“数据不够”，会误导测量计划；把它们都叫作“永远不可知”，又会放弃本可取得的进展。

让两个拟合相近的模型并排存在，能够暴露一张剖面隐藏的选择。比较应使用同一数据和误差尺度，列明两个模型各自允许变化的深度、形状与物性。若较小残差来自更多自由参数，就要检查改善集中在哪些观测、能否延续到未参与拟合的测点；若简单模型已有独立地质支持，则应说明这项支持排除了什么变化。竞争由此落实到参数与预测，不再停留于图像是否顺眼。

重力异常还建立在一系列参考扣除之上。纬度效应、高程与地形等影响若未处理，局部读数就混合了目标以外的变化；处理以后得到的异常，也要保留所用参考和密度假设。尤其在山地，用某一密度作地形校正，可能影响近地表异常形态。若随后又从该形态推断相同密度，必须检查是否形成自我确认。校正使目标显露，也可能把先验带进目标。

磁场资料中的时间变化同样需要区分。测线不同时刻的读数可能含外部场变化，若缺少适当校正，时间差异会被画成空间条带。这里与天气同时性问题有结构相似，却不能直接套用天气标准：磁场变化来源、基准站作用与目标波长都有自己的物理条件。多解性不仅发生在地下模型阶段，也可能已经进入异常资料的生成。

◆ 33.2 联合资料怎样缩小而非消灭多解

假设重力资料允许两种解释：盆地中的低密度沉积层较厚，或者沉积层较薄但其密度更低。若有钻孔抵达基底，局部厚度便获得直接约束；如果有地震反射资料显示基底形态，几何自由度还会进一步减少。剩余重力差异可以更多地用于限制密度。联合资料的力量在于，不同互动对同一候选的不同性质施加限制，使原有补偿不再任意。

然而，钻孔只约束它实际穿过的位置。两个孔之间的基底并不因此自动成为直线，孔外深部结构也没有被直接看见。地震界面同样需要解释：某个反射可能与岩性、孔隙或流体变化有关，未必就是所假定的地质接触。联合约束应说明每种资料究竟限制哪一个参数、覆盖多大空间、具有何种误差。把几种数据叠在同一张图上，还不是联合推断。

更隐蔽的危险是共享经验关系。若把地震速度通过经验公式换成密度，再把该密度用来拟合重力，最后宣布“地震与重力一致”，一致中有一部分可能是事先强制写进去的。经验关系并非不能用，但必须说明它适用于什么岩性与温压范围，其散布有多大，以及剩余不一致如何处理。如果模型只允许一条没有误差的速度—密度曲线，它可能把真实的岩性差别转嫁成虚构的几何变化。

联合反演有时以“共同边界”为约束，要求重力与磁场模型的变化发生在相似位置。这在同一地质接触同时改变密度和磁性时有合理依据；若一个接触只显著改变密度，而另一种化学变化只改变磁性，强制共同边界就会制造不存在的统一。相互约束不是要求不同物理性质永远同步，而是根据独立知识判断何时同步、何时应当分开。允许资料冲突，是判断联结是否合理的必要条件。

可以把候选集合想成几个条件的交集，但这只是逻辑图景，并非保证交集一定狭小。重力允许一批模型，地震允许另一批，钻孔再允许一批；如果它们仍留下很大的公共区域，结论就应保留宽范围。如果交集为空，也不能立刻宣布自然矛盾了，可能是误差估计过小、物性转换不当或者某类数据存在系统偏差。真正的联合研究既关心交集变小，也关心交集为何消失。

在思想史上，地下对象由这种交叉限制获得更丰富的身份。一个“重力低值区”可以逐渐变成具有深度范围、几何特征与材料候选的盆地结构，但每一步都需要新的约束承担。这里发生的不只是给旧对象多贴几个标签，而是改变它能够参与的判断：可以估计容积，可以比较演化方案，可以选择钻探位置。若这些新能力只依赖未经证实的共同先验，所谓对象增厚就可能只是图像增厚。

这里的 E 由不同性质的联系组成。埋深与密度、磁化强度与体积可以相互补偿；地震界面约束几何，钻孔约束局部层序，从而限制补偿的范围。其中一部分联系由新增资料建立，另一部分来自求解时采用的共同边界或物性关系。两者应分开记录：撤去一个先验后仍保持的结构，与撤去以后便消失的结构，承担的地质结论不同。这种区分也为下一轮观测指出目标。

33.3 正则化也是对象规定的一部分

即使明确承认多解，计算程序仍然常需输出一组具体参数。正则化就在这里进入。它在拟合资料之外，要求模型具有某种可接受形式，例如变化较平滑、异常较集中，或者尽量接近已有参考。一个常用的教学表达是：最小化“资料失配＋λ乘以模型惩罚”。λ调节两种要求的相对分量。这个式子没有告诉我们哪一种地下结构是真的，而是公开计算如何在许多候选中作选择。

平滑约束偏好缓慢变化，可以抑制噪声导致的细碎斑点，却可能把真实尖锐接触抹宽。稀疏约束偏好较少的显著异常，可以突出局部体，却可能把连续变化压缩成几个块。几何先验若规定某个界面只能单调变化，便排除了复杂褶皱。每一种选择都有可能合乎问题，也都有可能错过自然。正则化的责任不是保证无偏，而是让偏好可见、可替换、可检验。

有人据此认为，地下图像都是算法想象。这一结论过强。先验可以来自露头、钻孔、区域构造和实验，而非个人审美；资料也可能强烈要求某个大尺度结构，不论使用何种合理正则化都无法消去。应当比较的是，在允许的先验范围内，哪些特征持续存在，哪些随算法偏好而改变。如果一个异常只在极端参数下出现，证据等级应低于在多种合理设置下都稳定的结构。

分辨率分析因此具有认识论意义。研究者可以构造已知结构，模拟观测再反演，检查程序能否恢复目标尺度；也可以分析资料对参数的敏感方向，判断哪些组合受约束、哪些组合接近不可见。前者的恢复成功依赖测试结构与实际对象的相似程度，后者通常受线性化或局部近似限制。二者都不能变成一枚“已通过分辨率检验”的无条件印章，但能够防止把计算网格大小误称为真实分辨率。

不确定性也不应只画成每个格点的独立误差条。两个区域的密度可能强烈负相关：一处增加，另一处减少，就能保持几乎不变的外部场。若分别报告固定其他参数所得的很小条件误差，或者只给出各格点误差而隐去协方差结构，读者可能误以为每个位置都被独立确定。科学对象有时稳定的是某个积分量或参数组合，而不是局部值。把这种关联写清，才能知道模型实际允许怎样变化。

这里可以重新理解“好看的模型”。平滑、连贯、色彩清晰有助于阅读，却不是地质真实性的独立证据。视觉上不整齐的候选可能暴露资料冲突，反而更有研究价值；过度整

齐的图像可能只是惩罚项很强。出版与传播必须承担这一责任：图例应区分数据覆盖范围、反演区域与外推区域，正文应解释哪些边界由观测要求，哪些由参数化规定。图像的美不能偷偷替代证据的强。

反演对象连同观测路径、数学约束和辨认尺度一起进入公共讨论。新的钻孔可以揭示旧模型在局部的错误，新的测线也可以显示原先平滑约束掩盖了某种变化。修订时须确定被反对的是哪一层：若大尺度质量亏损依然存在而几何形态改变，保留下来的是较粗的对象规定；若异常可由校正误差解释，则连这一规定也需要撤回。SDE 所追踪的对象生成，包含这些不同幅度的修订。

参数化的尺度还影响“简单”的含义。用一百个小网格表示一个圆滑异常，可能需要许多非零参数；用一个几何体表示，则只需少数参数。于是，稀疏性并非脱离表示方式的自然美德。模型比较应判断参数化是否为目标结构提供公平表达，而不能因为某种表示让喜欢的解释显得简单，就宣布它获得普遍优势。

一个实用的报告可以同时给出两种结构：资料支持的共同部分，以及随先验变化的分歧部分。前者可以进入较强地质结论，后者作为后续测量的目标。这种表达比把不确定性塞进一句“结果仅供参考”更有内容，因为它告诉读者究竟什么可靠、什么还在竞争。地下世界由此成为分层可信的对象，而不是全真或全假的整幅画。

异常的符号同样需要参照。密度亏损是相对于周围或参考模型而言，不意味着材料具有负密度；磁异常负值也不意味着磁性绝对值为负。把相对量当成绝对属性，会把反演问题在开始时就规定错。每一次解释都应保留“相对于什么”的关系，使图上颜色能够回到可计算的物理量。

因此，图像中的一个像素表达的是整体推断在该位置上的结果。读取局部数值时，还需要知道哪些邻近参数与它共同变化，哪些积分量保持稳定；否则，区域模型就会被误读成许多相互独立的地下测量。

◆ 33.4 把下一次测量对准分歧

有两个模型同时符合现有资料时，下一步不必是继续争论哪一幅更符合直觉。可以先问：它们对尚未测量的地方会作出什么不同预测？如果甲认为异常源较浅，乙认为较深，两者在近地表高空间频率信号上可能不同；如果它们在现有测线下方很相似，却在侧向几

何上分歧，增加一条不同方向的测线可能比原线加密更有价值。候选之间的分歧把新观测的位置变成一个可计算问题。

但差别最大的位置未必最值得测。那里可能无法到达，噪声特别大，或者仪器对目标信号不够敏感。实际设计应比较预期模型差异与测量不确定性，还要考虑成本和风险。一个模型预测差十个单位、仪器噪声却达二十个单位的位置，不如差五个单位而噪声只有一个单位的位置有区分力。这样的教学比例不是实地方案，但揭示了“多测一些”与“更有信息地测量”的区别。

还可以选择不同物理量。两种密度分布的重力预测几乎相同，却可能对应不同的地震速度或地电性质；在具备独立物性依据时，换一种互动可能比提高同一仪器精度更有效。反之，若所谓新资料只是同一原始数据的另一种滤波产品，信息增益可能很小。研究设计需要追溯数据谱系，防止把重复加工误当作独立复核。

测量回来以后，也不应只问它支持最喜欢的模型吗。若两者都失败，说明候选集合本身需要扩展；若两者仍不可区分，则应报告这一结果，重新评估可识别尺度。失败的区分实验并非无用，它限制了此前关于差异大小或噪声水平的设想。只有在每次没有区分力时都事后改变标准，模型竞争才会退化成不可结束的保护行动。

地下世界的稳定因此可以是有边界的。我们可能可靠知道这里存在显著质量亏损，却不能唯一确定它由哪一种岩性和几何组成；可能可靠定位一个磁性接触带，却不能把它的每段起伏都解释成断层。清楚的有限结论能够支持后续行动，也能阻止过度投资或过度解释。承认未知不是撤销对象，而是把对象规定在它已经承担得起的范围内。

这一章的三元路径由候选 S 返回测量 D，再修订关联 E。它与前两章方向相通，却有自己的特点：对象未必收敛成一个，而可以稳定为一个经过约束的集合。科学进步不总表现为从多变一，有时表现为把原先混在一起的可能性分清，并知道需要什么证据才能继续缩小。下一章将把同样的问题带入地球内部的热与运动：即使几何结构已较清楚，什么力量驱动它，仍然需要另一组不可省略的联系。

第 34 章 内部热与地幔动力能量来源怎样成为运动解释

34.1 从放射性热到能量收支

知道地球内部有热，与解释板块为什么这样运动，是两个相隔很远的命题。前者涉及能量来源与温度状态，后者还需要说明热怎样转成浮力、材料怎样变形、应力怎样传给边界，以及阻力怎样消耗可用能量。若把这些环节压缩成“热推动板块”，就像说汽油推动汽车，却跳过发动机、传动和路面条件。地球内部没有一台可以直接拆开的机器，因而每一环都需要独立的观测与物理约束。

霍姆斯的放射性研究使热问题获得新的内容。放射性衰变不仅提供矿物时钟，也提供持续释放能量的机制。《放射性与地球运动》长篇发表于 1931 年的格拉斯哥地质学会会刊，学会材料将其前身追溯到 1928 年的讲演；1931 年《自然》上的同题短通信另有澄清目的。[R34a] 讲演、长篇论述与短通信属于不同文本节点。这里讨论的历史转折，是持续热源如何使地球运动获得新的能量解释，而非将这些文本合并成一个已经完成全部动力计算的模型。

能量收支首先要求分开“今天流出的热”与“今天产生的热”。地表热流既可能包含当下放射性产热的贡献，也可能来自地球内部储存热量的释放；不同圈层还有各自的交换。若把地表热流全部归给当前放射性产热，会低估冷却历史的作用；若把它全部归给早期储热，又会漏掉持续热源。把观测转换为动力解释之前，先要知道这些项怎样进入同一个守恒关系。

在最简化的整体教学模型中，可以把内部能量变化率写为热源输入减去边界输出。这个表达并不等于完整的地球热演化方程，却揭示一个基本逻辑：相同的地表热流，可以对应不同的内部热源与冷却速率组合。因此，仅测得一个总通量，还不足以唯一恢复热源分布。地壳富集某些产热元素，会改变近地表热流与深部产热之间的对应；空间平均又可能掩盖局部热异常。热预算与温度剖面互相约束，但不能简单互相代替。

热流测量也有自己的互动条件。测得温度梯度，再结合热导率，才能在适当假设下估计传导热流；地下水循环或海底热液活动会重新分配热量，使局部梯度不再只代表深部传导。样品热导率随温度、矿物与孔隙变化。一个地表点的热流，不是一支直插地幔的温度计。它必须放在边界与输运过程的联系中解释。

热由地球“尚未冷透”的迹象转成运动研究中的变量，意味着解释开始同时承担能量与材料约束。某种热源分布即使能提供所需总功率，若其温度剖面使假定矿物组合不再稳定，机制仍需修订。相反，材料状态允许的温度范围也会反向限制产热与冷却的组合。放射性提供新的能量来源之后，问题不再只是有没有足够热量，还包括热量在何处释放、以何种速率输送。

能量来源的空间分布还会改变对流方式。同样总量的热，若主要从底部输入，与主要在体积内部产生，可能形成不同温度剖面和边界层结构。因此，热预算中两个模型总瓦数相同，不代表其动力预测相同。把热源位置省略，只比较总能量是否足够，会错过能够区分机制的观测。地壳产热、地幔产热与核幔交换在模型中各有入口，不能作为可任意互换的一笔资金。

热流与时间尺度也必须相接。地表看到的是此刻输出，内部温度则保存较长历史。某次边界变化后，系统可能尚未达到稳态；用稳态模型解释瞬态，会把历史记忆误认成热源异常。相反，允许任意初始热状态又会增加自由度。热演化模型需要把地球形成与后续冷却的约束逐步引入，而不能只调今天的参数。

◆ 34.2 对流需要哪些物性条件

有温差并不意味着任何介质都会形成可持续的循环。热扩散可以在物质基本不作整体搬运的情况下传热；浮力推动的流动则要求温度或组成导致的密度差克服相应阻力。黏度越大，变形越慢；系统尺度、热扩散率、边界条件和内部加热方式，又会改变流动的出现条件。把地幔画成几只旋转箭头，只是提出一种组织方式，尚未给出它为什么能运行的证明。

一个常用的教学判别是瑞利数。对于简化的、由下加热的黏性层，可写成 $Ra = \rho g \alpha \Delta T\, L^3/(\eta \kappa)$，其中 ρ 是密度，g 是重力加速度，ΔT 是上下边界温差，α 是热膨胀系数，η 是动力黏度，κ 是热扩散率，L 是层厚。该式比较浮力驱动与黏性、扩散作用，具

体临界值依赖边界与几何，不能把某个教科书数值直接当作整个地幔的统一门槛。它的启发在于：温差必须与材料和尺度一起判断，单独说“地下很热”并不完成动力学论证。

尺度以三次方进入这个简化比值，提醒我们不要把小容器中的日常经验直接搬到地球。手中的岩石在几分钟里保持形状，不能证明它在百万年尺度下没有流变；地幔能够长期变形，也不表示它整体是一锅熔岩。固态矿物通过不同微观机制可以发生缓慢蠕变。这里的固态性质与长期黏性流动属于不同问题和时间尺度，两者能够同时成立。

温度与黏度的联系进一步使对流成为非线性问题。温度影响黏度，黏度影响流速，流动又搬运热量并改变温度分布。边界冷却可以形成较强的表层，内部较热区域则可能更易变形。不能先任意指定一幅温度图，再从另一处借来一幅流场图，把二者合称为模型；同一个耦合系统应当说明它们怎样相容。若想维持一条固定上升通道，也要解释周围流动与边界补偿如何配合。

组成差异与相变还会改变浮力。冷的板片通常具有相对于周围介质的密度优势，但沿途压力、相态与化学组成会影响这一关系。相变边界可能改变流动的连续性，黏度随深度变化也会影响物质停留或穿越。一个只有温度、没有组成与相态的教学模型可以帮助理解，但不能自动承担所有地幔结构的解释。简化的价值来自其明确边界，而不是把边界藏起来。

如何判断一个对流解释超过了示意图？至少要能够从材料参数、边界条件与能量输入推出某些可比较量：流速的大致尺度、热通量、温度分布或地表应力。如果自由调整每个位置的黏度就能追随任何图像，模型的区分力便很低。相反，受到实验与独立观测限制的参数，即使给出的预测不够完美，也更能暴露真实问题。

在 SDE 语言中，这一章的 E 具有明确内容：温度与密度共同规定浮力，温度又影响黏度，流动把热重新分布并改变后续浮力。D 包括不同空间和时间尺度上的热流、变形及速度差异。S 是满足这些关联的动力候选，而非“地下有循环”这样不受约束的图像。三方程在这里帮助检查联系有没有闭合，并不替代热力学、动量守恒或流变方程。

材料时间尺度可以用一个简化黏弹性比较说明。若用黏度 η 与剪切模量 μ 构成特征松弛时间 η/μ，那么远短于这一时间的作用可能更接近弹性响应，远长于它的过程则可能显出黏性变形。真实地幔流变比这个单一时间模型复杂，但它清楚地说明：同一介质对地震

波和地质运动可以呈现不同响应。把地震中的固态性质直接用于否定长期流动，是把两个时间问题错误合并。

这种区别还影响实验选择。短时间测得的弹性模量不能独自决定长期黏度，蠕变实验又需要面对应力范围与机制外推。若实验使用远高于地下的应力来缩短时间，必须检查变形机制是否保持。动力解释并非只需一张物性表，而需要知道不同时间尺度上哪些关系可以迁移。

◆ 34.3 板块运动并非一个马达

板块运动学可以在不了解全部驱动力的情况下取得成功。磁条带、地震机制与全球几何能够限制板块如何相对运动，但“如何运动”不能直接回答“为什么保持这种运动”。速度箭头不是力的箭头。若把两者混同，就会以运动符合某个方向作为力学机制已经成立的证据，而忽略多种力平衡可能产生近似相同的速度。

讨论驱动时，需要先说明系统边界。把板块作为研究单位，可以列出与周围地幔相互作用、俯冲板片的负浮力、重力势能差以及边界阻力。把整个地幔—岩石圈作为系统，某些“外加驱动力”则成为内部能量转换的一部分。所谓地幔对流与板片下沉并不必然是彼此排斥的两个发动机，板片本身可以是对流系统的重要组成。争论若不先对齐系统边界，就容易把同一过程在不同层次的表述误当作竞争理论。

板片拉力强调冷而较密的俯冲岩石圈下沉所产生的作用；洋脊相关的重力势能差则能够促使岩石圈从较高区域向外运动。把“洋脊推力”简单解释成岩浆像活塞一样把两块板挤开，会遗漏其常见的重力意义。地幔施加在板底的剪切作用可以驱动，也可以阻碍，取决于相对运动和流场。边界摩擦、弯曲和周围介质阻力，又会消耗或重新分配应力。

公开研究中，板片拉力与俯冲地震机制的联系提供了一条具体检验路径。关于俯冲岩石圈的研究曾结合地震资料与构造模型，讨论下沉板片中的张应力及周围阻力，而不是只根据地表箭头判断力的来源。[R34b] 本章不把某个研究情境中的贡献大小推广成所有板块、所有时期的固定排序。不同板块的边界长度、板龄、俯冲几何和耦合状态不同，力的相对作用也可能不同。

比较机制必须保持量纲和尺度一致。某个模型给出单位边界长度上的力，另一个给出区域总功率，不能直接比较数字大小。力对速度做功才涉及功率；瞬时力平衡与长期热预

算也不是同一检验。若一个解释能够提供足够总能量，却无法把应力传到需要运动的边界，它仍然不完整。若一个局部力学模型拟合某处运动，却需要不合理的全球能量输入，也不能仅凭局部拟合获胜。

可区分的预测因此比口号重要。改变板片连接状态，会不会显著改变板块运动？不同黏度结构会产生怎样的地表变形或动态地形？某种深部流场是否与观测到的各向异性相容？这些问题各有复杂依赖，不能把一项现象当作单一机制的专属指纹，但它们能推动模型提出明确的比较。若所有现象无论怎样都被解释为“对流作用”，对流便从可检验机制退化成总括名称。

这一段历史的开放性也值得保留。板块构造的成功不等于驱动问题从此结束；未解决的动力细节也不反过来取消运动学证据。不同层次的知识能够以不同成熟度并存。一个高质量的科学世界不是所有问题都同时闭合，而是已经成立的关系能够为未决问题提供严密入口，并允许后者独立接受修订。

观测地表速度还需要参考系。相对板块运动可以由边界关系约束，绝对运动则要选取某种参照；不同参照下，板底相对地幔流动的解释可能改变。讨论“地幔拖着板块走”时，若未说明速度是相对谁，就可能把参考系选择误当作动力结论。力学模型必须把运动学基准与深部流场假设一同公开。

此外，板块并非内部毫无变形的理想刚体。刚性近似在大尺度运动学上有用，边界带与大陆内部却可能发生分布变形。若把所有速度偏离都归给深部驱动力变化，可能忽略岩石圈自身的弱带和边界条件。动力与强度分布共同决定怎样运动，不能只问马达多大而不问材料怎样传递应力。

地表形变对动力也提供时间窗口。卫星测得的短期速度可能包含震后调整或弹性应变，而地质重建常代表长期平均；若二者不一致，不必立即引入新的全球驱动。先区分观测窗口与所建模型的时间尺度，才能判断差异是暂态、边界锁定还是长期运动改变。动力解释必须与它实际面对的时间对象相称。

地表速度对相对运动的约束，可以强于它对驱动力分配的约束。不同力平衡产生相近速度时，还需要应力、热流或流变资料；运动学上的精确并不会自动消除动力学上的多解。

34.4 动力解释如何接受失败

面对一个动力模型，可以从三组相互牵制的要求入手：它是否满足能量收支，是否符合材料响应，是否解释所针对尺度的运动。三者不应分别使用互不相容的参数。如果热预算要求一种温度分布，流变计算却借用另一种温度下的黏度，而速度拟合再引入没有依据的边界力，形式上虽有三项成功，实质上仍未形成同一个地球。

失败的形态不同，修订的层级也应不同。若误差随某个可独立校准的参数小幅变化，参数更新可能足够；若无论在允许范围内怎样调节，模型都不能同时匹配速度与热流，则可能需要改变耦合方式或机制。不能因为“参数还很多”而无限推迟机制检验，也不能因为一次局部失配就宣布整个热对流框架失败。修订应依据失配的空间结构、持续性与对假设的敏感程度。

历史研究尤其应避免以现代计算能力为早期模型补写精确性。霍姆斯提出热与运动联系的价值，并不要求他已经解决今天所有边界条件；后来模型增加的物性、计算和观测，也不应全部归入一位先驱的预见。把纲领、概念模型、定量实现与业务化计算分开，才看得见每个阶段真正新增了什么。科学创新既有提出问题的创造，也有把问题变成可失败计算的创造。

从热源到运动的联系可以分成几段检验。热源与边界交换限制热通量，热通量结合输运方式限制温度，温度与组成限制流变，流变再与边界力共同限制运动。SDE 将这些联系放入同一解释链，追问失配在哪一段出现、又怎样改变实验与观测选择。例如，速度拟合失败若来自黏度随温度变化的设定，继续增加热源自由度未必能解决问题；若热流与温度同时失配，则只修改边界力也不充分。各项成功只有在同一组条件下相容，才组成完整机制。

在研究实践中，开放模型仍需要共同参照。各组使用完全不同的速度、密度和层厚基线，可能让动力比较陷入无法定位的差别。共享参考地球因此成为必要工具。但参考一旦被广泛采用，又可能让某些假设悄悄变成不再被提问的事实。下一章将讨论这项张力：深部地球怎样成为公共对象，同时保留被新证据重新规定的能力。

第 35 章 深部地球的公共事实参考模型怎样稳定又怎样开放

35.1 PREM 为何是一种共同起点

不同研究者谈论同一个地球，并不意味着他们使用同一个计算对象。一组采用某种平均波速，另一组采用不同密度剖面；一篇论文把地壳简化为一层，另一篇保留几层。它们都可能在各自任务中合理，但结果相比较时，差异究竟来自新资料、新方法还是不同起点，便不容易判断。参考模型的价值，是把一部分起点公开并稳定下来，使分歧能够被定位。

Dziewonski 与 Anderson 于 1981 年提出 PREM，即初步参考地球模型。[B10] 它以平均结构及相关规定服务公共比较，将不同类型的地震资料与整体约束纳入共同模型问题，并处理各向异性与频率相关响应。[R35a] “参考”规定了它的用途：“初步”则保留了修订空间。一套平均结构可以成为许多研究的共同起点，使用者随后报告相对于它的偏离。因而，理解 PREM 既要看模型包含哪些性质，也要看哪些地方差异被留给了后续研究。

把三维地球压缩成径向模型，意味着同一深度上的许多横向差异被平均或吸收。大陆与海洋、冷板片与较热区域，并不会因此在自然界消失。参考模型只选择一种问题尺度：哪些径向变化足以描述广泛观测的共同部分，哪些偏离应留给区域研究。若读者把它当作每条射线实际经过的精确材料剖面，便超出了模型用途。

共同起点提供一种重要的节约。研究者可以用同一基线计算理论走时，再报告观测相对于基线的残差；也可以在新模型中指出哪些参数发生改变。这样，一项结果的意义不再只存在于作者自己的计算体系里。别人能够复现基线、替换某一环节并检查影响。公共性首先是一种可接续的操作能力，然后才表现为图表在教科书中被反复使用。

它还改变异常的身份。没有共同参照时，“比预期慢”可能只是一句局部判断；有明确参考后，慢多少、在哪些路径和频率上慢，便成为可比较对象。新对象并不是把自然差

异人为制造出来，而是把差异放进共同的计量与推断关系，让它能够跨研究组传播并接受检验。参考模型因此既是一项已有知识的压缩，也是一种生产新问题的装置。

参考模型还像一种共同的计算语言。自由振荡研究、走时研究和矿物物理研究可以把各自结果映射到相同的深度或半径坐标，再比较允许的变化。它减少了每次合作都重新定义变量的成本，也使不一致更容易集中到某个环节。但语言压缩会损失细节：一个区域的复杂层序被平均后，某些局部特征可能没有对应参数。因此，使用共同语言时仍需保留原资料，避免把不能表达的差异误认为不存在。

◆ 35.2 共同对象仍有模型依赖

参考模型越成功，越容易让使用者忘记它仍是一套模型。参数化决定允许哪些变化：若只允许随半径变化，就不能直接显露横向结构；若采用各向同性，就可能把方向差异吸收到平均速度；若允许各向异性，又需面对增加参数后怎样约束的问题。不存在完全不作选择的模型，关键在于选择是否与任务相称，并且能否公开改变。

各向异性尤其能说明这一点。同一材料对不同传播方向或振动方向可以有不同响应。若观测包含这种差异，却用一个标量速度概括，拟合可能出现系统残差。加入更多弹性参数可以改善表示，但参数增加不保证都被独立确定。资料对某些组合敏感、对另一些组合不敏感，因而模型应报告哪些特征由数据要求，哪些仍受简化约束支配。

频率也使“一个波速”不够完整。存在非弹性效应时，传播速度与衰减可以随频率联系变化。不同观测若采样不同频带，不加处理地放到同一表里，可能制造表面矛盾。参考模型需要说明其速度所对应的频率约定及相关修正，使用者也须保持一致。否则，新论文报告的速度异常可能部分来自频率不匹配，而不是地下真的换了一种材料。

密度与速度之间的关系同样不能被过度简化。地震资料对波传播敏感，整体质量与转动惯量则对密度分布提供另一类约束。二者相接可以限制模型，但密度某些细节仍可能依赖参数化或物性关系。一个共同模型把这些选择固定下来，不等于选择从此成为自然定律。研究者若把其中某条经验联系用于新区域，就应重新检查其适用性。

模型依赖并非反对公共知识的理由。地图采用投影，时钟采用时间标准，实验采用校准关系，它们都包含约定，却能够可靠服务客观判断。参考地球也一样：某些规定是为了

让共同计算可行，某些结构则被多类资料强烈要求。把二者全部称为发现，会掩盖约定；把二者全部称为人为建构，又会抹去观测的约束。思想史需要逐项区分，而不是先选一种哲学口号再把所有细节装进去。

从 SDE 角度，公共 S 并不是某个研究者头脑中的图像获得广泛赞同。它包括能够被重复调用的参数、计算规则、波相约定与比较程序。E 体现为这些规定之间的实际依赖，D 则包括多类观测的拟合和残差。只有当共享模型使不同主体能够产生可比较结果，并让不一致回到明确环节，公共对象才获得了超过文字共识的稳定性。

“平均地球”也不是把每条剖面作简单算术平均。不同资料对不同深度与性质具有不同敏感性，覆盖亦不均匀；拟合过程中参数之间会补偿。某个深度的模型值可能受到大量某类资料控制，而另一个深度更多依赖整体约束与正则化。理解平均的形成方式，才能判断它适合哪些任务。若一个区域恰好偏离全球平均，这不是参考模型出错的充分证据，而可能正是区域研究要解释的事实。

◆ 35.3 异常何以获得公共意义

设两篇论文都报告某一区域“速度低百分之二”。如果一篇相对于 PREM，另一篇相对于自己的区域平均，两个百分之二并不具有同一意义；若一篇是纵波、另一篇是横波，或频带与分辨率不同，差异更大。异常没有脱离参照的绝对身份。科学交流需要保留它从什么基线、经过什么反演和空间平均而来。

一种负责任的写法应同时交代模型版本、观测类型、频率范围、空间覆盖与分辨率。例如，某一异常只在长波长尺度上稳定，就不能在讨论中突然把它解释成一条狭窄通道；某个区域因射线覆盖稀疏而受强平滑影响，也不能与覆盖密集区域的细节等量齐观。这些限定会使描述显得不那么简洁，却防止比较在不同层次之间滑动。

共同基线的便利还伴随共同偏差风险。如果许多研究都采用同一错误的近地表修正，其残差中可能出现相似结构；这种重复不一定是独立确认。若模型训练、反演初始化与结果评价都依赖同一参考，循环依赖会进一步加强。研究者应当寻找能够越出共享链条的检验，例如改变参考、使用不同类型资料或检查未经同一修正影响的观测。

但更换参考后异常数值改变，也不自动证明旧结果虚假。若旧基线与新基线有整体偏移，异常幅度变化是正常的坐标变换结果。需要问的是，原先具有物理意义的结论是否保持：区域之间的相对差异是否还在，预测的独立波相是否仍成立，关于材料或动力的推断是否改变。如果只有颜色和百分比改变，而相关预测不变，那么变化主要在表示层；如果关键判断翻转，就必须重新审查依赖。

异常研究也会反向推动参考模型。持续出现的系统残差可能说明公共基线遗漏某种平均性质，而非每个区域恰好都异常。可是，不能只因新数据与模型不同就立刻更新基线。某一批资料可能有共同误差，某个频段可能需要特别处理。参考更新应当在代表性、误差与跨资料一致性之间权衡，防止把暂时偏差固化为新的公共起点。

这使公共事实具有双层稳定性。一层是制度与技术稳定：大家知道怎样调用、引用和复算；另一层是证据稳定：在允许改变的假设和新增资料下，关键结构仍然成立。前者能够促进后者，却不能替代后者。使用范围广只证明传播成功，不能单独证明每个参数真实。反过来，一个新模型尚未普及，也可能在某项明确任务上具有更强证据。

参考更新有一条可以追踪的证据链：系统残差先指向可能遗漏的性质，重新估计的参数再接受保留资料的检验，新版本随后改变区域异常的基线。链条中的任一环都可能中断。残差若来自共同校正误差，更新应落在资料处理；新增参数若只改善拟合资料，则尚不足以确定参考结构。历史上的具体更新应按这些环节查证，单凭版本更替日期无法恢复其因果过程。

公共异常还依赖符号的约定。速度扰动可以报告为绝对差，也可以报告为相对百分比；深度可以相对于地表，也可以采用半径；各向异性参数则可能有不同定义。数字相似却约定不同，会造成虚假一致；数字不同却经过同一变换，则可能实质相同。共享模型因此需要共享机器可读参数和明确单位，而不仅是一幅广为流传的彩色图。

一种有用的敏感性检查，是让同一研究分别采用两个合理参考，再比较其核心结论。如果某个大尺度异常在两种起点下都保持，局部数值有所变化，便可以把结构判断与幅度估计分开报告。若异常完全依赖某个参考，则应追查差异是资料覆盖不足，还是反演中的固定条件限制了可达解。这样的检查使参考模型保持工具地位，而不会变成隐形裁判。

参考模型的开放还依赖引用纪律。若软件包内部默认使用某一模型，而论文只写软件名称，读者可能不知道具体版本；软件更新后，同一命令又可能产生不同结果。报告模型、版本和关键设置，能够把公共事实从隐含默认变成可复查条件。这种细节看似技术性，却直接决定共同知识是否真的能够跨时间接续。

公共模型也需要明确退役条件：当新增资料与旧参数在允许误差内持续不相容，且替代方案在独立检验中改善，继续只因习惯而沿用旧值便不再合理。共享应促进修订，而不是保护默认设置。

◆ 35.4 稳定不是终审盖章

公共模型需要版本，也需要保留旧版本。后来的研究者若想复查一项历史结论，必须知道当时采用的参数与约定；仅提供最新模型，会让旧计算失去可重建性。保存旧版不是固守旧知识，而是保留知识怎样变化的证据。一个可以被替换却无法追溯的模型体系，会使误差归因和思想史写作都变得困难。

版本变化还应说明性质。有时是参数小幅调整，有时是新增各向异性、改变频率处理或重设资料权重。前者可能保留原有对象结构，后者则可能改变哪些差异能被辨认。不能把所有更新都称为科学革命，也不能把结构性改变说成纯粹数字维护。对象、操作和关联是否同时发生变化，是判断发生强度的具体尺度。

如果新资料只能由更复杂模型解释，还需防止复杂度成为万能避难所。新增参数应当承担可检验责任，最好能改善留出资料或不同观测类型的预测。一个模型若只是把每个残差单独吸收，可能拟合得很好，却没有形成更可靠的共同地球。公共参照的价值在于可迁移的约束，而不是把已知数据全部记住。

标准、参考与修订已经是科学史、计量学和模型研究中的成熟问题。SDE 在这里承担跨环节的组织工作：公共对象 S 如何规定残差比较 D，残差怎样暴露参数依赖 E 中的失配，依赖关系的调整又怎样形成新的共同基线。关键在于区分两类变化：在同一表示中校正数值，与改变表示所允许辨认的性质。前者通常延续原有比较，后者可能要求旧资料重新处理；版本记录因此也是理解对象变化的材料。

深部地球的公共事实最终具有一种有限而坚实的形式：我们能够共享起点、定位分歧、重复检验，也能够说明何时应当更换起点。它不要求所有人永远画同一幅图，而要求改变有依据，旧结论有谱系，新结论有责任。这样的稳定不是终审盖章，而是让后续研究不必每次从零开始，同时不被起点囚禁。

从地下到天空，问题的材料将发生巨大变化。深部结构常由许多事件累积约束，天气场则在资料传来时已经继续演化；前者追问远处不可直接到达的内部，后者需要把分散地方的读数压进同一个时刻。二者共同需要公共参照，却对时间提出不同要求。下一章将从测站、报文和时钟出发，讨论一场跨地域的“同一天气”怎样获得可比较的存在。

第 8 编

天气的动力学与可预报性

第 36 章 共同的天气怎样形成：测站、时刻与可比读数

第 37 章 天气图、气旋与锋面：从压力分布到生命史模型

第 38 章 从物理纲领到电子预报：1904、1922 与 1950

第 39 章 洛伦兹 1963：确定方程为什么不保证长期轨道预测

第 40 章 资料同化：观测与模型怎样共同生成分析场

第 36 章 共同的天气怎样形成

测站、时刻与可比读数

36.1 海事记录与共同格式

一位船长知道自己所在海面的风向，另一位船长知道远方的气压。两份知识都可以可靠，却未必能够拼成同一场天气。记录可能使用不同单位，观测时刻可能指地方时间，风的判断可能来自帆的反应，也可能来自仪器；船舶位置本身还有误差。共同天气不是把纸张堆在一起就自动出现的对象，它需要把地方互动的结果转成可以相互比较的关系。

1853 年的布鲁塞尔海事会议与 1854 年英国气象机构的建立，是这种协作制度化的重要节点。英国气象局的官方历史资料说明，国际合作以收集、共享海事资料和改进图表为目标，FitzRoy 主持早期工作。[B13] 这些事实可以固定机构与合作的时间线，却不足以证明所有船舶在机构成立那一天已经使用完全一致的格式。标准的提出、实际执行和记录质量的改善，是需要分别追踪的过程。

共同格式首先使遗漏变得可见。若表中明确要求时间、位置、风、气压与仪器信息，空白就能被识别为缺测，而不再与“当时没有值得记录的现象”混同。格式还规定单位和编码，让两个人对同一现象的记录不至于因写法不同而无法比较。但格式不能替代观察能力。一个填得整齐的表格，仍可能包含不可靠位置、误读刻度或主观判断不一致。

海上资料又有移动观测的特点。船舶不是固定站，它沿航线穿过不同天气；一次连续记录同时混合了时间变化与空间变化。若船向东行驶时温度下降，这可能是天气在当地转冷，也可能是船进入更冷区域，或者二者共同作用。解释必须把位置序列与时间序列相接。航海日志的价值因而不仅是积累许多数值，而是保存每个数值发生于何种移动过程。

共同格式也会选择什么进入公共知识。容易量化的气压和温度往往更便于汇总，云、能见度和海况则可能依赖分类训练。若为了统一而删去所有难以编码的地方描述，标准化可能失去关键线索；若完全保留自由叙述，又难以进行大范围比较。好的制度应当使基本字段可比，同时保存必要的原始描述与观测方法，让后来的分类可以重新处理材料。

地方读数能够传播，依赖仪器、感知和记录规范对自然过程的共同取样。公共天气场还要求这些地方结果具有可比较的联系。SDE 中的 E 在这里包括时刻与位置对读数含义的规定，以及仪器响应对可比尺度的限制。例如，船舶位置不清，即使气压读数准确，也难以判断它属于哪一处空间结构；响应过慢的仪器则可能把快速变化平均掉。共同表格使这些条件能够一并传递，条件得到核对后，地方记录才进入共同的 D。

风的记录尤其能显示格式与能力的区别。风向要明确是风来自何方还是吹向何方，风速则需要区分瞬时值与时间平均。船上观察还混合了船速与空气运动，只有知道船的航向和运动，才能把相对风与地理参照下的风联系起来。若记录表没有这些规定，两艘船报告相同风向，可能并没有描述同一运动；不同报告也可能只是参照不同。

共同格式在此不仅统一写法，而是统一读数所指的互动条件。它把“我这里感到的风”转成可与另一位置比较的空气运动估计。这个转换保留身体与仪器的贡献，又把可比较性从私人经验扩展到公共网络。其成败可以通过重复观测与交叉校准检验，不必依赖对机构权威的信任。

◆ 36.2 一张图中的同一时刻

天气图看起来是一幅空间图，内部却藏着严格的时间要求。如果北方站的读数比南方站早两小时，而这两小时里气压正在快速变化，图上画出的南北差异可能部分属于时间差异。所谓天气场，是在明确时间条件下，把不同位置的状态放到一起。没有同时性或合理的时间调整，空间结构就可能是报文到达顺序制造的幻影。

可以构造一个简单教学例子。设两站处于同一片缓慢移动的压力系统中，在各自位置上，气压都以每小时两单位的速度下降。甲站在上午八时记录，乙站在十时记录。如果不注明时刻，乙站比甲站低四单位的数值就可能被画成一个空间梯度，虽然这一差别在例子的设定中完全来自观测时间。数字纯为说明逻辑，不代表任何历史气象事件。

解决方法并非总是把所有记录机械平移到同一刻。需要知道变化速率及其不确定性；若气压突变，线性插值可能失效；若两站属于不同天气系统，用共同变化率又会制造新误差。因此，标准观测时刻很有价值，它减少对后期推算的依赖。但严格同步也不能消除仪

器响应时间、人工读数延迟或报文错误。时间是观测系统的一项实际测量条件，不只是表格顶端写好的数字。

还必须分开观测时刻、报出时刻与收到时刻。一份八时观测可能十时才传到中心；若中心在九时必须发布预报，它当时无法使用这份后来到达的资料。历史评价若拿事后完整资料检验当时业务判断，会把信息条件倒置。通信加速的意义，不只是让同一张图更快画成，而是改变预报发布时实际能够拥有的状态估计。

同时性也有尺度。研究一个移动缓慢的大范围系统，几十分钟差异可能尚可处理；研究快速发展的局地对流，同样延迟便可能重要。没有一个脱离任务的“足够同时”。观测频次、传输时延与目标过程的时间尺度必须相称。若新天气对象具有更快演变，旧网络即使位置不变，也可能因为采样过慢而不再适用。

这使一张天气图具有四条时间线：自然天气继续演变，仪器在某刻取样，报文在另一刻传播，分析者在指定截止时刻形成判断。历史写作应保留它们的差别。气旋不会等人收齐数据才移动，而分析图也不能因标了某个时刻就被当成没有估计误差的自然切片。公共天气的科学身份，正来自在这些不同时间之间建立可检查的对应。

同时天气场 S 由带时间戳的读数序列 D 形成，其 E 包括时空对齐、仪器响应与天气演变之间的约束。通信提速缩短了资料到达时的年龄，却不修复错误时间戳；反过来，时间戳准确但报文迟到，资料又可能错过本次预报的截止时间。可靠的场图需要这两个环节同时满足任务要求。保留原始时间与延迟，也让事后的重建能够区分采样误差和信息不可得。

地方时间与统一时间的转换还牵涉经度。若报告采用地方太阳时，同样写作正午的两个站，实际可能并非同一全球时刻。转换规则清楚以后，这种差异可以被处理；若旧档案没有说明时间制度，后人重建天气图就会遇到系统不确定性。时间标签本身也有历史，需要知道它在当时的记录实践中究竟意味着什么。

◆ 36.3 标准化不能抹去站址

同一型号温度计，在阳光直射、通风遮蔽和靠近墙面的条件下，可能给出不同读数。差异并不都属于仪器坏了，其中一部分来自它与周围能量交换方式不同。若要比较空气温

度，就必须规定暴露方式，并保存站址变化；若要研究墙边微气候，则这种差异本身又可能是目标。标准化始终相对于研究对象，而不是让所有仪器在任何地方都读出一样的数字。

海拔也是如此。高山站与海平面站的气压不同，不能直接把数值低解释成低压天气系统。把气压归算到共同高度，有助于比较大尺度分布，却引入温度结构等假设。归算不是把误差洗掉的手续，而是一个物理推断。当地形复杂或近地面温度结构特殊时，推断可能出现系统偏差。天气图上的同一类数值，必须知道它是直接观测还是经过归算。

站址周围的变化也会改变记录的含义。植被、建筑、地表覆盖和仪器位置变动，都可能影响局地读数。对于当天的预报，这些信息影响代表性；对于长期序列，变化还可能被误判成气候趋势。虽然长期气候问题将在后文专论，这里已能看到一条共同原则：元数据不是资料之外的行政附件，而是判断两个数值能否比较的证据。

校准与代表性应分开。校准良好的温度计可以准确测到它所在微环境的温度，却未必代表模型网格内的平均状态；两个仪器完全一致，也可能共同受到不合适暴露的影响。若只追求仪器精度，而忽略空间尺度，观测反而容易被赋予过高权重。一个数值的小数位数不能直接转换成对区域天气的确定程度。

共同标准也不能消灭地方知识。熟悉海岸、山谷或港口的人，可能辨认出仪器网络稀疏处的特征；这些经验需要转成可记录、可比较的条件，但不应因尚未采用统一编码就被一概排除。另一方面，地方经验也必须接受跨事件检验，不能以“这里一直如此”免除证据责任。公共气象的成长，在于让地方能力进入共同检验，而不是简单用中心规则取代地方观察。

这项工作具有主体能力的生成面向。观察者要学会稳定识别云类、判断仪器异常、记录缺测原因；中心分析者则要学会区分真实天气差异与站址差异。标准一旦形成，不只是物被测得更一致，人也被训练成能够产生可比证据的主体。SIO 复合体的变化在这里有具体落点：操作能力、对象特征与公共比较同时被重新组织。

但“统一”不等于压低自然多样性。若各地确实存在差异，合格标准应使差异更清楚，而不是把它们修正到同一个平均值。任何均一化或校正都需要保留原记录、说明依

据，并检查是否误删真实变化。否则，公共天气可能变成制度期待的天气，失去自然对分类的反作用。

降水观测又带来积分时间问题。一个雨量读数通常是某段时间内累积的量，不是读数瞬间的雨强。两个站同在八时读表，若累积起点不同，其数值也不能直接比较。积雪、风对收集效率的影响以及蒸发损失，还会改变仪器实际捕获的水量。共同天气场由许多不同类型的量组成，不能把全部变量都当成同一种瞬时采样。

网络还会受到交通与人口分布影响。容易到达、通信方便的地区往往资料较密，海洋与偏远地带则可能稀疏。因此，一张图中的“主要天气特征”有时也反映哪里更容易被看见。研究者应比较观测覆盖与目标过程的分布，防止把资料空白当作天气平静。补站的价值正可以从这种可见性偏差中提出，而不只是追求地图形式上的均匀。

观测空白还可能随天气改变。风暴使通信中断，恶劣海况改变航线，极端低温使仪器失效，都会让缺测与目标现象相关。若只分析成功收到的报文，某类天气可能被系统性低估。网络设计因此还需要冗余和失效记录，使“没有报告”保留其可能的物理与技术含义。

原始报文与后期订正应分别保存。这样，某个异常究竟在当时已被发现，还是后来比较邻站资料后才被修正，就能够重新辨认；历史天气的再分析也不会抹去当时业务实际面对的信息条件。

◆ 36.4 新场图回写观测网络

当分散记录组成天气图，空白便获得新的意义。单个地方的观测者只知道自己没有邻站；中心图则能显示某个海区或山地长期缺少约束，导致某类系统的路径难以判断。新的场对象因此可能反向提出补站、增加频次或改变报告项目的要求。对象不只是观测网络的产品，也能够成为重排网络的理由。

网络改变还受航海安全、军事通信、行政覆盖与科研任务共同影响。一个新站后来有助于定位风暴，并不说明当初建站只由这种需求推动。要确定具体的回写，需把天气分析中暴露的空白，与建站理由、站址选择和实际采样项目连接起来。若前者要求增加海区约

束，而后者只改变陆上通信覆盖，二者就尚未形成所讨论的因果链。一般机制与特定历史行动因而需要分别说明。

教学上可以比较两种网络优化目标。第一种追求地图覆盖均匀，第二种追求降低关键天气系统的不确定性。二者可能给出不同站点：均匀网格填补最大空白，目标观测则可能把资源集中在系统上游或变化迅速区域。没有一种目标在所有任务上最好，航海服务、研究与公共预警需要协商。网络之善不只是数据最多，而是其产生的差异能否回应明确任务。

反馈也可能暴露标准本身的问题。某类仪器在强风雨条件下反复失效，说明缺测不是随机的；最需要观测的天气恰好最容易丢失。若忽视这种选择性，统计资料会系统低估极端情形。改进可以是更换装置、增加冗余或调整质量控制，而不是简单把异常读数删除。天气对象对观测条件的回写，在这种失败中最容易被看见。

共同天气由可比读数、时间对齐、空间代表性与使用反馈持续结合。一个站点能够准确测量的局地状态，经过位置与时刻标记，才可能成为区域关系的一部分；区域图又显示哪些地方和时段仍缺少约束，推动网络调整。标准在这一往返中有双重作用：它维持已有资料的可比性，也必须让观测失败留下可供修订的记录。地方差异由此得以保留，并进入共同判断。

当共同天气场已经能够被绘出，新的问题随之出现：图上的低压中心是否只是某一时刻的形状，还是具有生成、移动和消亡过程的系统？等压线怎样与温度、风和降水共同形成可以追踪的气旋与锋面？下一章将从同时图走向时间序列，讨论天气对象如何获得生命史。

第 37 章 天气图、气旋与锋面 从压力分布到生命史模型

37.1 压力中心与移动过程

一张天气图上，几条闭合等压线可以围出低压中心。它给分析者一个位置，却还没有给出完整气旋。中心可能加深、减弱、分裂，外围风与降水也不必简单围着中心同心排列。若只把等压线当成对象，就容易把一个动态系统压成某一时刻的轮廓。气旋思想的增长，在于让压力、风、温度和天气现象进入可以沿时间追踪的共同结构。

先看一个连续图示的教学问题。第一时刻在西边有一个低压，第二时刻东边出现更低的中心，能否断言同一个低压向东移动？还要比较中间时刻、周围风场与压力变化。有时是旧中心减弱、新中心发展，连接两点得到的“路径”并不代表一个物质团块从西走到东。天气系统的身份不是身份证式固定标签，而是通过连续演化与动力联系维持的过程同一性。

等压线具有诊断价值，却不是风的全部原因说明。在旋转地球上，压力梯度、科里奥利作用、摩擦以及加速度共同限制风场。某些尺度上可以采用近似平衡，近地面或快速变化区域则需要保留额外作用。若把风画成永远沿等压线流动，就会漏掉跨等压线的辐合；若把风一概画成从高压直冲低压，又会抹去旋转效应。图形应当由适用的动力关系支撑。

这里也藏着空间分辨率问题。稀疏站点之间的插值可以画出光滑中心，但中心位置可能随新增资料显著移动。某条狭窄风向转折带，若站距太大就会被平均掉；相反，孤立误读也可能被插值放大成小尺度结构。图上的线越细，不代表证据越细。气旋的公共身份需要把插值规则与观测覆盖一同带进讨论。

从同时图走向时序，还带来新的检验方式。一个形状解释如果只对某一张图成立，很容易依赖偶然；一个生命史模型则要求相邻时刻的变化具有可预测联系。暖区怎样移动，风向何时转变，降水带与压力中心的相对位置怎样变化，都可以成为下一张图的检验任务。对象因此不仅概括已有记录，还开始要求未来记录以某种方式出现。

压力中心的加深还需要质量收支解释。在静力近似下，地面气压与上方空气柱的重量有关；气柱净失去质量，地面气压才会下降。地面附近辐合把空气带入气柱，若同时有更强的高空辐散，气柱仍可以净失去质量。由此可见，低层空气向中心聚集与低压加深并不矛盾，但必须放在整个柱体的收支中判断。只看近地面风箭头，无法独立决定气旋是否发展。

这条教学关系也说明为什么高空观测后来如此重要。地面图能够提示结构，却不能完整给出垂直交换与上层输送。两个地面压力形状近似的系统，可能因上层配置不同而走向不同未来。天气生命史因此不是把地面图按年龄排序，而是追踪三维质量与能量转换怎样允许某个阶段出现。

◆ 37.2 1919 年模型的新规定性

卑尔根学派的组织形成以 1917 年为重要时间点，Jacob Bjerknes 关于移动气旋结构的论文发表于 1919 年。不能把组织建立、结构模型发表与后来成熟的锋面生命史全部压到同一年。[B14] 这些节点之间的关系，正说明一种研究共同体如何逐步把观测密度、物理解释与图式表达连接起来。

1919 年论文的原文开头指出，日常图表的站距和观测间隔不足以直接支持所需的细致动力分析，作者因此利用更详细的北欧记录，并结合天气图与天空观察。文中以当时的术语讨论两类辐合线及其与暖区、上升运动、云雨分布的联系。[R37a] 这里最值得保留的不是把旧术语全部换成今日名称，而是看见一个具体研究改变：系统内部不再只剩环绕低压中心的近似螺旋，温度结构与不同天气带获得了独立辨认地位。

这可以用一段现代教学解释展开，而不把它冒充原论文的逐句重述。相邻冷暖空气具有不同密度和热力性质，边界附近的运动会使暖空气抬升或改变位置。抬升如果达到凝结条件，便可能形成云与降水；抬升的范围和强弱不同，天气带的宽窄与性质也会不同。于是，温度差、风场、垂直运动与云雨不再是四张互不相干的图，而是能够相互限制的结构。

这一联系并不表示只看见云带就能唯一画出锋面。湿度不足时，显著边界也未必产生强降水；地形抬升可以形成类似云雨；局地加热又可能产生对流。需要联合温度、风与时

序，才能把“有雨的地方”与锋区分开。一个新概念的力量，不在于把所有现象都归给自己，而在于提出何种组合才算它的证据。

锋面也不是一堵没有厚度的墙。图上的线通常是一个有限宽度、具有三维倾斜结构的过渡区在某种分析面上的表示。地面线的位置，不等于整个高空边界的位置。若忽略这一点，就难以理解为何降水可以提前出现在地面锋线之前，也难以理解同一位置在不同高度可能处于不同气团关系中。天气对象因此从平面上的线获得立体规定。

更深的变化在于研究主体的辨认能力。过去某站记录的风向突变、温度变化和降雨时段，可以分别放进不同栏目；在锋面模型中，它们成为判断一个过程经过的联合线索。观察者不是仅多记住一个名字，而是学会以新的关系阅读原有记录。这种能力又会影响下一次观测：哪个时刻值得加密，哪种突变不能当作误差删去，哪些高空资料能够区分解释。

气旋—锋面系统的形成，把原先分列的风向、温度和云雨报告组织成共同判据。在 SDE 意义上，这种对象 S 的规定又改变了差异序列 D：时间加密应对准预期的锋面经过，垂直观测则检查地面线与高空结构是否相容。E 表现为温度差、运动和凝结条件的相互限制，断开其中任一联系，云带就难以单独承担锋面识别。新的辨认与协作能力由这组操作形成，自然气团边界则有其独立的物理历史。

锋面位置还可以从一个固定站的时间序列与空间图互相检查。若地图预期暖区将经过，站上应当出现与该解释相容的温度、风与湿度变化；若只有气压变化而热力特征不符，可能是锋线位置、气团分类或移动速度有误。这种检查把单站经验与区域图连接起来，避免中心分析完全脱离地方记录。时间序列不是地图的附属，它能够反驳地图。

锋区的倾斜则让“经过”具有高度差异。高空的暖空气可能先到，地面冷空气仍滞留；降水和云的先后因此不能只按地面线是否到达解释。对居住在山谷或高地的观察者，相同系统也可能呈现不同顺序。统一模型需要把地形和垂直位置作为条件，而不是要求所有地方都复制标准图中的一段天气。

◆ 37.3 数学表述与锢囚扩展

一个结构模型仍不足以说明系统怎样发展。1922 年 Bjerknes 与 Solberg 关于气旋生命史和极锋理论的研究，使阶段变化成为明确主题；Solberg 的数学工作与 Bergeron 对

气团、锢囚过程的研究，各自参与了模型的扩展。[R37b][B14] 这是一条多人、多阶段的研究链，不能把所有概念都归给单篇早期论文。

以常见的理想化生命史作教学说明：在冷暖空气交界附近，扰动可以发展成波状结构，低压与锋区共同演变，暖区在地面上形成一定范围；随着相对运动和结构变化，暖空气可能逐步离开地面，冷暖锋的关系转变，出现锢囚阶段。这种描述把几幅天气图组织成一个有方向的过程，而不只是按形状分类。可是，阶段序列是理想模型，不能要求每一个实际气旋按同一速度、同一几何依次走完。

“冷锋追上暖锋”是便于入门的说法，但若把它当作完整动力机制，会过度简化。锋区本身有三维结构，温度分布与风场共同演化，锢囚也不能只由地面两条线的距离判断。系统可能受地形、海陆差异及高空流场影响，出现弯折、分裂或不同的暖区结构。一个生命史模型若无法容纳这些差异，应当缩小适用范围或发展新的类别，而不是把所有例外当作绘图不规范。

数学表述的价值，是让这些结构承担守恒与运动的责任。空气不能任意汇聚又不改变质量分布；热力差异不能在没有输运、混合或加热的情况下消失；风的弯曲与加速需要相应力学联系。概念图能帮助辨认，方程则限制概念图允许怎样变化。两者并非彼此取代：没有可辨认结构，方程结果难以解释；没有动力限制，结构图又可能成为任意叙事。

由温度水平差异引出的热成风关系，提供一条理解立体结构的现代教学入口。在适当的静力与地转近似下，水平温度梯度与地转风的垂直切变相联系。因此，地面冷暖分布不是只影响体感，它还限制高空风结构。反过来，高空资料可以检查地面锋面解释是否与三维场相容。这不是把后来的全部动力理论倒投给 1919 年的作者，而是说明为何生命史模型最终必须越出单层天气图。

失败案例在这里具有创造性。如果某类气旋反复不按标准序列发展，问题可能是尺度、区域或强迫条件不同，也可能是旧分类把两种机制混在一起。研究者需要比较这些失败的共同特征，而不是逐例找借口。对象的分化有时比维持统一模板更重要：一个大类被分成几种可检验类型，可能比一张“适用于所有气旋”的图更有解释力。

理想生命史的适用范围，应由实际气旋的演变来限定。共同物理关系使不同个案可以比较，偏离标准序列的个案则帮助辨认哪些阶段取决于特定环境。若某类系统持续缺少预

期的暖区演变，分类需要解释这种缺失，而不能只把图形修饰得更接近典型。共享模型的稳定，来自它既能组织常见过程，也能把重复出现的差异交给新的研究。

暖空气与冷空气并不是两种永远保持身份的液体。气团沿途与地表交换热和水分，并发生混合，其特征会逐渐改变。气团分类的意义，在于某些性质在一定尺度和时段内具有相对一致性；超出这一范围，分类就需要更新。若一团空气离开发源区很久，仍只凭来路给它贴固定标签，可能忽略它已经取得的新热力结构。

锋生与锋消也不只是两条线靠近或远离。流场可以加强温度梯度，使原先宽缓的过渡变窄，也可以通过拉伸、旋转与混合使其减弱。一个边界的身份因此来自持续维持的差异，而不是在地图上画下后便永远存在。这个事实为发生学提供了格外清楚的案例：对象不是每一时刻都由同一批空气分子组成，其相对稳定来自过程关系的维持。

这种过程同一性还使“生命史”成为有用但有限的比喻。气旋没有生物个体那样固定的遗传程序，它的阶段来自能量、质量与环境流场的耦合。用出生、成熟、消亡帮助理解，可以提示时间方向；若把它变成必然命运，就会遮蔽系统可能的再发展与相互作用。生命史模型必须保留分支，而不能只保留一条故事线。

能量观点进一步说明温度结构为何重要。水平上并置的冷暖空气包含能够在适当运动中转化的可用势能；空气重排时，一部分可以转成动能，又被摩擦和其他过程耗散。并非任何温差都能无条件释放同样能量，旋转、稳定层结和边界共同限制转化路径。因而，锋面模型不只是给天气现象找到位置，也把系统的发展连接到能量来源与转换。

这里可以与地幔动力作有限比较。两者都需要说明温差怎样转成运动，但大气中的旋转、层结与水相变化改变了能量转换条件，地幔的长期运动则受到黏性流变和材料相态的强烈限制。相同的“温差驱动”说法，并不对应相同的计算。对气旋而言，温度梯度与风的垂直结构需要共同解释；对地幔而言，温度与黏度的反馈则影响热的输送。比较应落实到这些不同的联系。

锋面模型的另一个收获，是把预报中的空间顺序转成时间预期。若一个系统以一定方向和速度接近，前方云带、持续降水、暖区与后方突变可能依次影响某地。然而，这个转换要求系统结构在经过期间没有发生过快改变。若气旋正在迅速发展，不能把一张静态图

按固定速度扫过站点，就当作完整预报。移动与演变需要分别估计，再在同一过程中结合。

这也是连续天气图比单张“典型图”更有教育价值的原因。读者可以比较哪些特征保持、哪些特征新生，检查上一时刻的解释能否承受下一时刻。图式的统一性由这种跨时检验维持，多样性则由真实演变保留。共同模型不是答案册，而是能够持续生成并接受检验的问题结构。

因此，气旋研究的对象不是孤立的压力低值，而是跨变量、跨高度和跨时间维持联系的过程。这种联系一旦断裂，分类便应重新检查。

◆ 37.4 从认识天气到重排业务

锋面和生命史模型进入预报实践后，图上的标记会改变注意力的分配。预报员不仅看低压中心在哪里，还看温度边界怎样移动、暖区怎样收缩、云雨带与高空结构是否一致。某个站的突然风向转变不再只是孤立读数，而可能成为更新锋线位置的证据。图式由此具有操作意义：它把分散报告连接到下一步判断。

业务能力不能仅凭符号传播证明。许多单位采用同一种锋线画法，说明语言统一，却未必说明预测更好。应比较模型是否在独立天气过程上改善位置、时间或天气类型的判断；还应考虑预报员经验、观测增加和通信改善的共同作用。若新方法只在事后解释上灵活，却不能在事前限定结果，它的实用贡献就需要重新评估。

一个清楚的检验是保存发报时的分析，而非只保存事后修订图。预报员当时把锋面画在哪里，用了哪些已收到的观测，后来又因什么资料移动它，这些记录能区分实时能力与回顾性合理化。若出版图已经吸收随后到达的资料，再拿它说明当时预报准确，就会把分析修订误当作预见。思想史与技能评价在这里共享同一证据纪律。

业务失败也可以回写观测。若地面资料足以定位锋线，却不能判断发展速度，就可能需要高空信息；若快速变化在六小时图之间发生，增加时间频次可能比增加静态站点更有价值。具体网络改变应由档案确认，但其方法逻辑已经清楚：新对象产生新的未知，新的未知推动有针对性的互动，互动再检验对象定义。

公开资料中可以审查的，是特征重新归组与操作改变之间的关系：锋面身份如何使若干单站变化成为联合证据，预期何种经过又怎样提出加密观测要求。SDE 将这些关系组织为复查路径，并不据此推定科学家曾按同一顺序思考。具体业务改变还包含试错、协作和技术条件，需要由当时记录说明。如此，模型的传播与它真正改变的判断能力才能分开评价。

气旋与锋面世界的成熟，也没有使数值计算成为多余。生命史模型提供整体辨认，却难以逐点计算复杂状态的未来。若要把当前三维状态推进到下一时刻，需要更明确的方程、初值和求解方法。1904 年的物理纲领、1922 年的手算尝试与 1950 年的电子计算，将把这一要求放到另一条发展线上。下一章讨论的不是计算机如何取代看图，而是天气怎样获得一种新的可计算、可检验存在方式。

第 38 章 从物理纲领到电子预报

1904、1922 与 1950

◆ 38.1 1904：天气成为初值问题

在一张天气图上辨认气旋，与从当前状态计算下一张图，是两种不同任务。前者可以依靠形态、经验和概念模型，后者要求说明状态由哪些变量构成、变量之间遵循什么关系，以及怎样把这些关系转成有限时间内可执行的计算。数值天气的思想史因此不能从电子计算机出现才开始，也不能在写出流体方程时就宣布已经完成。

Vilhelm Bjerknes 在 1904 年的纲领中，将天气预报联系到两项条件：充分知道某一时刻的大气状态，以及充分知道状态演变所遵循的物理规律。该文英译本还保留了对当时海洋和高空资料缺口的讨论。[R38a] 预报因此被规定为具有观测前提的初值问题。物理规律说明应当计算什么，观测网络决定其中哪些状态能够取得。两项条件需要分别推进，又须在同一个状态表示中相接，才能形成可执行的预报。

“当前状态”比一张地面气压图丰富。风有不同方向的分量，温度、压力、密度与水分具有三维分布；边界也会影响演变。某些变量通过状态方程或平衡关系相联，不是所有数值都可以独立任意指定。初值必须形成一个物理上相容的状态，而不是把各台站最可信的读数逐项拼起来就结束。每个数值单独合理，合起来仍可能要求不合理的加速度。

以现代教学语言表达，若状态记作 x，模型可写为 $dx/dt = M(x)$。形式上，知道初值 x_0 便可以沿时间推进。但 M 不仅包含动力方程，还包含采用的近似、边界与未解析过程的表达；x_0 也只是有限观测形成的估计。数学符号把问题压缩得很短，实际责任却没有减少。若状态维数很大，资料不足；若模型遗漏重要过程，规律近似不足；若求解耗时太长，预报失去时效。这几种失败不能用同一个“算不出来”混合。

方程的闭合尤其关键。大尺度运动方程可以写出，但云、湍流、辐射和地表交换等过程跨越不同尺度。某些过程需要用参数化把未解析作用传给可计算变量。参数化不是任意

补丁，也不是自然定律的完整替身，而是有适用范围的关系。若分辨率或目标过程改变，原有表达也可能需要改变。这一问题后来持续存在，并不会被计算速度单独消除。

初值问题还改变观测的目的。测量不再只是为了描述过去天气，而是为了填充一个能向未来推进的状态。哪些高度、哪些海区、哪些变量最缺少约束，能够从模型敏感性中获得新的解释。前两章所讨论的共同天气场与结构辨认因此成为必要基础，却不再是终点。观测与方程必须互相提出要求：方程说明缺什么状态，观测说明哪些状态实际上能被可靠取得。

1904 年的纲领重定了目标 S：未来天气成为从当前状态演变而来的计算对象。E 包括状态变量之间的物理关联与可解条件，D 则需连接观测取样和计算推进。此时三者之间仍有实际断点。高空状态缺测，使方程所需初值不足；求解方法与资源不足，又使已知状态难以及时推进。纲领明确了这些断点的位置，后来的手算与电子计算才有可能分别检验哪一处已经接通、哪一处仍在限制预报。

初值的空间完整性也并非“每个网格都有数字”这么简单。缺测处可以插值填充，但插值可能只让数组完整，没有增加相应信息。若某个区域的风场完全依赖邻近站外推，随后计算对那里很敏感，预报不确定性就应反映这一弱点。模型状态的数值完整与证据完整不同；前者是程序运行的必要条件，后者才关系到预测信心。

边界条件也属于初值问题的实际范围。有限区域大气受区域外流动影响，只知道区域内部此刻状态，未必足够决定很久后的演变。若采用全球模型，侧边界问题改变了形式，地表与上边界等条件仍在。把问题称为初值问题，不意味着所有边界与外部强迫都可以删除。它是在明确系统边界以后，对状态演变所作的数学规定。

◆ 38.2 1922：失败计算留下什么

Richardson 于 1922 年出版《用数值过程预报天气》，把方程、网格、计算表与具体算例推进到可以检查的程度。[B15] 他的尝试著名地给出了失真的气压变化。后来针对这一算例的数值重建研究，将巨大误差与初始风压场不平衡所激发的快速振荡联系起来。[R38b] 本章不把这一后来的诊断伪装成作者当年已经完整掌握的解释，也不把算例失败简单归结为手算太慢。

为什么初值不平衡能产生巨大倾向？在较大尺度的大气运动中，某些力常接近相互抵消。压力梯度很大，旋转作用也很大，净加速度却可以较小。若两项各有不大的观测误差，它们相减后的误差可能相对于真实净值十分显著。类似地，几个方向的质量输送大体抵消时，小的风误差也可能造成虚假的强辐合或辐散，进而产生不合理的气压变化率。

可以用一个教学数字说明，而不代替 Richardson 的原表。设两个相反作用分别为一千和九百九十九，净值为一。如果每项各有百分之一误差，误差量级约十，已经远大于净值。于是，“每个输入都很准确”与“输入组合后给出的倾向准确”不是同一判断。天气计算对相互平衡的场提出要求，不能只逐变量检查小数点。

这并不意味着真实大气永远处于精确平衡。快速波动和局地不平衡本来就是大气的一部分。问题是，当前模型与资料能否正确表示这些快过程，以及预报目标是否需要解析它们。若目标是较缓慢的大尺度天气演变，而初值误差激发了强烈但不真实的快速振荡，计算就可能被无关噪声支配。如何保留重要动力、过滤不适当的自由度，成为后来近似方法的关键。

离散化还带来第二类困难。连续方程中的空间导数，被相邻网格数值之差近似；时间导数，被有限步推进近似。网格太粗会漏掉结构，步长不当会导致误差放大。即使初值完全正确，一个不稳定的算法也能产生爆炸式增长；即使算法稳定，过粗网格也可能让波传播速度错误。数值结果失真时，必须区分初值、物理模型和算法三种来源。

Richardson 留下的价值，正在于失败以一种可追踪形式出现。计算过程、变量安排与结果被公开，后人才能研究哪里需要改动。如果只公布一个错误答案，没有方法，失败难以累积知识；如果只公布方法纲领，不承认具体结果失败，也无法暴露初值与求解之间的张力。科学创造有时不是给出可立即使用的成品，而是把困难变成别人能够接手的明确结构。

计算时效当然仍是实质限制。预报六小时后的天气，若计算需要远超六小时，就不能服务实时判断。但提高速度只是使计算有机会赶上自然，并不保证结果正确。把原来失真的流程加速一万倍，可以更快地得到失真。电子计算的历史成就必须与物理近似和初值处理共同理解，不能写成机器终于替一个已经完全成熟的方法按下快进键。

这里还要区分回报与预报。使用历史某日资料计算其后天气，可以检验方法，却不代表当时真的在事件发生前发出了预报。回报有利于反复研究同一案例，也容易产生调参泄漏：研究者若已经知道结果，可能不自觉地选择最有利设定。因此，从可复算尝试走向业务技能，需要在未见结果的情境中保持方法和评价规则。

Richardson 算例中的 D 同时包含输入差异、差分操作和随推进出现的误差。E 则把初始场的平衡、算法稳定性与所保留过程的时间尺度联系起来。三者中任何一项不相容，都可能使未来状态 S 失真，而且相似的失真图像未必来自同一原因。失败留下的计算记录因此具有诊断价值：后人可以分别改变初值处理、求解方案或物理近似，比较哪一项调整真正改变异常倾向。技术突破由这种拆分逐步获得方向。

离散导数为何容易放大观测噪声，也可用教学例说明。相邻两点的差除以网格间距，用来估计梯度；如果两点各有独立噪声，差分中的噪声量并不随着间距缩小而自动减小，除以更小间距后反而可能更显著。网格变细而观测质量不变，不保证初始梯度更准确。因此，分辨率、资料密度与误差处理需要共同推进，不能把更密网格当作无条件进步。

质量连续关系又把这种梯度误差传播到其他变量。水平风的散度若估计错误，垂直运动和压力倾向会受到影响；温度输送的误差则可能改变后续平衡。一个局部读数的误差并非只留在该读数所在栏目，它通过方程进入整个状态。数值预报的困难与机会都来自这种耦合：既能用物理联系补足信息，也会把不相容输入放大成系统问题。

◆ 38.3 1950：简化方程与电子计算

Charney、Fjørtoft 与 von Neumann 的 1950 年论文讨论有限区域上的正压涡度方程数值积分，报告了利用 500 毫巴资料进行的四次二十四小时试验，并分析边界条件、差分稳定性和误差问题。[B16][R38c] 这个节点的重要性，不在于把完整大气原封不动交给电子机器，而在于选择一种受到物理判断支持、又能在当时条件下求解的近似系统。

正压模型把垂直结构大幅简化。在教学性的单层、近似无辐散情境中，绝对涡度可以随水平运动输送，写成 $\partial \zeta/\partial t + u\cdot\nabla(\zeta+f)=0$，其中 ζ 为相对涡度，f 反映地球旋转随纬度的变化。这是理解思路的理想形式，不代替原论文全部坐标、离散与边界推导。它说明模型选择了什么：通过一个代表性层的旋转结构，描述较大尺度流动的演变。

为什么涡度适合这样的简化？直接推进全部原始变量时，系统包含多种时间尺度，某些快速波动对目标天气的慢演变未必必要，却会增加初值和计算负担。经过适当近似，可以突出较慢的大尺度关系，减少不希望被初值噪声支配的自由度。这不是说快过程不存在，而是把预测任务限定到模型能够承担的尺度。简化因而是一项有物理内容的选择，而非只为节省机器时间随意删项。

涡度与风之间还需要一个反演关系。若以流函数表示无辐散风，涡度与流函数通过空间微分联系；推进涡度以后，需要解相应的椭圆问题恢复流场。这样，一个时间步并不是把每个网格点单独往前推，而包含全区域的空间联系。边界条件会进入解，区域外的影响不能凭空消失。有限区域模型因此必须说明边缘怎样处理，以及边界误差可能向内部传播多远。

500 毫巴面也不是固定几何高度，而是等压面。其位势高度分布与流场之间可以在相应近似下建立联系。把这一面用作大尺度流动的代表，并不意味着地面天气、高空急流和水汽过程都被完整解析。模型能够较好表示某些高度场演变，不能自动转化成每个地点的降雨预报。预报对象的范围必须与状态变量一致。

数值稳定性是另一项核心创造。以最简单的一维平流教学问题为例，若信息以速度 c 移动，网格间距为 Δx，时间步为 Δt，许多显式方案需要时间步与网格、传播速度满足相应限制，使数值信息传播能够跟上物理传播。常见的 $c\Delta t/\Delta x$ 比值体现这一点，但具体稳定条件依算法而异，不能把一个比值门槛当成所有方程的通用定理。重要的是，网格加密后，时间步常常也必须缩短，计算成本不会只按格点数线性增加。

稳定也不等于准确。一个带强数值扩散的方案可能非常稳定，却把锋利结构抹平；另一个方案可能保持振幅，却产生相位误差，让系统位置逐步偏移。验证需要同时考察误差大小、空间形态与随时间的增长。若只看最终图像大体相像，就可能漏掉对实际预报十分重要的时间和位置偏差。

四次试验能够证明一种路线具有可操作前景，却不是对所有季节、区域和天气型的统计认证。正压近似对某些流动有价值，对具有重要垂直差异和热力发展的问题则有限。网格与资料误差也会影响结果。原论文将这些限制纳入讨论，说明早期电子预报并不是一个无误机器取代人类的故事，而是一种新的、可公开比较的计算实践。[R38c]

1950 年的试验把规定尺度和时效的预报 S，与初值输入、差分推进及误差比较 D 连接起来；E 由物理近似、边界条件、网格和初值质量的相互限制构成。机器使更多时间步能够及时执行，涡度近似则减少需要推进的自由度。两项改变共同成立，才使计算得到的场具有天气意义。若只提高计算速度，却保留初值与边界条件之间的不一致，计算仍会更快地传播输入偏差；若近似删除了目标系统的发展机制，增加时间步数同样不能恢复它。

正压近似的边界可以从热成风关系理解。在相应平衡近似下，水平温度梯度对应地转风的垂直切变，单一代表层便难以同时描述不同高度的运动。气旋发展的能量转换又可能依赖这种垂直结构。因此，正压模式即使能推进某些大尺度形势，对快速发展系统的描述仍可能不足。后续增加层次的意义，在于让被简化掉的垂直联系重新进入状态演变，使上下层之间的相互影响能够参与预测。

涡度方程还让地球旋转的纬度变化进入预测。即使相对涡度变化不大，空气向不同纬度运动时，绝对涡度关系也会限制流场。这使大尺度波动具有不同于平面水槽直觉的传播性质。计算不能只把当前天气图整体平移，因为系统内部的相互作用会改变形状与移动。数值预报真正新增的能力，是按模型关系生成变化，而非更快地执行经验外推。

在网格上解涡度—流函数关系时，一个地点的涡度变化可以通过空间求解影响较大区域的风。这种非局部性来自数学边值问题，不表示信息在真实大气中无限快传播；它反映所采用平衡近似把某些调整过程压缩进诊断关系。任何近似都在改变显式保留的时间尺度。理解这一点，才能判断它为什么高效，也知道何时不适合使用。

计算成本的变化则可以作数量级教学比较。二维网格若两个方向的间距都减半，格点数约增加四倍；若稳定性又要求时间步减半，同一预报时长的推进次数再增一倍，基本操作量可能至少增加约八倍。加入更多垂直层、物理过程与迭代求解，还会继续增加成本。这个估算不代表某一历史机器的实际性能，却解释了为什么物理简化、算法改进与硬件发展必须一起考虑。

早期电子计算还有程序与数据转换的劳动。连续物理量要变成离散数组，观测图要转换为机器可以读取的初值，输出数字又要恢复为可供气象判断的图场。每一次转换都可能引入单位、坐标、舍入或编码错误。数值天气的复合主体因此包括编程、资料处理与物理解释能力；把历史只写成数学家写方程、机器自动出结果，会隐去这些实际承担可靠性的互动。

这一点与 Richardson 的计算表具有连续性。手算表格与电子数组虽然媒介不同，都在规定变量放在哪里、每一步读取哪些量、结果如何传给下一步。电子化扩大了规模并改变了时效，却没有取消计算组织。流程的可审计性，是从个别算例走向持续业务的重要条件。

◆ 38.4 可计算怎样转成可检验技能

一个模型能够输出明天的图，不代表它比不计算更有价值。最简单的比较对象是持续性：假设明天与今天相近。另一个可能是气候平均，或已有经验方法。不同目标需要不同基准。若某种天气变化很慢，持续性本来就不差；若模型只是复制当前场，误差看似小，也可能没有真正预测变化。技能必须相对于明示基准，而不是只相对于完美真值的距离。

评价指标也会改变判断。平均误差可以反映系统偏差，均方误差会突出较大偏离，位置误差则对移动系统尤为重要。一个锋面预测得很强却偏移几十公里，可能在格点误差中受到双重惩罚；一个过度平滑的场反而可能取得较小均方误差，却对局地风险没有帮助。因此，指标应与任务一致，并报告它忽略了什么。没有一项分数能够替代全部物理检验。

评价还要注意训练与检验的时间顺序。若反复根据同一组案例调整参数，案例分数改善可能来自适配，而不是可迁移能力。保留未参与改进的事件或时期，可以提供更严格检验；若气候背景与观测网络随时间变化，还应说明新时期是否与旧时期可比。一次成功试验为路线提供证据，持续跨情境表现才逐步形成可靠封装。

空间平均也可能遮蔽有实际意义的失败。广阔区域中的多数格点变化较小，少数强风暴区域即使预测失准，对全域平均分数影响也可能有限。因此，整体指标与特定天气型检验应并行。前者防止只挑成功事件，后者防止平均掩盖关键过程。统一评价不是只用一个数字，而是让不同指标围绕明确任务承担各自责任。

预报误差还可以分解为系统位置与强度的不同问题。假设模型生成了一个形状大体正确的低压，却比观测偏东；另一个模型把中心位置算对，却明显低估加深。二者可能得到相近的总体误差，但改进方向不同。前者需要检查平流、波速与边界，后者可能涉及发展机制和垂直结构。误差图因此不仅用于评分，也能成为寻找模型缺口的材料。

类似地，误差随预报时效的曲线能够帮助区分问题。初始时刻就存在的显著偏差，提示分析或变量转换需要检查；早期较小、随后系统增长的偏差，可能暴露模型过程或数值传播；某一时段突然恶化，又可能与特定天气型或边界影响相连。没有一种曲线形态能够唯一归因，但它们比一个最终分数保留更多可诊断结构。

数值天气由此改变“失败”的可读性。传统预报失准可能只留下“下雨没有发生”的评价；数值系统能够保存从初值到各中间状态的完整路径，让研究者逐步定位偏离何时出现。这样的透明度并非自动实现，前提是保留计算记录并使用一致版本。计算的增长不仅生产更多答案，也有可能生产更细致的错误史，使改进具有可累积的依据。

预报的可用性还具有发布尺度。区域高度场的成功，需要经过额外转换才能成为站点风雨信息；转换中会加入地形、局地统计或其他模型。用户看到的最后一句预报，可能已包含多个层次的知识。评价某个数值核心时，应分清哪些技能来自核心演变，哪些来自后处理，否则既可能高估核心，也可能忽略后处理的创造。计算对象如何成为使用对象，本身是一条需要检验的传递链。

业务预报还必须在实际截止时刻运行。资料迟到、仪器故障、计算失败和发布延迟，都影响用户能否利用结果。研究试验可以使用完整回顾资料，业务系统则要在不完整状态下按时给出判断。两者之间的转化不是纯粹行政推广，而是新的可靠性问题：方法在缺测和异常条件下是否仍能维持合理行为，失败时能否被识别并处理。

可检验技能需要持续档案。保存初值、模型版本、参数化、观测来源和最终预报，才能判断改进究竟来自哪里。若新版本同时更换资料、网格与物理方案，分数提高并不能自动归因于某一算法。受控比较或分阶段试验可以帮助拆分贡献；无法拆分时，应把成效归给整个系统，避免制造单一英雄技术的叙事。

这种档案还使科学对象继续变化。早期目标可能是一张代表层的高度场，后来目标扩展到降水、风暴、概率与影响。每次扩展都需要新的变量、观测和评价，不是把旧模型结果换个标题。一个预测系统在高度场上有技能，不足以证明它在极端降水上也有技能；一个概率预报可靠，也不等于某条确定轨道能够被长期精确指出。

三个历史节点分别改变了预报链上的不同环节。1904 年明确初值与物理规律的共同要求，1922 年的可检查算例暴露输入平衡和计算推进的困难，1950 年则把适当近似与电

子求解连接成试验。S、D、E 在这里组织的是这些关系的复查：预报目标怎样限定计算，计算误差怎样检验近似，近似的调整又怎样改变所需观测。后来的成功承接了早期问题，同时加入新的技术工作，各节点的贡献由此保持连续而可区分。

但即使资料更好、算法稳定、计算足够快，未来是否就能无限延长地被准确推出？这个问题已经超出 Richardson 式的具体失败。某些确定系统对初值的微小差别会持续放大，使有限信息无法支撑任意长的单轨道预测。下一章将进入洛伦兹 1963 年的简化系统，分清计算不足、资料不足与动力敏感性之间的区别。

第 39 章 洛伦兹 1963
确定方程为什么不保证长期轨道预测

◆ 39.1 三变量系统的建模代价

确定性常被误解成一种无条件预见能力：只要规律不变，未来就应当能够算得足够远。这个推论漏掉了初值信息与误差传播。规律可以明确规定每个状态怎样继续，实际研究者却只能用有限精度知道状态。若相近状态的未来逐渐分开，知道规律与知道自己究竟处在哪条轨道上，就成为不同问题。洛伦兹 1963 年的《确定性非周期流》把这种张力放进一个可计算的简化系统。[B17]

洛伦兹系统的常见无量纲形式为：$dX/dt=\sigma(Y-X)$，$dY/dt=rX-Y-XZ$，$dZ/dt=XY-bZ$。X、Y、Z 来自受热流体运动与温度结构的截断描述，与某个城市的风速、温度和气压并无直接对应。σ、r、b 是模型参数。三个方程既包含线性增长或衰减，也包含变量乘积所体现的非线性耦合。有限变量使反复积分成为可能，耦合又使一个变量的增长改变其他变量的变化率；逐个变量朝固定平衡衰减的图像，因而不足以概括系统行为。

模型的简化有明确代价。从连续流体到有限变量，许多空间自由度被删去；边界、对流形式和参数范围也已经规定。因而，一条计算轨道不能直接当作真实大气的三维演变。它是研究动力性质的模型，不是完整天气预报器。其思想价值在于：即使在这样小而明确的系统中，确定性与长期单轨道可预见性也不能被轻易画等号。

方程的非线性可以从两项乘积看见。X 与 Z 相乘影响 Y 的变化，X 与 Y 相乘影响 Z 的变化。某个变量增长会改变另一个变量的变化方向，后者又反馈到原来的运动。这种联系使系统可能在不同区域间转换，而不是一直朝一个固定值靠拢。但非线性本身不等于混沌，某些参数下系统可以趋于平衡或形成其他较简单行为。不能看见乘积项就宣布不可预测。

可以对平衡点作一个简短教学分析。取 $\sigma>0$、$b>0$，令三个时间导数为零，首先得到 $Y=X$；再由后两式得到零平衡点，以及在 r 大于一时可能存在的两个对称非零平衡

点，在 $b>0$、$r>1$ 时，$X=Y=\pm\sqrt{b(r-1)}$，$Z=r-1$。这个代数步骤说明，系统方程具有明确结构，绝不是随意随机游走。至于这些平衡点是否稳定，以及轨道在指定参数下怎样行为，还需要线性稳定分析与数值或更严格的动力研究。

耗散也能从相空间散度看见。三个方程右端对各自变量的偏导相加，为 $-\sigma-1-b$；在正参数下为负。它表示局部体积元在相空间中受到总体收缩作用，但不能单凭这一条就证明存在混沌吸引子。收缩与某些方向的伸展可以并存，这种组合为复杂的有界行为提供可能。把“总体收缩”误读成所有轨道互相靠近，或者把“局部伸展”误读成轨道必然无限发散，都会失去动力几何的关键差别。

洛伦兹意义的混沌是在明确方程与相空间中规定的动力性质，SDE 三态理论中的混沌态则有其关于规定性与发生的定义。两者使用相同名称，却没有因此建立数学同一性。以下数值实验只讨论洛伦兹系统的扰动演化；SDE 分析另在预测任务、比较操作与评价条件的变化中展开。这样，轨道是否敏感与研究对象如何重新规定，分别由动力证据和研究实践的证据承担。

对相邻状态作线性化，可以更清楚地看见误差怎样进入方程。记微小差异为 $(\delta X,\delta Y,\delta Z)$，则其变化由沿当前轨道变化的雅可比矩阵控制。第一行含 $-\sigma$ 与 σ，第二行含 $r-Z$、-1 与 $-X$，第三行含 Y、X 与 $-b$。系数随轨道上的 X、Y、Z 改变，所以同一个大小的初始误差放在不同状态或不同方向上，可能经历不同增长过程。初值敏感不是一个与状态无关的神秘放大器。

这一线性化也指出“最大误差方向”的意义。某些扰动方向可能暂时收缩，另一些方向被伸展；经过一段演变，误差形状会被旋转和拉长。真实预报中的误差因此不宜总用每个变量相互独立、大小相同的圆球表示。即使只在小系统中，这种几何也要求研究者关心误差之间的关联，而非只关心每个数的小数位。

◆ 39.2 相邻轨道与精度复查

要讨论初值敏感，最直接的计算是从两个很近的初始状态出发，使用同一方程观察差异。但这个程序有一个陷阱：数值算法自身也会引入误差。若步长过大，两条轨道分离可

能主要来自计算不准确；若积分很久，即便相同初值用两种算法也可能不再逐点重合。因此，展示一幅分叉图之前，必须把动力差异与数值差异分开考察。

本书为说明这一点执行了一项小型教学计算，不把它当作 1963 年原算例的复现。参数取 σ＝10、r＝28、b＝8/3，初值分别为(1,1,1)与(1＋10^{-5},1,1)。采用经典四阶 Runge—Kutta 方法，分别用 0.01、0.005、0.0025 与 0.00125 的时间步长积分到无量纲时间 20，并记录三个变量的欧氏距离。[R39a] 这里的时间单位属于模型，不能换算成真实天气的小时或天。

结果不是从一开始就按同一倍数持续增长。以最细步长为例，时间 5 的距离约为 1.25×10^{-5}，时间 10 约为 1.76×10^{-5}，时间 15 约为 2.38×10^{-4}，时间 20 约为 9.24×10^{-3}。初始距离为 10^{-5}，经过这段计算后放大了约九百倍。这个数字只描述本次指定初值和区间，不是系统的通用增长率，更不是大气的统一预报极限。

步长比较使解释更有节制。在时间 20，0.005、0.0025 与 0.00125 所得相邻轨道距离均在约 0.00924 附近，而 0.01 给出约 0.0107，偏离较明显。相同初值轨道在不同步长下也存在差异，且到后期比早期显著。因此，这项实验支持在所选参数与有限区间内存在稳定可见的扰动放大现象，却不构成对精确轨道的严格误差界证明。细步长间的相近结果是一项数值检查，不是数学定理。

为什么需要同时比较两类距离？第一类是在同一步长下比较两个初值，主要用于看扰动演化；第二类是保持初值相同而改变步长，用来估计算法敏感性。只做第一类，很容易把算法问题归给动力；只做第二类，又可能把动力敏感导致的长时逐点失配误认为方法完全无效。两类比较共同限定我们能够说什么。

有限精度也不只有步长。浮点舍入、输出截断与初值记录位数都会改变起点或计算路径。若研究问题是很长时间后的具体位置，需要更严格的高精度与误差控制；若问题是某种统计特征，则可以采用不同检验。但不能因为长期轨道无法逐点重合，就自动宣布统计量可信。不同问题需要独立的收敛检查。

还可以改变初始扰动方向。仅沿 X 方向扰动是一种选择，沿其他方向可能在初期表现不同；误差增长与状态所在区域也有关。某些阶段差异暂时缩小，之后再扩大，并不反对初值敏感。所谓指数增长常是某种局部或平均描述，不是要求每一瞬间距离都单调增加。

把轨道距离画在对数坐标上可以帮助观察，但拟合一条直线之前仍需说明区间与饱和影响。

当距离接近系统典型尺度时，增长不能继续按同一指数无限延伸。有界吸引域使两条轨道最终可能都留在同一大区域，距离出现起伏或饱和。此时“失去单轨道可预测性”不等于“系统飞向任意远处”。我们失去的是对具体状态位置的精细确定，并非失去关于允许区域的全部知识。这个区别对天气预报和气候统计的关系尤其重要。

步长和精度比较也规定了这个教学实验的可撤回条件。若在更小步长、更高精度或其他可靠方法下，所观察的分离消失，就需先检查算法；若扰动放大在多种受控设置下保持，则得到有限区间的数值支持。参数、算法及结果见附录四，可据此复查两类距离。它提供的是计算证据：严格的混沌判定及精确轨道误差界，还需要相应的数学分析。

离散算法实际上定义了一个与连续方程相关、却不完全相同的更新映射。四阶 Runge—Kutta 在每个时间步评估几个中间斜率，并加权组合得到下一状态；它的局部误差阶数描述步长趋于零时的近似性质。这个阶数不能直接保证任意长积分后仍与精确解逐点接近，尤其在误差会被动力放大的系统中。算法理论与动力理论需要一起理解。

因此，比较步长时应先固定实际积分终点，再逐步缩小步长，检查同一有限区间上的结果。若不断延长终点，同时要求两种步长轨道永远重合，就是把两个极限混在一起。有限时间收敛不等于长期精确跟踪，长期逐点失配也不反对有限时间数值解的有效性。这种区分让敏感系统仍然可以被认真计算，而不落入“既然混沌，一切数值都没意义”的误解。

本次教学计算还有一个可学习之处：早期距离几乎不增长，后期才明显放大。如果只截取时间零到五，可能低估敏感性；如果只截取最后一小段，又可能把局部增长率过度推广。实验设计必须说明观察区间，而不能挑选最戏剧化的画面。把所有指定检查点同时报告，比只展示最终分离更能体现计算的真实内容。

◆ 39.3 非周期不等于一切随机

确定性指的是，在适用条件下，同一初值与同一方程规定同一演变；周期性指轨道经过一定时间后重复。两者没有逻辑上的等价关系。确定系统可以非周期，随机系统也可能

呈现显著周期成分。把“无法长期精确预测”直接翻译成“没有规律”，等于混淆方程结构、初值知识与预测任务三个层次。

一条非周期轨道也不必满足所有混沌判据。准周期运动就是提醒：不严格重复可以来自多种机制。讨论洛伦兹系统时，应该明确关注其非线性耦合、有界演化与相近轨道的敏感差异，而不是只以图像看起来复杂为据。图像复杂度会随投影、分辨率和绘制时长改变，不能独立承担动力分类。

随机性同样有不同来源。外部噪声、未解析过程或观测误差可以在实际天气模型中表现为随机项；确定混沌则可以在没有外加随机驱动的方程中出现。二者在观测上有时不易区分，但概念上不能互相代替。若从一段有限、带噪声的自然时间序列中判断混沌，需要处理采样、非平稳性与替代模型，远比画出一条不规则曲线困难。

轨道预测与统计描述又是不同任务。即使无法确定很久以后某个时刻的具体状态，仍可能研究轨道在不同区域停留的比例、某个量的长期平均或极值分布。但这些统计性质是否稳定，需要证明或计算检验；不能因为某个吸引子看起来有两个翼，就断言两侧停留概率永远相等，也不能把有限样本均值当作系统唯一不变量。

从天气到气候的推广尤其需要小心。天气预报关心给定初始条件后某段未来的状态，气候问题常涉及在一定边界与强迫条件下的统计分布。初值敏感可以限制前者的远期精确轨道，却不自动取消后者；边界强迫变化也可能改变统计分布。另一方面，统计预测并非天然容易，模型偏差、慢变量和极端尾部都可能造成困难。不能用“天气不可预测、气候可预测”这种对称口号替代具体论证。

对哲学而言，洛伦兹的启发不是宇宙突然丧失因果，而是人的有限信息进入了预测能力的定义。方程中的确定性，是关于给定条件的演变关系；科学实践中的可预报性，还取决于初值误差、模型误差、目标变量和允许误差。把这些条件写进判断，并不降低物理规律的地位，而是使规律如何成为可用知识更准确。

SDE 在此可以揭示一种对象重定：原先想问“任意远的未来究竟在哪里”，后来转而问“在当前信息下，哪些变量、哪些时距能够维持可接受误差”。S 不再只有单条轨道，也可以是条件化预测区间或状态集合；D 不再只有一次积分，而包括扰动实验与误差

比较；E 是非线性耦合及有限信息对预测范围的共同限制。这种重定发生在研究任务与表示方式上，不是认识活动改变了洛伦兹方程本身。

用概率描述初值，不意味着把确定方程改成随机方程。可以让一个初始概率分布中的每个状态都按同一确定模型演变，分布随之被拉伸、折叠和搬运。概率反映我们对起点的有限知识，动力仍然确定。如果再加入随机模型误差，则是另一层假设。区分这两层，能够避免把集合成员之间的差异一概解释为自然本身的随机性。

同样，集合平均未必是一条真实可行轨道。若一半成员在一个状态区域，另一半在另一区域，二者平均可能落在少有成员经过的位置。平均值对某些损失函数有用，却可能不代表最典型的未来。概率预报需要同时报告分布形态、可能的多峰和事件概率，而不是用一条平均线重新伪装成确定未来。这是预测对象从轨道走向分布后新增的表达责任。

长期统计还要面对取样相关。沿一条轨道每隔很短时间取值，样本很多，却不等于相互独立；相邻状态可能高度相关。估计平均值或极端概率时，应考虑有效样本量和等待时间。若某类稀有转换很少出现，短积分无法可靠估计其概率。复杂动力不会自动赠送丰富统计信息，仍需要合适的实验长度与不确定性分析。

本章也不把蝴蝶图形当作因果解释。某个微扰后来与大差异相伴，并不意味着现实中的任意微小行动都能按意愿控制远方天气。敏感依赖说明预测需要精确条件，不说明控制者知道怎样选择微扰以获得指定结果。控制、因果归因与可预报性属于不同任务，不能从“微小差异会放大”跳到“任何小事都决定一切”。

方程、参数、初值与算法使敏感依赖成为可以复查的问题。蝴蝶比喻便于传播，却没有给出扰动方向、增长区间和允许误差；把它直接用于其他对象时，这些缺项仍需补足。例如，某个观测量对微扰是否敏感，可能与全状态轨道是否分离不同。洛伦兹模型带来的思想变化，应落实到指定预测任务能够维持多久，而不能仅由一幅复杂图形扩展成普遍结论。

敏感系统中的可预测性还具有变量选择问题。两个状态在三维欧氏距离上已经明显分开，某个粗粒度量却可能仍然接近；反过来，一个阈值事件可能对小的状态差别十分敏感。因此，必须先说明允许误差是针对全状态、某个分量、区域平均还是事件发生。没有目标变量的“预测极限”，往往只是把不同任务混为一谈。

观测精度与模型精度也不能任意互相替代。若初值已经很精细，而模型结构有系统偏差，继续增加初值位数可能收益有限；若模型在目标尺度上较可靠，而初始场缺少关键方向的信息，改善观测又可能很有价值。一个有效研究计划应比较这两类误差在指定时距上的贡献，而不是把所有不确定性归给混沌。混沌解释不能成为放弃修复普通错误的借口。

概率预报的检验进一步改变成功的定义。若系统在许多类似场合报告某事件概率为七成，那么这些场合中事件发生的频率应与之相称；单次没有发生，并不能独自证明七成概率错了。与此同时，一个永远报长期平均概率的系统可能校准尚可，却缺少区分不同情境的能力。可靠性与分辨能力需要一起判断。这是从轨道到概率后新增的任务，不是降低标准以逃避错误。

这种变化也影响主体的学习。预报者必须学会表达有条件的不确定性，使用者则要根据损失与收益决定行动阈值。科学并未退回含糊，而是把不确定性本身变成可检验对象。若所有概率都无法校准，或信息不比简单基准更有用，新的表达方式仍应接受失败，不能仅凭承认复杂性而获得豁免。

误差阈值还取决于行动。航线规划、农业安排与基础研究允许的偏离不同，同一预测系统可以对一项任务已失去技能，对另一项仍然有用。这里没有必要把价值判断塞进动力方程，却需要把评价标准放在实际任务中说明。确定系统的敏感性给出物理约束，主体与工具复合体则据此选择可承担的问题。二者共同规定可用预测的边界，但不能互相替代：行动需要不能让不可识别的轨道变得可识别，物理限制也不直接决定所有行动选择。

◆ 39.4 预测任务怎样被重新规定

如果初值误差在某一阶段近似以正增长率 λ 放大，可以写一个教学关系：误差约等于初始误差乘以 $\exp(\lambda t)$。当允许误差为 ε，初始误差为 δ 时，达到阈值的时间近似为 $\ln(\varepsilon/\delta)/\lambda$。这个公式依赖增长率近似恒定、误差尚未饱和等条件，但它说明一个重要现象：把初值误差减少十倍，预测时距增加的是一个对数意义上的量，而不是自动延长十倍。

这并不意味着提高观测精度没有价值。更准的初值可以改善短期预测，也可能把关键事件的可用时距向后推进；某些状态的误差增长较慢，改善还会更有收益。问题在于收益

应按具体动力与任务衡量，而不是以“只要更精确就没有极限”为原则。有限可预报性把资源投入问题变得更细：在哪里提高精度、提高哪种变量、为了哪一种决策，往往比抽象地追求所有数值更精细更重要。

集合方法是一种自然的后续方向：用一组合理初值或模型设定表示不确定性，比较它们的演变。集合分散可以提供风险线索，但不能直接等同于真实概率。若初始集合没有覆盖关键误差方向，全部成员一致也可能共同错误；若模型误差未被表达，集合可能过于自信。概率任务需要校准、可靠性检验与足够多的独立案例，而不是把多条线画在一起就结束。

历史上，现代集合预报与概率业务经过后续多项研究和系统建设，不应全归结为1963 年一篇论文的直接发明。洛伦兹的节点在于明确展示某种确定动力与长期轨道预测之间的张力，为后来重定任务提供了强有力的思想资源。思想影响、方法形成与业务实施属于不同时间线，保持区别才能尊重后续工作的创造性。

预测目标也可以转向事件。用户可能并不需要二十天后某一小时每个网格的风，而需要某段时间是否更可能出现某类形势，或某个阈值被超过的风险。目标变粗不保证一定可预测，但可能与系统保留的信息更相称。科学任务的成熟，不只是延长同一目标的时距，也包括选择在给定条件下真正具有可检验意义的问题。

这些变化为 SDE 的分析相位提供了不同层次的材料。动力性质的认识使原有预测对象受到限制，扰动与精度比较改变研究操作，有限时距与概率检验则重新组织评价关系。这些层次可以先后出现，也可能只有其中一部分成立。例如，增加扰动实验而仍使用原有目标和评分方式，首先改变的是操作；只有目标与判据也被重新规定，才能讨论更广的关系重组。发生强度由这些实际变化决定。

本章也留下一个实践出口。初值不可能无限精确，却仍然可以通过观测与模型联合估计得更好。观测不完整、有误差，模型也有偏差，二者怎样在明确误差条件下形成下一次预报的起点？下一章讨论资料同化。它不承诺消灭混沌，而是在有限信息中尽量形成相容、可用并能够持续修订的分析场。

第 40 章 资料同化
观测与模型怎样共同生成分析场

◆ 40.1 背景与观测为何都不是答案

预报需要一个初始大气，但没有任何仪器能够在同一时刻、每一个位置、每一个高度直接测出全部状态。测站提供局部读数，卫星记录辐射，探空沿上升路径采样，飞机沿航路报告。它们的时间、位置和所测物理量不同。上一轮短期预报可以填补空白，却也积累了误差。资料同化的任务，是把这些不完整信息放进共同的误差与动力关系，形成可用于下一轮预报的分析场。

“背景”通常是由先前分析推进而来的短期预报，它不是无知起点，而包含此前观测与模型的信息；“观测”也不是未经解释的真值，它受仪器误差、代表性与处理影响。两者相遇时，首先需要观测算子，把模型状态转换成与仪器读数可比较的量。卫星辐射不能直接与某个格点温度相减，必须经过辐射传输等联系；固定站读数与网格平均之间，也需要考虑尺度差别。

一个单变量教学例子可以把权重说清。设背景温度估计为十度，误差标准差为两度；独立观测为十二度，误差标准差为一度。在无偏、误差不相关且采用简单高斯二次估计的条件下，分析值更靠近精度较高的观测，得到十一点六度。它不是两者简单平均，也不是宣布观测绝对正确。若观测误差更大，权重就应下降；若背景与观测误差相关，这个简单结果便不能照搬。

写成符号，分析为 $x_a = x\_b + K(y - x\_b)$，其中 $K = \sigma\_b^2/(\sigma\_b^2 + \sigma\_o^2)$。$y - x\_b$ 是观测相对背景的差，通常称创新；K 决定把差异吸收多少。这个式子揭示，更新不只由“差多少”决定，还由我们怎样评价两类误差决定。同样的两项数值，若误差条件不同，可以形成不同而各有依据的分析。

在多变量问题中，背景误差协方差还决定信息怎样传播。一处温度观测可能影响邻近位置，某一压力信息也可能通过平衡联系影响风。但传播范围不能凭空设得无限大。相关

尺度、方向与地形等条件若不合适，一项正确观测也可能被错误扩散。分析场的空间连续性不是由插值美感保证，而是由误差结构与动力联系承担。

分析场首先是指定资料、算子与误差设定下的状态估计。它填补了单独观测留下的空白，并以协方差关系限制更新，却也继承这些关系的适用范围。检验需要越出本次拟合：比较后续预报、留出观测和不同类型资料，可以检查更新是否改善了未被直接约束的部分。因而，分析场的价值应由它支持的判断来衡量；全网格都有数值，并不使每处状态都得到同等强度的观测确认。

在前述标量例子中，分析误差方差也可以计算。在相应独立、无偏假设下，它为背景与观测方差乘积除以二者方差之和，得到零点八平方度，小于各自原有方差。这个改善有条件：若两者共享偏差，公式会过于乐观；若观测实际误差不是一度，权重也会错误。数学上的最优只是在所声明的误差模型内最优，不能作为现实正确性的自动保证。

把一条观测复制十次尤其能说明这一点。若程序错误地认为复制品相互独立，它会不断增加观测的信息权重，把方差压得越来越小；实际却没有获得任何新信息。重复不是多样性，数量不是独立性。资料同化中的可靠性，取决于是否把观测之间的共同来源纳入 E，而非只让 D 在数量上增长。

◆ 40.2 4D-Var 把时间怎样带入

如果观测分布在几个小时内，不能全部当成同一时刻的读数。一个移动气旋在早时刻位于西边、晚时刻位于东边，二者并不矛盾。四维变分同化把时间窗内的观测通过模型演变联系起来，寻找一个初始状态，使它推进到各观测时刻后能够较好符合资料，同时不无依据地偏离背景。时间因此不是事后校正标签，而进入了状态估计的内部结构。

ECMWF 于 1997 年 11 月 25 日将 4D-Var 投入业务，官方实施记录保存了这一节点。[R40a] 早期采用六小时时间窗，后来的窗口扩展、物理方案与模型误差处理属于进一步发展。[R40b] 因而，这个日期标记的是特定系统开始运行。比较业务效果时，应把当时资料、窗口长度和近似方式作为整体条件；现代版本的能力需要由后续改进分别说明，不能全部记入初版的成绩。

用强约束形式作教学表示，代价函数可写为：$J(\mathbf{x}_0)=\tfrac{1}{2}(\mathbf{x}_0-\mathbf{x}_b)^{\mathrm{T}}\mathbf{B}^{-1}(\mathbf{x}_0-\mathbf{x}_b)+\tfrac{1}{2}\sum[\mathbf{y}_i-\mathbf{H}_i(\mathbf{M}_i(\mathbf{x}_0))]^{\mathrm{T}}\mathbf{R}_i^{-1}[\mathbf{y}_i-\mathbf{H}_i(\mathbf{M}_i(\mathbf{x}_0))]$。第一项限制初始状态偏离背景的程度，第二项

比较模型推进到各时刻后与观测的差异。B 与 R 分别表达背景和观测误差的协方差，M_i 表示从起点推进到第 i 个观测时刻，H_i 把状态映射到观测空间。这里假设背景误差与观测误差不相关，不同 i 组的观测误差按块不相关，同组内部相关性由 R_i 表达。若跨时刻或跨组相关显著，应把残差堆叠，并使用含非对角块的整体观测误差协方差。

这一式子中的“强约束”，意味着时间窗内的模型演变被当作精确约束来使用。它不宣称现实模型真的无误，而是在求解中作出该近似。若模型偏差显著，系统可能被迫把模型错误吸收到初值调整中：为了在后时刻符合观测，起点被改得不合理。这正是模型误差需要单独处理的原因。弱约束方法允许在一定误差规定下偏离模型轨道，但会增加未知量和估计责任。

求解也不是逐个观测轮流把状态拉近。变分方法寻找整个代价函数的较优状态，后时刻观测可以通过模型敏感性影响初始时刻。伴随方法有助于有效计算目标对大量初始变量的梯度。它在计算上常涉及反向传播敏感性，却不能被理解成真实大气倒着流回过去。物理时间演变与梯度计算的方向属于不同层次。

非线性又使问题更复杂。观测算子和模式演变可能非线性，代价函数未必只有一个简单最低点；实际实现常使用线性化、增量形式与多次迭代。近似的有效范围、内外循环安排和计算成本互相限制。一个理论上表达完整的目标函数，仍需变成能在业务时限内稳定运行的算法。1997 年的节点之所以重要，正在于理论、程序、资料处理与运行条件形成了可持续组合。

四维关系也改变观测价值。一项单独看来信息有限的观测，若位于系统对初值敏感的时空位置，可能帮助约束一段轨道；大量高度相关、位置相近的观测，则未必提供同样多的独立信息。数据数量、覆盖与动力可观测性不是同一个指标。模型在这里不仅填空，也规定某项观测能够影响哪些状态方向。

时间窗长度也有双重作用。窗口较长，可以利用更多不同时刻的观测，并让动力关系帮助传播信息；但模型误差、非线性和线性化困难也可能随之增加。窗口较短，近似较容易维持，却可能无法充分利用时间联系。不存在脱离模型能力与资料分布的最佳窗口。业务系统改变窗口，应同时检查计算成本、分析质量与后续预报，而不能只以吸收资料更多作为理由。

观测算子的误差还会与初值误差竞争。例如卫星辐射偏差可能来自温度剖面，也可能来自云、地表发射率或辐射传输近似。如果系统只允许改温度，就可能用错误的温度调整去解释其他来源的偏差。多变量与多资料约束有助于区分这些候选，但同样可能留下补偿。分析场的形成与地下反演一样，需要认识可识别方向。

◆ 40.3 误差与重复资料的风险

加权更新依赖误差设定，而误差本身需要估计。仪器噪声可以较容易校准，代表性误差、处理误差和模型偏差则常与状态有关。若某类观测在云雨条件下误差更大，却始终使用晴空误差权重，同化就可能在最复杂天气中最自信地吸收错误。质量控制与误差模型因此是状态估计的一部分，不是数据进入算法之前的纯粹清洁工作。

相关误差尤其容易制造虚假精度。假设一百条记录共享同一个偏差，把它们当作一百次独立测量，误差平均会被错误地认为大幅下降。卫星相邻通道、相邻像元或同源处理产品可能具有相关性；背景也已吸收前一循环的观测，不能把每轮都想成完全独立。资料谱系和相关结构决定信息增益，不能只按条数计算信心。

重复利用还有更直接形式。同一原始观测被不同产品包装，若未识别其共同来源，就可能在代价函数中被多次计数。联合资料看似丰富，实际只是在同一证据上加码。解决需要追踪来源、处理版本和误差依赖，或通过适当抽稀、相关误差建模等方法处理。采用哪种方法取决于系统与任务，不能把简单删除重复项当成全部答案。

偏差与随机误差也应分开。一个温度计总是偏高一度，增加样本不能把这种偏差自动平均掉；模型长期在某种状态下偏冷，也不能只用更大的随机方差概括。偏差订正可以改善一致性，却有把真实信号当成偏差移除的风险。若订正依据同一个有偏模型，观测可能被调整到模型期待的世界。独立校准和不同资料交叉检验因而十分重要。

同化残差可以提供诊断，但残差变小不必然意味着分析更真实。增加观测权重会让分析更贴近观测，即使观测有偏；增加模型自由度也会提高拟合。应同时检查背景相对观测的创新、分析相对观测的残差，以及后续预报对未参与更新资料的表现。几种指标共同使用，才能区分真正信息吸收与过度拟合。

用分析场验证预报十分方便，因为它覆盖完整，但分析场与预报可能共享模型和资料。某类错误若在二者中同向存在，验证误差可能偏小。独立观测、不同分析系统及留出

实验能够补充检验，各自也有代表性与覆盖限制。没有绝对无模型的全场真值，并不意味着无法评价；它要求评价公开依赖，采用多种有边界的证据。

这项纪律与第 33 章的联合反演相通，却有新的时间特征。地下反演常把一批资料放进一个结构模型，同化则持续循环：上一轮分析生成背景，背景与新观测生成下一轮分析。循环能够积累信息，也能够积累偏差。若一个错误被每轮重新强化，系统可能形成表面自洽而持续偏离的状态。检测这种失败，需要跳出当轮残差，考察长期与外部表现。

质量控制应当允许真正极端的状态进入。若系统把一切偏离背景较大的观测都当作错误，快速发展的真实天气可能被过滤掉；若所有大创新都被接受，仪器故障又可能污染分析。判断需要结合邻近观测、仪器状态、物理一致性与背景的不确定性。一个固定阈值可以提高效率，却不能在所有情境中代替证据比较。

观测影响试验则可以检查某类资料的实际贡献。保持系统其他部分尽量一致，比较使用与不使用某类观测时的后续预报，可以看到它对任务的增益或损害。结果仍受试验时段和资料依赖影响，不能把一次区域改善推广为普遍效果。但这种受控比较能够把“资料越多越好”转成可检验判断，也为误差权重与筛选规则的修订提供依据。

ECMWF 的 2017 年技术说明给出一个具体难题：当时弱约束处理需限制在一定高层范围，以避免把飞机观测误差错误归因于模式误差。[R40c] 这表明增加一个误差变量并不自动使估计更真实，新的自由度也可能吸收错误来源。区分观测与模式误差，需要新的独立约束；方法越灵活，来源识别的责任越大。

分析场还需要注明时效：实时分析只能使用截止时刻已到达的资料，事后重分析可以吸收更完整记录并使用统一版本。二者对同一历史时刻给出不同状态，并不必然矛盾，而是信息条件不同。评价当时业务能力时，必须使用当时可得的信息，不能以事后优势替代实时判断。

◆ 40.4 同化回路的发生学边界

同化循环包含两类性质不同的变化。一类是在固定算子与误差模型下，用新资料更新状态；另一类是持续失配促使研究者修订算子、误差结构或筛选规则。前者让既有系统继续运行，后者才改变状态怎样被估计的条件。SDE 对回写的分析需要辨认这一差别：同样表现为新分析场，背后可能只是输入改变，也可能涉及观测与模型之间的联系重组。

例如，某类观测在特定条件下持续产生偏差，研究者可能修改观测算子、误差权重或筛选规则。这几种改变不是等价的。修改算子意味着重新规定仪器读数与状态变量的联系；修改权重意味着调整可信程度；改变筛选则改变哪些差异进入估计。若只看到分数提高，而不知道采取了哪一种改动，就无法解释新的分析场究竟怎样形成。

官方技术记录中，强约束与弱约束对模型误差的不同处理，提供了这种变化的一个明确方向。ECMWF 的说明将 1997 年的强约束业务起点与后来对模型误差的扩展区分开来。[R40c] 这不是自然大气突然换了一套规律，而是估计系统承认了原先被忽略的误差来源，并把它纳入状态与轨道之间的关系。具体成效仍应按相应版本与检验报告判断。

三式在这里组织一条可核对的关系链。S 是带版本、资料与误差条件的分析场；D 是观测创新、迭代差异与循环预报检验；E 是观测算子、协方差和动力传播之间的特征纠缠。分析场规定下一轮预报的起点，继而影响哪些观测显得异常；持续异常又可能暴露算子或误差权重的问题，促使 E 修订。区分正常状态更新与这种条件修订，才能说明回路究竟在积累信息，还是在改变其获取信息的方式。

同化还展示一种复合主体能力。预报员、观测网络、误差模型和计算系统共同承担状态判断，任何单一环节都不拥有完整大气。人的责任没有消失，而是转到识别系统失效、比较独立证据与决定修订条件。仪器和模型也不是被动工具，它们通过可测范围与可解结构实际参与判断的形成。这种复合能力既能扩展认识，也可能封装偏差，因此需要可追溯版本与失败记录。

分析场是观测与模型在误差条件下共同形成的可用状态。它为下一次预报提供起点，又由随后到达的资料和实际演变接受检验。这种循环使初值估计持续更新，作用范围仍由观测、动力与误差模型限定。接下来转向流域时，状态估计之外还会出现另一项前提：应当把哪些空间范围划入同一个核算对象？大气中的初值相容与水文中的边界相容，将以不同方式要求观测、过程和研究单元彼此匹配。

第 9 编

流域海洋与跨圈层交换

第 41 章 流域不是一个现成的容器：研究边界的形成

第 42 章 从入渗到状态记忆：降水怎样成为径流

第 43 章 从温盐水团到风驱环流：海洋对象的两种形成

第 44 章 水循环从一幅图到可核算的交换系统

第 45 章 任务改变，水文对象是否必须改变

第 41 章 流域不是一个现成的容器

研究边界的形成

41.1 从出口反推集水区

在一幅地形图上找到河流出口，再沿山脊画出闭合线，这似乎已经完成了流域的定义。线内降水流向同一出口，线外降水流向别处。这个朴素图景有很强的生产力：它允许研究者将分散的雨量记录与一处流量记录联系起来，使降水与出流不再是互不相干的数字。然而，闭合线的意义始终取决于水怎样运动。山脊能够分开地表水，并不因此保证也分开山脊以下所有深度的地下水。

划界首先是一项有方向的操作。地形高程不是只供观看的形状，而被转换成某一点可能向哪些相邻位置排水的约束。汇流方向连接成河网，河网再把出口与上游范围连接起来。研究者移动出口，所问的问题便改变了；把出口移到支流汇入口上游和下游，不只是图中一个点的位置不同，参与水量核算的全部上游来水也不同。一个流域的面积因此必须与出口位置一起报告，脱离出口的面积数字不能完整表达对象身份。

数字高程资料提高了这种操作的可重复性，却没有使判断消失。洼地是测量噪声造成的伪洼地，还是具有真实蓄水能力的地形低点？道路构筑物拦断了地表坡向，但其下是否存在涵洞？低平区几乎相同的高程值应当如何确定排水方向？这些问题决定了哪些连线具有水文意义。若算法把真实封闭洼地全部填平，计算得到的汇水面积可能非常整齐，却悄悄引入了并不存在的排水连接。若没有补入涵洞，又可能制造一条不存在的分水线。

地表划界失配还不同于水量失衡。前者是边界与所要表达的排水关系不符，后者是所用储量与通量不能在规定误差内闭合。即使边界画对了，雨量站漏掉山地降雨、出口流量关系发生变化、蓄水量没有记录，都能造成收支不闭合。反之，多个错误也可能相互抵消，使一张不恰当的边界图获得漂亮的总量平衡。因此，平衡检验是必要约束，却不是边界正确的充分证明。

考虑一个教学情境：同一出口上游存在跨山引水渠。雨量仍只落在地形流域内，出口流量却包含外部输入。若把引水全部误认为流域降雨的结果，径流系数可能显得异常。正确处理可以保留地形流域，另外列明人工输入；也可以为了某项调度任务扩展控制系统，把取水口与渠道一并纳入。两者没有天然的高低之分，但都必须说明流量究竟穿越了哪条边界。对象的可选择性由任务产生，对选择的约束则由真实通量承担。

流域由出口、地形连通和过程假设共同限定。出口回答在哪里核算，连通关系回答哪些来水能够抵达，过程假设则决定哪些跨界交换必须单列。山脊与河道提供了划界依据，水量观测检验这条界线能否承担预定任务。科学流域的形成，就在于把已有的地貌与水流组织成能够区分内部交换和边界交换的研究单元。

出口本身也可以具有不同截面定义。以桥位为标志便于重复定位，但河道分汊、漫滩与旁侧渠道可能使桥下主槽只承担部分出流。洪水时有效过水范围扩大，平水期成立的单断面代表性可能改变。因此，固定坐标并不免除对实际过水结构的检查，测量位置的稳定与控制边界的完整性应分别判断。

◆ 41.2 水文单元不必是完整流域

美国地质调查局的流域边界数据集规范特别提醒：一个水文单元可以代表某出口以上全部汇水区，也可以只代表其中一部分；完整上游范围可能需要组合多个水文单元。[C01] 这条看似技术性的规定，防止了一个重要的概念滑动：编码单元、地图多边形与完整出口流域不能无条件互换。

设一条干流由北向南流经三个区域，上游两个支流区域分别编号，干流下游的局部集水区再编号。第三个单元的出口接收到的水包括本地降水汇流以及前两个单元的来水。倘若只以第三个多边形内的降水解释出口流量，上游输入便消失了；如果又把三个单元各自出口流量相加，下游断面已经包含的上游流量便被重复计算。同一套编码资料可以便利管理，也可以在错误使用时制造守恒矛盾。问题不在编码本身，而在使用者没有把单元的拓扑位置转成核算规则。

层级编码解决的是另一类难题：怎样让不同机构谈论同一片区域，怎样把局部记录汇总到更大的区域，以及怎样在地图修订后追溯资料的所属关系。它让流域从一位调查者手

中的局地图件，成为可以交换的公共资料对象。但是，层级中的父子关系主要表达空间包含，水文模型还需要表达流向、储留、边界穿越与时间滞后。包含关系并不自动提供这些动态信息。

一张水文单元图若用于土地覆盖统计，区域面积与分类一致性可能是主要问题；若用于洪水演算，支流汇合的顺序和洪峰到达时间便不可省略。前一任务可以允许把一些内部过程留作黑箱，后一任务必须打开其中部分连接。把管理单元直接移作过程单元，相当于把一种任务中可以忽略的差异，未经检验就判定为另一任务中也可忽略。许多模型误用因此发生在输入资料看似完全合法的地方。

沿海与湖岸地带尤其能暴露出口想象的局限。陆地排水未必通过一个集中河口进入海洋或湖泊；分散小沟、湿地与地下交换可能让边界具有多个出流位置。此时仍可建立资料单元，但不能用一条河流的单出口图式强迫它满足并不存在的结构。闭合区域在几何上是明确的，并不表示其中的水只有一种离开途径。

水文单元的转换使对象身份的形成过程变得清楚。出口位置规定汇水关系，汇水关系规定可汇总的降雨与地表特征，这些特征又规定哪种流量记录有资格代表该单元。若拿掉出口，面积统计就失去出流归属；若拿掉上游连接，下游流量就失去完整输入。E 所指的特征纠缠由这些相互限定构成，S 是由此显露的可复核流域，D 则包含划界、组合、测量和比较中实际运行的差异。

水文规范已经区分单元与流域。思想史还需考察这一区分在资料转用中怎样被保持或丢失：管理表中的多边形进入过程模型时，原先省略的上游连接是否恢复？地图修订进入长期序列时，旧观测是否重新归属？沿着这些转换追踪，同一名称下对象身份的变化才有可检查的内容。规范规定怎样使用，使用史则显示哪些规定曾被忽略，又因何种失配重新成为问题。

这里还可以区分外边界与内部划分。外边界回答哪些交换穿越研究系统，内部划分回答同一系统中的哪些差异需要单独记录。二者常被同一个“流域划分”名称掩盖。改变内部单元而保留外边界，整体水量预算可能不变，但对子区贡献的认识会改变；移动外边界则直接改变输入与输出的身份。前一种调整不应被误写成整个流域换了一个出口，后一种调整也不能仅用编号更新来处理。

版本映射因此至少需要保存面积变化、河网连接变化和编号变化三项信息。编号改变可能只属于资料整理，面积不变也可能隐藏河道连接修正。若历史序列按旧边界汇总，新资料按新边界汇总，比较前需要决定重算旧资料还是建立明确转换。直接拼接会使边界修改在曲线上表现为水文变化。一个可更新的对象不是永不变动，而是变化之后仍能说明旧记录与新记录的关系。

◆ 41.3 测站与边界互相修改

边界建立后，研究者通常据此安排雨量、河道与地下水观测。测站并不是均匀撒在地图上的点。上游分支是否需要独立测流，取决于研究者要不要区分支流贡献；地下水井需要布在哪里，取决于跨边界交换是否可能影响所研究的时间尺度。对象先给出观测方向，观测随后又可能使对象原先的界限受到质疑。

这种双向关系需要实际记录，而不是一句“相互影响”。一份可检查的资料卡应至少包含出口坐标、边界版本、观测时间窗、流量换算方法以及引排水记录。相同站名不能保证资料始终可比：断面迁移、河床冲淤、测量方法调整，都可能改变水位与流量的关系。如果某年以后收支残差突然增大，首先应核对这些变化，再讨论流域是否出现了新的地下连通。

时间窗口的选择同样重要。一次暴雨之后，流域可能仍在把先前储存的水排出。若只比较降雨期间的输入与出流，许多水尚未到达出口，残差表现为暂时蓄水；若延长到退水结束，结果又不同。年度核算也不能默认年初与年末储量相同，尤其在连续湿润或连续干旱阶段。把储量变化从公式中删去，并不会使储量变化停止发生。

可以把简化水量关系写成“储量增加＝降水输入＋外来输入－蒸散输出－河道出流－其他外排”。这里每一项都对应控制边界和时间窗。若用流域面积归一化，各项可以表示为等效水深；若保留体积单位，面积积分就必须已经包含在降水与蒸散项中。单位一致只解决第一层问题，第二层问题是各项是否覆盖同一片区域与同一时段，第三层才是误差是否能够支持机制判断。

假定残差在汛期持续为正、枯水期持续为负。这个模式可能提示储留时间处理不足，也可能来自降水空间估计、蒸散季节偏差或出口测流误差。它并不唯一指向地下水。研究

者可以提出几种候选解释，再寻找它们预期不同的观测：地下交换假说需要水位梯度、地层连通和水化学等独立约束；雨量偏差假说要求检查地形位置与降水测站代表性；储量遗漏假说则需要土壤、积雪或水库记录。一个残差可以发出问题，却不能替代解答。

新增测站的价值因此取决于区分力。若三种假说都能预测出口流量的相似改变，再增加一台同类出口仪器主要提高精度，却未必区分机制。相反，一条穿越疑似地下分水界的水位观测剖面，可能在精度较低时仍带来更强的判别能力。研究对象不是随着数据数量简单膨胀，而是随着某些关键差异变得可辨认，才需要重新划分过程单元。

上述残差诊断区分了两条修订路线。若水位、地层与水化学共同否定原边界所假定的隔离关系，需要改画控制边界或增设跨界输入；若异常与断面迁移、雨量估计或引水记录同步，首先应修订资料链。两条路线可以用同一出口的边界版本图、测站变更记录和同期引排水资料作比较。它们是辨认边界修订原因的条件，尚不指向某次具体调查的历史经过。

◆ 41.4 边界成立的有限条件

一个可用流域应当附带自己的适用范围。它可以适合某个出口的长期水量评估，却不适合短时洪峰预报；可以用于地表径流统计，却不能直接界定区域含水层。适用范围不是成果后的免责说明，而是对象定义的一部分。缺少这一部分，图层跨任务传播得越快，误用也可能越快。

据此，边界调整有三个层级。地表连通、主要边界通量与观测误差均满足任务要求时，可以保留原边界；局部支流贡献或人工调水需要辨认时，可以细分内部单元，或在既有边界上增加通量接口；当地下交换范围与研究时距使原有控制体持续失衡时，则需重建控制范围。三种处置分别改变内部细节、跨界账目或整体范围，比较修订成效时也应采用相应层级的证据。

比较边界版本时，应把收益和代价同时计算。细分可能减少对支流峰值的平均化，却增加难以观测的参数；扩展可能补入外来水，却使资料短缺更加严重。只报告新图“更完整”，不足以说明新对象“更有效”。需要观察新划界能否改善独立事件的解释，能否让此前混同的过程分开，以及改善是否超过新增自由度带来的拟合优势。

流域由研究操作形成，不意味着水遵从研究者愿望。恰恰因为实际水流可以越过不恰当的边界，边界才有被反驳和修正的可能。可操作的科学世界包含这种抵抗：地图允许我们提问，测量允许自然过程否定某些地图用途。把主体、操作和对象属性共同纳入分析，是为了追究责任所在，而不是解除物理约束。

流域边界由此成为一项能够保存版本、接受测量检验并随用途调整的科学成果。水量平衡说明账目是否闭合，误差分析说明残差是否值得解释，单元规范保证不同使用者谈论的范围可以对应。沿着边界、测站与核算规则的变更，才能进一步辨认对象形成的具体历史。下一章将保留已经限定的边界，把注意力转向边界内部：同样的降水总量，为什么会成为完全不同的径流过程？

第 42 章 从入渗到状态记忆
降水怎样成为径流

42.1 1933 年的入渗问题

雨落到土地之后，并不立即成为河流中的水。它可能停在叶面，进入土壤，充满洼地，沿地表移动，也可能在更晚的时候经地下路径抵达河道。如果只记下雨量与出口流量，这些路径便被折叠在一个比例里。比例有时有效，却很容易把不同的发生过程当成同一种转换。水文学的重要问题之一，就是怎样打开这个比例，识别雨水进入不同路径的条件。

霍顿 1933 年的《入渗在水文循环中的作用》把入渗放在这一转换的中心。[C02] 进入土壤的水由此成为分配后续路径的关键环节：地表供水有多少能够被接纳，剩余部分又以何种方式积留与移动？本章以该论文提出的问题为历史起点，以下采用现代水文学概念分析入渗、储留与产流；这些教学推演不等同于原论文的公式、实验程序或术语原貌。如此分层，既保留 1933 年的问题位置，也能说明后来形成的分析尺度。

理解这一转向，可以从能力与实际通量的区别开始。入渗能力表达在相应条件下地表允许水进入土壤的速率，实际入渗还受到供水约束。地表没有足够水，即使土壤能够接纳更多，实际通量也不会凭空增加。因此，雨强小于入渗能力时，不能把能力直接当成进入土壤的水量；雨强超过相应能力时，超过的部分也要先考虑洼地储留与其他路径，才能讨论出流。

这一差别使研究从“降了多少雨”转向“何时以多大强度降雨，地表当时处于什么状态”。总雨量的量纲是长度，雨强与入渗速率的量纲则是长度除以时间。前者可以比较一次事件的总体规模，后者才能检验特定时刻是否出现供水超过入渗的条件。如果不先区分量纲，研究者容易把一个总量阈值当成过程阈值，解释也就失去明确的作用位置。

土壤表面还不是一个恒定孔洞集合。含水程度、表层结皮、植被根系与可连通孔隙，会影响进入路径；积水压力与水分再分配又可能改变实际入渗。雨水到来之前的土地因此

不能用“同一块土”四个字完全概括。同一地点具有连续身份，却可能处于不同状态。水文学将状态纳入解释，并不是向外部添加一个抽象的“记忆”，而是承认先前过程留下的物理条件仍在参与眼前过程。

这里也须防止另一种简化：入渗能力被超过不是所有地表径流的唯一形成机制。某些地方由于土壤已经接近饱和，或者地下水抬升与地形汇聚而出现地表出水，即使局地降雨强度没有超过一块较干土壤的潜在接纳能力，也可能有径流。不同机制可能同时存在，并随区域、季节与事件阶段改变其份额。把一个机制讲清楚，不能以删除其他机制为代价。

以发生学语言说，研究者面对的 S 不再只是某场雨后的总出流，而是具有前期条件与路径约束的响应事件。D 包括雨强、土壤水分和河流过程的顺序差异。E 的关键不在“土壤、雨、河流都有关”，而在土壤含水改变可用储留空间，储留与入渗改变地表积水，积水和连通路径再改变到达河道的时序。每一项的解释身份由其他项共同规定。

这种拆分也改变了调查安排。总量比例只要求比较雨量与出口出流；要区分入渗受限、储留填满与路径连通，就需要同时记录雨强、前期含水和产流位置。变量的增加有明确分工：它们分别约束输入速率、可用容量和空间连接。思想史所要追踪的，正是这一问题转换怎样使原先未被记录的状态进入观测，而非把所有新增变量当作同一种资料积累。

还可以把地表产流与河道出口出流进一步分开。地表某处出现积水，不意味着水立即抵达测站；它可能先填满小洼地，随后连接到浅沟，再进入河道。连通性具有阈值：孤立积水区尚未连成路径时，局地有水而出口响应有限；降雨继续后，原来分离的区域连通，出口可能迅速增加。于是，响应的非线性不必全由某个土壤参数突然变化造成，也可能来自空间连接的形成。

这使“贡献面积”成为随事件阶段变化的对象。地形流域面积可以保持稳定，而真正向河流提供快速径流的区域并不固定。将二者混同，会使研究者误以为整片土地始终以同样比例响应降雨。更细的解释需要辨认坡脚、近河区与高处土壤在不同状态下的作用，但是否采用动态贡献区，仍要由观测和任务决定。

入渗、贡献面积与汇流时延共同打开了雨量—出流比例。比较两场雨时，首先需要保持总量相同，再分别改变输入速度、前期状态或空间连接，才能看清究竟是哪一项使响应分化。

42.2 相同雨量的不同过程

以下用教学情境隔离总量与速率。甲、乙两场雨都提供三十毫米降水，甲均匀分布在六小时内，乙集中在半小时内，相应平均雨强分别为每小时五毫米与六十毫米。若土壤入渗和地表储留能够接纳甲的输入，却无法及时接纳乙的输入，两场雨便可能产生不同的积水与径流。三十毫米仅用于保持事件总量一致，不代表某地预警阈值；能否形成快速出流，还取决于积水是否接入河网。

现在再改变前期条件。一次较小降雨落在已经湿润的土壤上，可能比一次较大降雨落在干燥土壤上产生更大的即时出流份额。然而这一判断也不是处处适用的经验律。干燥土壤的表层状况、疏水性、裂隙与结皮可能使实际响应偏离简单想象。正确的问题不是“湿土一定比干土产流多吗”，而是哪些可测状态在什么地表条件下改变了输入到输出的分配。

洪峰还取决于汇流同步性。若两个支流分别产生同样的出流总量，但峰值到达干流的时刻相差很大，下游峰值可能被摊开；如果到达时刻重合，峰值则可能增加。雨量空间分布与河网传播时间因此共同参与洪峰形成。只比较全流域平均雨量，会把“下在哪儿”这一差异消去。对年度水量而言可以容忍的平均化，对短时洪峰而言可能恰好删掉最有决定力的条件。

这给事件比较提出了严格要求。首先，雨量必须与所研究的汇水面积相匹配。其次，降雨开始、结束与出流观察窗口需要明确。再次，前期状态不能只由一个方便的统计指标代表而不检验其适用性。最后，要说明出口流量已经经过何种水位—流量转换，洪峰期间是否超出了实测率定范围。否则，“同量不同流”可能部分源于资料误差，而不是新机制。

一个有解释力的事件表应同时保留降雨过程、前期状态与输出过程。降雨过程记录强度、持续时间和空间覆盖；状态记录土壤含水、积雪、水库或其他相关储量；输出过程则记录峰值、到达时间、总量与退水形状。不能把这些指标全部揉成一个总分，因为不同机制在不同指标上留下的痕迹不同。一个模型可能正确预测总量，却错误预测峰值；另一个模型可能抓住峰值，却误判退水所反映的储留时间。

若要判断加入状态变量是否有用，应先提出它应改善哪一部分响应。把前期湿度加入模型，可能改善径流总量；引入河网传播可能改善峰值时序。若声称前期湿度同时解释全部变化，就需要更强证据。模型里增加变量总会增加调整空间，只有在没有参与调参的完整事件上保持改善，才较有理由认为抓住了可迁移关系。

同样重要的是“没有改善”的结果。如果土壤湿度观测只代表一小片地表，其信号可能与真正参与洪水形成的区域不一致。此时无改善并不自动证明土壤状态无关，而可能说明测量对象与作用对象失配。反过来，统计改善也可能只是湿度变量代理了未明确记录的季节性。需要通过季节、土地覆盖和降雨类型的分组检查，把机制意义与预测便利分开。

同量事件的推演得到的是一种可区分关系：总雨量保持不变时，改变输入历时或传播时延，峰值仍可变化。真实流域中的主导原因则要靠完整过程线判断。如果短历时输入增强而峰值时序基本保持，产流分配值得先查；如果总出流变化较小而峰形显著展宽，汇流时延就是重要候选。两者也可能同时改变，因而需要把总量、峰值和时序作为相互约束的结果。

用卷积思想还能说明汇流怎样改变洪峰。在一个线性近似的教学模型中，出口响应可表示为此前各时刻有效输入，经由不同传播时延加权后的叠加。权重函数若较窄，输入集中到达，峰值较尖；若较宽，水被分散到更长时间，峰值较低而持续较久。即使输入总量完全相同，传播时间分布也能改变过程线。这个推演隔离了汇流作用，并不假定真实流域必然线性。

如果权重函数随前期状态变化，问题又进一步复杂。较湿状态下部分路径可能更快连接，原本适用的传播关系便改变了。此时仅调整一个产流比例，可能正确改变总量却仍误判峰值。把产流与汇流的误差分开，能够避免一个参数代替两个不同机制。

退水段提供另一种信息。降雨停止后仍有出流，说明储存的水继续释放。在无新输入的线性储库教学近似中，设 $Q=kW$，W 为可排储量，Q 为体积流出率，k 的单位为时间的倒数；由 $dW/dt=-kW$，得到 W 与 Q 按指数衰减。释放时间尺度为 $1/k$。退水斜率因此能够约束有效释放时间；不过，多储库叠加或随状态变化的 k 也会改变退水形状，仅凭局部拟合尚不能唯一辨认内部储库数目。

降雨与流量的比较还涉及面积加权。三个雨量站的算术平均，不一定代表流域平均降水，因为它们覆盖的地形与面积不同。若高处降水显著而站点集中于谷地，平均值可能系统偏低。对同量事件的筛选应保留面雨量估计方法，否则所谓相同总雨量可能只是站点平均相同。

蒸发损失在不同事件时距中的地位也不同。对很短的强降雨过程，它可能不是峰值形成的主导项；对长时段湿润与干燥演变，它却影响前期状态。一个过程在事件方程中可以近似忽略，并不意味着它在准备该事件的状态形成中无关。区分直接作用与通过前期状态进入的作用，有助于解释为什么同一变量在不同模型中时而显眼、时而隐去。

事件结束也不是纯粹日历判断。若前一场雨的退水尚未结束，后一场雨已经开始，两场事件的水在出口叠加。把它们人为切开后分别计算径流系数，可能把储量延续误当作某场雨的异常效率。完整过程分析应保留事件边界的物理理由，而不是只按固定间隔划分。

◆ 42.3 历史记忆怎样被误读

水文学中的记忆至少有三种含义。第一种是物理状态的延续：先前降雨使土壤、水库或地下含水系统保留了水。第二种是统计表达：模型用前几个时刻的变量预测后一个时刻。第三种是资料处理造成的伪记忆：训练和检验之间共享了本应隔离的信息。这三者可以同时出现，却绝不能混为一谈。

物理记忆具有可追问的载体。问“流域记得上一场雨”，实际应问什么储库改变了、改变维持多久、通过哪条路径影响本次出流。储库的时间尺度不同，记忆也就不同。浅层土壤可能对近期降雨敏感，深部地下系统可能保持更长的响应；河道中的传播与水库调度又产生其他时间尺度。一个滞后项能够在数学上描述迟到，不意味着已经识别迟到的物理位置。

统计记忆有其正当价值。即使不能逐项测量内部储量，出口流量的历史序列也可能含有可预测信息。但历史系数必须在研究范围内解释。某个模型使用过去三天流量，并不证明自然系统的记忆恰好为三天；三天可能只是资料分辨率与模型结构的折中。若换一个时间步长，系数随之改变，也不能直接说自然机制发生变化。参数身份依赖模型形式，需要与可观测状态建立桥梁。

时间泄漏则属于证据组织错误。比如把同一场洪水的相邻时刻随机分配到训练和检验集合，模型已经从训练片段看过该事件的大部分结构，检验结果却被解释成对一场全新洪水的预测。这种方法可以回答局部插值问题，却不能直接回答新事件泛化问题。若研究目标是未见事件，就应按完整事件或完整时段划分，并保证标准化、参数选择和缺测填补不提前使用检验期信息。

更隐蔽的泄漏发生在特征构建中。所谓“前期降雨指标”如果包含预测时刻之后的雨量，名称再合理也失去预报资格。某个经过居中平滑的流量序列，可能同时利用过去与未来时刻；在事后重建中这可以有用，在实时预警中却不可直接使用。需要分别说明信息何时真实可得、何时经过处理进入数据库，以及模型在何时作出判断。

验证设计因此应先于结果比较。如果研究者在看过一组结果后不断改变事件边界、剔除困难样本或选择最有利时段，最后漂亮的性能不能承担原本的证据责任。预先写清分组原则并保留失败事件，既是统计纪律，也是思想史研究的重要提醒：被保留下来的成功模型往往有清晰故事，被排除的失败条件却可能揭示其真正范围。

误差之间还会相关。一个流域连续数天的流量误差可能来自同一场雨量低估，不能把每小时误差都当作相互独立的失败或成功。对有效样本量的判断应考虑事件关联。否则，几场长期过程会被拆成数千个貌似独立的证据，置信程度随记录频率虚增。数据更密不一定意味着独立信息更多，这是水文学与古气候研究共同面对的问题。

从 SDE 的差异序列 D 来看，物理演变、仪器采样、数据库处理和模型读取各有自己的次序。它们之间需要明确映射：一小时平均值归属于哪个时段，缺测补值用了哪些邻近资料，预测时刻以前究竟有哪些信息可用。E 则指储量与通量维持响应的具体约束，例如同一含水状态通过排水改变退水，又被新降雨改写。将这两类关系分清，才不会把处理程序引入的连续性当作流域自身的记忆。

状态重建还存在“等效状态”问题。两个流域时刻可能具有相同出口流量，却有不同土壤与地下储水分布。它们在没有新降雨时的短期退水相似，遇到下一场雨却可能分化。出口流量因此可以携带状态信息，却未必完整标识状态。模型若只依赖一个滞后流量，容易把这些不同内部配置压成同一输入条件。

这种压缩是否可接受，取决于预测时距与事件类型。若短时预测已经足够准确，使用有效状态可以是合理封装；若某类新事件反复失败，就应检查压缩是否丢失关键差异。研究进步不是永远拒绝黑箱，而是知道何时黑箱的已验证范围被越过。打开内部结构的理由应来自具体失败模式，而非对复杂机制的抽象偏好。

前期降雨指数也是一种状态近似。它通常让较近降雨获得较大权重，较远降雨逐渐衰减。这样的指数能够概括湿润程度，却不直接测量土壤水分；相同指数在蒸散、土层深度或排水条件不同的地方，可能对应不同真实状态。把它当作方便变量可以，把它当作普适物理量则需要额外论证。

◆ 42.4 何时需要重建对象

模型失误之后，最节制的解释通常是先检查测量、参数与已有过程。若某段异常可以由雨量空间代表性不足解释，就没有必要立即宣称流域机制改变。若补入一个物理上已有的前期湿度状态即可改善独立检验，首先应肯定成熟水文学的解释能力，而不是把改善归功于新的哲学框架。

但也有情况需要重建对象。例如同一出口的响应在土地开发、排水工程或水库运行改变后呈现系统差异。要把差异归因于土地利用，至少需要变化的空间与时间记录，并排除同时发生的气候、测站与管理变化。仅凭某时期径流增加，不能判定不透水面积是唯一原因。对象重建要求新的连通路径确实进入了作用链，而不只是参数数值变大。

重建也未必意味着推翻整个流域模型。研究者可以把原先合并的地表分成不同响应区，可以把人工调度从自然出流中显式分离，也可以为变化前后设置不同结构。应根据证据决定改动层级。把所有失败都解释为“模型不够复杂”，会使任何反证被更多自由度吸收；把所有变化都塞进一个固定参数，也可能使模型失去物理意义。

一个可检验的重建应留下旧、新模型的共同比较面。它们应使用相同的训练信息和相同的独立事件，报告总量、峰值、时序等不同指标，而不只挑选新模型最有利的一项。新增结构还应给出独特预期，例如某类短历时降雨响应改变明显而另一类事件改变较小。若新模型只提高总体拟合，却不能产生区别性预期，机制解释仍然有限。

洪峰异常与响应区重划之间可以建立一条待检验的研究链：某类事件反复失配，促使研究者区分原先合并的区域；分区提出新的雨量或土壤观测需求；新观测再决定哪些区域可以合并、哪些必须区分。这里的关键是后一步能否修正前一步。如果新增测站并未改变分区判断，便只能确认原有结构；若要把这条研究链写成历史经过，还需模型版本、调查安排与事件记录相互印证。

水文对象的成熟，表现为能够把失败分层处理。有的失败要求纠正数据，有的要求调整参数，有的要求增补状态，有的才要求重建连通结构。这样的区分让理论既能够修正，又不至于每遇困难就失去身份。下一章将把状态与通量的区别带入海洋：温度和盐度可以帮助辨认水团，但要说明水怎样流动，还需要另一类约束。

第 43 章 从温盐水团到风驱环流
海洋对象的两种形成

43.1 航线怎样切开海洋

调查船驶过海面，留下的不是海洋本身，而是一串带有位置、深度与时刻的记录。相邻两站之间可能隔着未被采样的水域，相邻两个深度之间也有仪器没有直接读出的结构。海洋学的任务因此不只是把测量数字画成漂亮断面，还要说明哪些连线具有物理依据，哪些只是在稀疏资料之间作出的估计。

温度与盐度能够构成强有力的识别坐标。某一深度范围内的温盐组合，与另一区域相近时，研究者可以提出水体具有共同形成条件或经历混合的假说。但温盐图中的一个点不是海水的身份证号码。它不直接记录这部分水从哪里来、何时抵达、以多大速度移动。分类所要求的属性相似，与输运所要求的方向和速率证据，承担不同责任。

一条航线还会把空间与时间缠在一起。船舶先测西端后测东端，断面上两处差异既可能来自空间结构，也可能包含期间变化。若研究对象足够缓慢，这种近似可以合理；若局地变化快，断面就可能把不同时刻的状态拼成一个实际上没有同时存在的海洋。采样时刻不能作为文件中无关紧要的一列，而是判断断面意义的条件。

由此可见，绘图中的连续性并不自动保证对象的连续性。等温线或等盐线连接的是在一定插值规则下具有相近数值的位置，水团边界则需要结合成因、混合与空间结构判断，流线更需要运动信息。把三者画在同一张图上可以促进理解，却也容易使线条之间的视觉相似掩盖物理差异。一条等温线不是一条水粒子的轨迹。

海洋测量把船位、探测深度、仪器响应与海水属性连接为一个观察系统。某个水团能否显露，既取决于温盐差异，也取决于站距和深度间隔能否分辨它。若薄层结构落在两个采样深度之间，扩大航区并不能补回垂向信息；若仪器响应滞后，快速通过锋面还可能平滑真实梯度。SDE 在此关注的，是可测差异如何取得水团身份，以及这种身份需要怎样的重复观测条件。

这里还有一个重要的不对称：一份好断面可以证明某些属性结构，却未必足以证明动力解释；一个好动力模型也可能在观测稀疏处产生多个可行解。将两者相接，不是让其中一方取代另一方，而是让属性分布与运动约束相互限制。斯韦尔德鲁普 1947 年的论文正提供了一个观察这种连接的历史入口。[C03]

测量方向也改变可见结构。沿岸航线容易显示沿岸变化，却可能错过离岸梯度；垂直于锋面的断面能够辨认两侧差异，平行航线则可能长期停留在一侧。站位选择不是在一个已经完整呈现的海洋上抽样，而参与决定哪种空间结构能被发现。发现某条锋面之后再布置横断面，正说明对象显露能够重新组织观测方向。

另一方面，不能把调查的空白处理解为自然没有结构。空白只是观测系统没有分辨该处。绘图时以平滑面填补空白是必要估计，但应与实测点区分。历史海洋图的演变，因此既反映海洋解释的变化，也反映船舶能力、航线覆盖与仪器深度范围的变化。新的科学世界有时来自一个新方程，有时来自终于能够把此前的空白变成可比较记录。

因此，航线既提供地理断面，也为属性比较选定了样本范围。把这些样品移入温盐图，能够跨越地理距离辨认相近水体，但航线没有测到的深度和区域仍然留在解释边界之外。

◆ 43.2 水团的身份与混合

把温度放在横轴、盐度放在纵轴，海水样品便从地图上的位置变成属性空间中的点群。这种转换有解释力量：两个相距很远的样品，可以在属性空间中彼此接近；地理上相邻的样品，也可能属于显著不同的温盐组合。研究者由此能够在复杂海洋中寻找具有共同特征的水体，而不是只按经纬度把海水切成小格子。

但共同特征可能有不同来源。两处海水可能由相同源区输运而来，也可能在不同地区受到相似冷却或蒸发作用，还可能由不同端元混合后到达相近的温盐值。属性相似给出候选关系，不给出唯一历史。在这里，历史海洋学与矿物年龄研究面对同一种逻辑问题：今天相似的读数，可能来自不同过程；要重建来源，就必须增加能区分过程的约束。

以理想混合为教学例子，若两个端元的盐度分别已知，且混合过程中没有额外盐源汇，那么混合物的盐度可按质量份额计算。温度关系也可在适当的热量近似下讨论。这使

温盐图中的混合线成为有用线索。然而实际海水可能与外界交换热量，端元也可能随时间变化。若把所有偏离混合线的样品都归为测量错误，就会错过真正改变水体身份的过程。

密度进一步把分类连接到动力。温度与盐度都影响密度，但影响方向与大小受状态条件约束。因此，相同密度并不保证相同温盐组合，相同温度也不保证相同浮力性质。水团分类不能只以单一属性为标志，更不能把“冷”直接等同于“下沉”。决定浮力与稳定性的，是相对于周围水体的状态及其垂直排列。

当研究者从温盐变化估计密度结构，再讨论压力梯度与流速关系时，一条新的推理链出现了。温盐不再只用来识别水团，还被用来约束运动。但是，某些方法只能获得相对流速或垂直切变，仍需要参照条件决定绝对流动。假设某一深度静止可以方便计算，却必须说明其理由；若参照层实际上有流，整个输运估计会受到影响。

这表明分类与动力不是彼此排斥的两种海洋。分类提供可追踪的属性关系，动力说明这些关系在何种运动中能够维持或改变。水团可以混合，分类边界可以变模糊，流动也可以不断把新水送入同一属性范围。一个名称所代表的水团，不必对应永远由同一批水分子构成的实体。它可以是持续生成、持续交换而仍可辨认的特征组合。

水团身份能够在持续交换中保持，是因为某些温盐区分仍可重复辨认。混合与输运中的差异构成 D，温盐对密度的共同约束，以及密度、压力与运动之间的联系构成 E。若只保留温度，盐度补偿可能使浮力判断改变；若只有密度而没有参照速度，运动结构仍不能唯一转成绝对输运。水团分类与动力推断相接时，正是这些欠缺决定下一种观测应当补在哪里。

水团标签还提供继续追踪的方向。若某一温盐范围被用作来源线索，研究者可以检查它在不同深度和区域是否延续，并以其他示踪约束区分输运与局地形成。这样的分类既组织搜索，也接受搜索结果修订：一旦相似温盐对应不同示踪组成，原来的单一来源解释就需要拆分。这是由属性标签走向来源假说的研究逻辑，具体航次如何沿此推进仍须由调查记录确定。

温盐混合还显示非线性状态方程的意义。两个端元按照质量比例混合后，温度和盐度可以在一定近似下按比例组合，但密度不必等于端元密度的同样平均，因为密度对温盐的

关系并非处处线性。因而，两股密度接近的水混合后，也可能改变相对于周围水体的浮力。这个一般推理提示，不能把属性空间中的直线直接当成所有动力属性的线性变化。

水团的体积也不是静态分类表中的固定常数。某类水可以在形成区不断产生，在输运中逐渐混合，最终不再满足原分类范围。若要讨论它的收支，就需要定义进入与离开该属性范围的过程。这样，水团可以从“地图上的一块水”转为“属性空间中的生成与消耗对象”。这种转换使分类与通量再次相遇，也提醒读者分类边界同样需要物理解释。

示踪物可以为温盐相似增加另一类约束。若某种物质的源区、反应或衰变行为已知，其空间分布可能帮助判断水体路径与交换时间。但示踪年龄不是每个水分子的共同旅行时间，混合会使一个样品包含不同历史。解释时应把来源分布与混合过程纳入，而不是给水团贴上一个唯一出生日期。

这与水团作为持续生成对象的身份一致。海洋中某类水的“年龄”可以指自最后接触大气以来的时间、模型中的通风时间或某种示踪推断量，彼此未必相同。名称相似时尤其需要说明定义。技术精确性并非妨碍思想表达，而是使“海洋具有历史”这句话获得真正可检验内容。

对海洋思想史而言，新的测量往往不是简单补充旧地图，而是让原来一个名词分裂为几个可区分对象。水团位置、来源、混合比例与输运速率逐渐分工，正使海洋从静态地理空间成为具有多重时间和运动结构的研究世界。

◆ 43.3 1947 年的动力转折

斯韦尔德鲁普 1947 年论文的题目同时标明风驱、斜压海洋和东太平洋赤道海流的应用范围。[C03] 这不是一份无条件描述所有海洋运动的宣言，而是试图把风场与海洋内部输运放进可计算关系的研究。它的思想史意义，在于海流不再仅由航测中“哪里流得快、哪里水更冷”的分布来组织，还可以由动力平衡给出尚未直接测出的输运约束。

本节以现代符号概括这一动力联系，作为教学推演。在大尺度、适当稳态及相应边界条件下，风驱的深度积分经向输运可由斯韦尔德鲁普平衡约束：$\beta V = (\partial \tau_y/\partial x - \partial \tau_x/\partial y)/\rho$。这里 x 向东、y 向北，V 为经向速度沿深度的积分，$\beta = df/dy$，f 为科氏参数，

τ_x、τ_y 为表面风应力分量，ρ 为参考密度。右端是风应力旋度的垂直分量。关系把风的水平空间差异与水柱输运相连，尚未给出海盆环流的全部闭合方式。

风应力旋度与风速大小不是同一个量。两片海区即使平均风速相近，若风应力随空间变化的方式不同，所施加的旋转作用也可以不同。海洋响应因此不能简单说成“风向哪儿吹，全部海水就向哪儿走”。地球自转、压力结构、垂向传递与质量连续性共同限制流动，使局地风与大尺度输运之间存在需要推导的关系。

β 项尤其说明地球几何不是无关背景。科氏参数随纬度变化，使流体在南北移动时面对不同的行星涡度条件。在合适近似中，风施加的涡度作用可以与经向输运所伴随的变化相平衡。因此，环流的方向与强度不是从温盐图外观直接读出，而是由位置相关的动力约束限定。海洋对象在这里增加了新的可检验属性：给定风应力结构后，某种输运分布是否可能成立。

这个关系中的 V 是单位横向宽度的体积输运，单位为平方米每秒；再沿断面横向积分，才得到立方米每秒的总体积流率。β 的单位为每米每秒，因而 βV 的单位为米每平方秒，与风应力旋度除以密度一致。量纲不仅防止代数错误，也划清三类对象：表面局地速度、水柱积分输运与整个断面输运。把漂流浮标速度直接与 V 比较，或把 V 当作全海盆总量，都省略了必要的积分范围。

近似的边界必须和关系一起保存。强烈非稳态、显著非线性、复杂底形、强摩擦以及边界流等情况，都可能要求保留被简化掉的项。它们不是数学上的小注脚，而决定在哪里可以期待关系有效。若模型在某处失效，应检查那里是否仍满足原先的尺度与平衡条件，再判断失败是资料问题还是理论范围问题。

特别要避免以内部海洋的平衡解释全部闭合环流。一个海盆中的水不能无止境向某方向输送而不返回；返回路径和边界动力需要进一步说明。把后来的边界流理论成果全部记在 1947 年论文名下，会抹去海洋学持续解决剩余困难的历史。思想史的节点之所以重要，常常正因为它重新分配了问题：某些问题得到有效约束，另一些此前不清晰的问题因此暴露出来。

这里的“斜压”还提醒我们，密度结构不能从动力中剔除。不同深度的压力梯度与流动关系涉及垂向结构，温盐调查于是获得新的任务。它不仅告诉研究者海水怎样分类，也

帮助检验动力模型所需要的状态。分类与动力之间的桥梁，并不是把两张图叠在一起，而是让某组观测在方程中承担明确角色。

这一转折使研究者能够提出方向更清楚的反证。如果根据风场估计的输运与独立断面约束明显不符，可以追查风应力资料、参照速度、非稳态或忽略项。不同错误产生不同检查路线。相比“海流很复杂”，这种分解把复杂性变成可操作问题。这正是新科学对象生成的标志之一：它不仅让一些现象变得可说，也使某些说法更容易被判为不成立。

地转近似可以帮助看清温盐与运动之间的桥梁。在离赤道足够远、适当大尺度条件下，水平压力梯度与科氏作用可以近似平衡。压力随深度的变化又受密度影响，因此水平密度梯度能够约束地转流的垂直切变。这里的关系不是“密度高的地方一定流向密度低的地方”，因为旋转使流动方向与压力梯度的关系不同于无旋转直觉。

这一推演同时解释为什么参照速度重要。知道各层之间的速度差，并不知道整根水柱是否共同平移。若给所有深度加上同一个合适的速度分量，切变关系可能不变，而积分输运却改变。于是，温盐断面可以强烈约束运动结构，同时留下绝对输运的不确定。这种不完全决定不是理论失败，而是方程提供的信息类型所致。

赤道附近又提醒我们不能机械套用近似。科氏参数接近零时，某些常用地转表达会变得不适合，必须重新检查尺度与保留项。1947 年论文题目包含赤道海流应用，正使读者更应注意原研究的具体处理，而不能把一条中纬度教学公式无条件覆盖全部研究区。不同纬度并非只换了坐标数值，有时改变了主导平衡的形式。

风驱输运关系中的边界条件，也具有同等地位。微分关系说明局地变化怎样受约束，求完整输运场还需要积分常数和边界信息。若不知道哪里不允许水穿越，或怎样处理沿岸输运，仅有局地方程就可能存在多个解。海岸并不是模型外部的装饰线，它参与决定环流能够怎样闭合。

这使海洋动力学的历史很难被写成单一定律一次解释全部现象。新的平衡关系通常打开一片可计算区域，同时把边界层、赤道区或非稳态问题推到前台。科学进步并不总是剩余问题数量减少，有时是问题被重新分类，原先笼统的不理解变成几类具有不同技术路线的困难。

43.4 新预测怎样改动航测

动力模型产生预测之后，观测不应只是寻找一幅相似图形。真正有力的比较，需要明确哪些量由模型输入，哪些由模型推出，哪些来自独立测量。如果风场、温盐场与流速产品共享同一套同化资料，把它们全部当成独立支持，证据强度就会被高估。输入与检验的区分必须沿着资料来源追到底。

一次合适的调查可以针对竞争解释设计。若两个模型在海盆内部给出相似输运，却在边界附近或某个深度范围给出不同预期，那么有限船时应优先用于这些差异最大的地方。若模型争议主要来自参照流速，增加同类温盐站可能仍无法解决问题，需要直接流速或其他独立约束。调查价值来自对疑问的判别能力，而不只来自站位数量。

调查也可能失败。海况使关键站位缺失，时间变化使断面无法代表同一状态，仪器误差使两个预期落在同一不确定范围内，这些结果都不应被叙述为已经验证理论。可以保留未决结论，并说明下一步最有区分力的资料是什么。思想史若只追踪成功航次，会把真实研究中的犹疑、折中与测量限制删除，理论的形成便显得比实际更直接。

当新预测确实改变航线，D 的组织便发生了变化：先前均匀覆盖的调查，可能转向特定边界、深度或季节。若新观测再显示原模型遗漏关键过程，E 也随之修订。可是，只有模型版本、调查计划与测量结果之间有可核时间链，才能认定发生了这种回写。现有理论能够合理解释某航线，不等于它就是该航线当年的设计原因。

从航测到模型检验，海洋对象经历了几次责任转换：温盐差异先承担分类，密度结构再约束切变，参照速度与风应力则把问题推向积分输运。新的调查若针对其中尚未确定的一环，所增加的就不只是另一幅断面图，而是排除某类运动解释的能力。发生学分析应保留这种转换的层级，使一个属性上相似的水团、一个来源假说和一个动力输运解不会被同一名称混为一谈。

一个可操作的模型检验还应区分速度误差与输运误差。某深度的局地流速偏差很大，若影响厚度很小，对总输运未必占主导；相反，一个幅度不大但贯穿深水层的偏差，可能显著改变积分结果。观测应与模型输出处在同一积分与平均层次，否则看似矛盾的结果可能只是在比较不同对象。

对研究共同体而言，统一这种比较层次本身是一项成果。船测断面、漂流轨迹与固定位置时间序列各有优势，不能用其中一类的标准贬低另一类。它们的和谐来自承担不同证据责任：一类辨认空间结构，一类追踪移动，一类观察时间变化。共同海洋对象正是在这些不同操作能够彼此转换和约束时成立。

海洋因此既不是静态属性图，也不是无属性的流动。可研究的海洋在温盐、密度、压力、风与输运的共同约束中取得身份。下一章要把这个结论带到更大的交换系统：河流抵达海洋、海水蒸发进入大气、大气降水回到陆地，怎样从一幅熟悉的循环图变成可核算的研究对象？

第 44 章 水循环从一幅图到可核算的交换系统

◆ 44.1 循环图的解释能力

水从海面升入天空，形成云，降落到陆地，再沿河流返回海洋。这样的图很容易记忆，也确实表达了一种重要联系：陆地上的水不只属于陆地，海洋与大气不是隔绝的领域。但画出一个圆环，并不意味着已经解释水怎样循环，更不意味着知道每一条箭头的数量、速度与变化原因。

首先需要区分储量与通量。湖泊里现存多少水，是储量；某段时间有多少水流入湖泊，是通量的时间积分。巨大的储库可以只具有缓慢交换，较小的储库也可以具有快速周转。如果把图中箭头画得一样粗，把储库画成面积相近的方框，读者很容易从视觉上得到实际并不存在的数量关系。图示中的均衡，需要有单位和时间尺度才成为科学上的均衡。

循环图还常把不同时间尺度压在同一平面。河道水可能在相对短时间内离开一个地区，深部地下水或冰体中的水则可能停留更久。水最终能够回到大海，并不表示一次暴雨的水会在研究窗口内完成这个循环。研究者必须先问：要解释的是一次事件、一个季节，还是更长期的储量变化？不同问题所需保留的储库不同。

霍顿的入渗问题已经说明，一条降水箭头进入陆地后会分成不同路径；斯韦尔德鲁普的研究则提醒，海洋中的运动也不是一只没有内部结构的水箱。[C02][C03] 将两者联系起来，可以看见水循环图的双重身份：它是提出跨区域问题的起点，却不是免除内部过程分析的终点。

一个箭头若要承担研究责任，至少要标明输送什么、穿过哪里、以什么单位计量，以及在什么时间窗内求和。蒸发可以用单位面积单位时间的质量通量表示，河流出海通常涉及穿过断面的体积流率，降水资料又常以水深表示。它们可以转换，却不能不经面积、密度和时间处理就相加。统一图式必须伴随统一量纲。

此外，箭头存在不等于箭头具有主导因果地位。陆面蒸发与降水可能相关，但某地降水的变化也受到外部水汽输送控制。若只根据同一季节中两者一起增大，就宣布当地蒸发

控制当地降水，便忽略了共同驱动与远距离输运。循环图告诉我们哪里可能相连，机制研究则要判断这些连接在何种条件下实际发挥多大作用。

把循环图转成核算系统，需要先明确相态、储留、输运与能量之间的特征纠缠 E，再比较不同空间和时间窗口中的差异序列 D，才得到具有适用范围的系统 S。沿这条分析路线，箭头的一端发生改变，就必须检查另一端是否仍可使用旧定义。例如把小时蒸发改为年度总量，既要累积通量，也要重新考虑储量变化；整体对象随核算尺度取得具体身份。

储留时间提供另一把理解循环的尺子。若某储库体积为 W，平均体积流出率为 Q，则 W/Q 给出一个周转时间量级；例如 W 以立方米计、Q 以立方米每年计，所得时间以年计。这个比值有助于比较储库交换快慢，却不意味着每一滴水都停留同样久。真实系统可能同时有快速通道与缓慢区域，平均周转时间相同的两个储库，其停留时间分布也可以不同。

因此，示踪物到达时间与水量周转时间不应混同。某种示踪物可能优先沿快速路径到达出口，而大部分水仍停在慢区；也可能因吸附或反应比水移动更慢。研究者利用示踪理解水循环时，需要把示踪物自身的行为加入解释。观测工具的可用性来自它与目标过程之间的明确关系，而不是它天然等于目标过程。

循环的闭合是物质交换的约束，并不意味着每条路径具有相同速度，也不意味着全部过程已经获得解释。

◆ 44.2 陆域出流怎样进入海洋

一条河流的出口既是陆地流域的边界，也可能是海洋控制区的输入界面。只要两边使用同一位置、同一时间窗和兼容单位，这个接口就能把两套核算连接起来。困难在于现实的河口不是一根单向水管：潮汐往复、盐淡水混合、支汊与湿地交换，都可能使瞬时流量和净输入明显不同。

假设在河口断面上观察到一小时向海流动、下一小时向陆流动。以向海为正，对流量绝对值按时间积分，得到累计往复通过体积；对有符号流量按时间积分，得到该时段的净输运体积。若需要时均通量，再除以总时长。两者都可以有意义，但回答的问题不同。若

研究淡水对海区的长期输入，需要适当定义净淡水通量；若研究污染物或盐分交换，往复运动及浓度分布又不可忽略。一个“出海量”名词不能替代这些区别。

重复计算也容易发生在空间汇总中。上游测站记录的水继续流过下游测站，把两者直接相加便重复计算同一股水。水文单元可以是完整流域的一部分，这一规范提醒在跨陆海汇总时同样适用。[C01] 应区分累积出流与局部新增贡献，并根据河网关系决定求和方式。表格看似整齐并不保证它符合网络结构。

遗漏则可能发生在没有显著河道的边界。分散地表排水、地下水与海水的交换、人工排放，都可能与研究问题有关。但发现残差后不能把它们当成便利的填空项。若一个未知通量恰好被设成让预算闭合的数值，这只是残差定义，不是独立测量。要赋予它机制身份，还需要水位、示踪、地层与局地流动等约束。

盐度提供了连接水量与混合的另一种方式。淡水输入能够改变海区盐度，但盐度也受外海交换、降水、蒸发和混合影响。若用盐度变化反推淡水通量，就必须同时处理其他过程，否则同一个盐度下降可能被赋予多个不同原因。反演并非不可靠，而是依赖明确的控制条件。没有这些条件，精细的数字可能只是多个未知数之间的一种分配。

时间对齐同样不能省略。河流出流高峰发生在上游降雨之后，近岸海水对输入的响应还可能继续延迟。如果把陆域月平均出流与海洋日资料直接比较，短时信号会被不同程度平滑。某种看似相位差的现象，可能来自记录方式而非真实传输。研究者应同时报告测量时刻、平均窗口与所研究的响应时间。

可以把接口核验分成四层。几何层确认两边是否使用同一边界；计量层确认单位与方向；时间层确认窗口；机制层确认交换项是否具有独立证据。前面三层未解决时，直接讨论复杂反馈很容易把资料错误升级成自然机制。相反，扎实的接口工作虽然不够戏剧化，却常常决定系统研究能否成立。

在发生学意义上，陆海系统不是把“陆地水文学”与“海洋学”两个名称并列。它是在接口处使两边对同一通量承担一致责任。一个部门提供的出口数据，会限制另一部门能够解释的海区变化；海区无法闭合的预算，又可能促使陆地重新检查遗漏的入海路径。只有这种互相约束实际发生，跨圈层才从口号成为研究对象。

河口还显示浓度与通量之间的差别。某种物质浓度降低，可能来自淡水稀释；若水量增长更快，其总输出仍可能增加。在断面浓度可由代表值近似时，输送率是体积流率与浓度的乘积，累计输送量还需对时间积分。若断面上浓度与流速分布不均，则应对局地速度与浓度的乘积作面积积分。若采用普通平均值，均值相乘可能遗漏共同变化项；只有相应协方差为零时，乘积的平均才等于平均值的乘积。采用流量加权浓度可以保留输送所需的信息，但须明确权重与积分范围。

这个简单统计差别具有实质后果。洪峰期间浓度可能升高或降低，取决于冲刷、稀释与来源变化。若只在平水期采样，估计出的年度物质输入会偏离真正交换。于是，陆海接口的研究会反过来改变陆域采样时间：为了核算营养盐或悬沙，需要在水量变化最快的阶段取得相应组分记录。系统问题由此具体改变了 D 的组织。

河流与海洋控制体连接时，还需要检查边界是否移动。潮滩淹没与露出改变有效水域，水库或湖泊水位变化也会改变面积。若把储量变化全部归为水深变化，却忽略面积变化，预算可能产生系统偏差。固定边界模型可以近似处理，但必须说明被忽略的几何变化是否小于任务允许误差。

同理，冰川退缩或积雪覆盖改变时，原来位于储库内部的水可能进入新的输运路径。此时不仅通量数值改变，控制系统的连接也可能改变。水循环的整体结构并非永远固定的环，它可以在地貌、气候与人为工程作用下重组，核算规则需要随实际连接更新。

◆ 44.3 相变与能量约束

蒸发使液态水转变为水汽，同时涉及能量交换。凝结使水汽形成液态水，也改变周围空气的能量状态。因此，水循环不能只靠水量守恒解释。两处具有相似水量输入的地表，若可用能量、空气湿度与湍流交换不同，蒸发响应也可能不同。水的可得性与能量的可得性共同约束实际通量。

可以用简化地表能量收支帮助理解：净辐射在适当条件下分配给感热、潜热、地表向下传热以及储能变化。这里潜热通量可以与蒸发质量通量通过相变潜热相联系。这个关系提供量纲桥梁，使水量交换与能量交换能够相互核查。但它不意味着测到净辐射就能唯一确定蒸发，因为其余分配项仍可能未知。

例如，湿润地表具备充足水源，增加的能量可能较多进入蒸发；干燥地表缺水，类似能量变化可能更多表现为温度上升与感热交换。这里的差别不应被简化为“水越多蒸发永远越强”。实际蒸发还受到大气输送能力、植被调节和近地层条件影响。状态决定过程的可用路径，过程又改变状态。

植被使这种关系更具层次。植物可以通过根系取得水分，通过叶面交换把水汽送入空气，同时改变地表粗糙度、遮阴和辐射吸收。若只把植被写成一个增加或减少蒸发的固定系数，就可能忽略季节、生长状态和水分胁迫。研究中是否需要打开这些过程，取决于任务尺度与资料是否足够支持辨认，并非所有模型都应无条件增加生理细节。

海面蒸发同样涉及空气与海水两侧条件。水汽梯度、风与稳定度影响交换，海洋内部输运与混合又影响海表温度。一个局地蒸发异常可能既是大气变化的响应，也反过来影响热量与盐度。要判断反馈，就必须给出明确回路，而不是把任何双向相关都称为反馈。

例如，一个候选回路可以写为：某种扰动改变海表温度，温度变化改变蒸发，蒸发引起潜热损失并改变海表温度。第一箭头的原因、第二箭头的交换条件和第三箭头的热量效应必须分别检验。若温度变化实际上由水平输运维持，而蒸发只占很小份额，就不能把该回路描述成主要调节机制。反馈强弱是相对于其他过程而言的判断。

尺度转换也可能改变解释。瞬时局地蒸发与大范围季节降水之间隔着水汽输送和凝结条件，二者不应直接相连。若采用大气控制体，水汽储量变化与水平通量辐合需要进入预算；若采用地表控制体，则要明确哪些过程位于边界之外。同一现象可以用不同控制体分析，但不同公式中的项不能在不说明转换的情况下拼接。

教学中常把水循环的动力归为太阳辐射和重力。这是有用的高层概括，却不能预测某一项通量。太阳提供能量，不意味着每个蒸发变化都直接由太阳变化引起；重力支持流动，也不意味着路径与速度不受地形、阻力及压力影响。机制解释要从高层能量来源走向具体条件，才能形成可检验命题。

相态、温度、湿度、能量与输运的相互规定，在蒸发中尤为集中。水分充足而空气近于饱和时，水汽梯度仍可能限制交换；空气干燥而地表缺水时，可用水量又成为限制；交换增强后，潜热损失还会改写地表温度。这些特征共同决定蒸发通量作为 S 如何显露，时

空比较构成 D，限制条件之间的关系构成 E。因而同一蒸发异常必须同时放进水量与能量账目，不能仅由其中一项推出全部原因。

海水结冰与融化也说明相态不仅改变能量，还改变盐分分配。冰形成过程中盐通常不能简单按原比例进入固相，周围水的盐度与浮力可能随之改变；融化又提供淡水。若只在水量图中画一条“海水—海冰”箭头，就会遗漏其对密度与流动的后果。海冰因此可以同时进入质量、能量和动力预算，但各预算中的量必须分别定义。

陆地积雪也具有类似的多重身份。它储存水，改变地表反照率，并以融化时机影响径流。降雪总量相同而融化速度不同，河流响应可以显著不同；温度变化不仅改变相态，也改变水释放的时序。由此可见，水循环中的“同量”从来不自动意味着“同过程”，第四十二章的结论在跨圈层尺度继续成立。

对能量与水量共同闭合的要求，可以排除某些看似合理解释。如果某项蒸发估计需要的潜热长期超过可用能量，却没有合理的能量输送或储存变化补充，那么至少一项估计或边界定义有问题。反过来，能量收支合理也不能唯一决定蒸发，因为感热等分配仍可能不同。多预算的价值在于交叉约束，而不是让任何一条预算包办全部机制。

◆ 44.4 闭合之后仍有哪些未知

水量预算近似闭合，是研究的重要成绩，却不能直接宣布机制已经唯一确定。两种模型可能以不同方式分配蒸散、地下交换与储量变化，而都能重现同一出口流量。输出一致只能说明它们在该观测上暂时难以区分，并不能证明内部过程相同。

同样，机制合理也不保证预算已经可靠。研究者可能正确识别蒸发需要能量，但缺少代表整个区域的净辐射与湍流资料；可能正确识别地下交换，却没有足够资料估计其数量。机制证据与数量证据应分别评价，不能因为其中一项较强，就把另一项的不确定性隐藏起来。

最有效的下一步通常不是盲目增加变量，而是寻找竞争模型预期不同的观测。若两个模型在出口流量上相同，在土壤水分或地下水位上不同，那么内部状态观测比更多同类出口资料更有区分力。若差异集中在退水阶段，整年平均值则不是合适检验。证据矩阵应把“模型不同处”与“可观察差异”一一连接。

误差传播也不能只靠独立误差平方相加。降水与蒸散产品可能共享输入资料，多个储量估计也可能来自同一模型。相关误差会使预算看起来比实际更精确。研究者需要说明资料之间的依赖，尤其在多个卫星产品和同化结果同时使用时。来源多样不自动意味着证据独立。

系统研究的成熟并不以箭头最多为标准。增加连接可能揭示遗漏过程，也可能引入无法约束的自由度。一个边界清楚、误差可追踪的简化系统，有时比一个囊括所有圈层却不能检验的模型更有解释力。整体性应体现为关键交换被正确连接，而不是所有名词都进入同一张图。

可以把两种未决状态作更精确比较。第一种是预算闭合但分配等效：不同内部通量组合给出相同总变化，下一步需要内部状态或示踪约束。第二种是预算不闭合但机制明确：各过程方向已有证据，下一步主要需要提高数量估计或补齐边界。两者对应的观测投资不同。如果把它们都叫作“数据不足”，研究计划就无法辨认最值得补测的地方。

核算结果也应按用途保留不同细节。教学图可以突出主要路径，研究图则需要让储量、通量和残差分开可读。尤其不能把由闭合条件求得的未知项，画得与独立观测项具有同等证据地位。若后来新增测量使残差分配改变，读者应能辨认是总预算更新，还是原先合并的路径被拆开。图示保留这一区别，才有助于知识继续修订。

水循环由此成为带有尺度、边界和误差的交换系统。它的共同性来自接口两侧使用同一交换量，解释力则来自不同预算能够相互限制：水量允许的分配，还要接受能量与物质组成的检查。预算闭合以后，内部路径仍可能未决，下一步观测应落在这些候选路径预期不同的地方。下一章将进一步追问：当任务从理解水循环转为洪峰预警或水资源管理，这个系统需要改变到什么程度?

第 45 章 任务改变，水文对象是否必须改变

◆ 45.1 三种问题，三张需求表

同一条河流可以同时进入三个工作现场。预警人员关心未来几小时的水位是否超过某个行动条件；水资源研究者关心未来一段时期可利用水量与储量如何变化；资料管理者关心不同年份、不同机构的数据能否归入一致的空间单元。三者都谈论流域，却不因此研究同一个尺度下的完整对象。

洪峰问题需要保留时间细节。一次事件总量预测正确，如果峰值到达时间错误，仍可能无法支持及时行动。长期水量问题则必须处理季节与年际储量变化、蒸散与外部输入；个别短时峰值误差可能对总预算影响不大。资料归档需要稳定编码、版本映射与来源追踪，它的首要标准不是即时预测性能，而是记录在流转中不失去身份。

这三张需求表决定了不同的省略权。年水量模型可以把部分短时过程平均化，洪峰模型不能随意这样做；归档系统可以保存一项未完全解释的观测，过程模型却必须判断它如何进入计算。任务不同不意味着自然规律不同，而是哪些差异必须进入对象发生了变化。

USGS 关于水文单元的规范为这种区分提供了具体依据：资料单元可只包含出口上游汇水区的一部分。[C01] 这种设计对组织资料很有用，但若直接作为洪水模型的孤立控制体，就可能遗漏上游输入。由管理用途走向过程用途，需要重新连接拓扑关系；由过程用途走向行动用途，还需要说明误差造成什么后果。

三种任务也对应三种验证对象。洪峰模型需要检验在规定提前量下的峰高与到达时间；水量模型需要检验同一时段内的输入、输出和储量变化；归档系统则需要证明旧记录在编码或边界更新后仍能定位。输出变量、时空范围、可用信息和验证目标共同限定成果的身份。只有这些条件能够对应，两项都称为“某流域模型”的成果才有直接比较的基础。

在 SDE 语言中，任务进入 S 的定义，并不使科学任意化。洪峰与年水量都是可测的，但它们要求 D 保留不同差异，也要求 E 展开不同的储留与连接关系。任务对对象的作用，首先体现为哪些差异必须辨认；真实过程则限制这种辨认能否成功。

预测时距还决定什么信息有机会发挥作用。对短时洪峰，当前降雨与河道状态可能最紧迫；对长期水量，季节气候、储库变化与管理情景可能更重要。一个变量在某任务中几乎没有预测增益，不意味着它在自然过程中无关，只可能是其作用在该时距内尚未显露，或已被其他变量间接包含。

归档任务则关注另一种持续性。若测站名称改变、坐标修订或边界更新，旧记录必须能够追溯。它不直接生成洪水机制，却为任何后续机制研究提供可复核资料。把归档视为低于理论的机械劳动，会忽略对象身份一旦丢失，再复杂模型也无法恢复的事实。

◆ 45.2 空间细化不总是进步

把一个流域划成更多格网，看起来总比粗略划分更接近真实世界。它确实可能保留坡度、土壤、土地覆盖与降雨的空间差异，减少平均化造成的错误。但分辨率的增加同时扩大了参数与资料需求。若每个小格都需要独立参数，而观测仍只有一个出口，精细结构可能无法被数据辨认。

这种困难可以通过一个教学比较理解。甲模型只区分两个响应区，乙模型区分二十个响应区。如果只有出口流量，两者可能都拟合良好。乙模型的许多参数组合却可能产生同样输出，研究者难以知道哪一组合对应真实过程。更多细节提供了更多可调空间，却未必提供更多可验证知识。

空间细化还可能与输入分辨率不匹配。把粗雨量资料插值到很细格网，地图变得光滑精美，但新增格点并没有带来独立降雨观测。如果后续模型把这些格点当作真实小尺度差异，便把插值假设转化成了自然事实。需要区分数据的显示分辨率、有效信息尺度与模型网格尺度。

类似问题也出现在土壤和土地覆盖图中。一张分类图上的边界可能有位置误差，分类名称也不能唯一决定水力参数。若模型过度精细，边界的小偏差会改变汇流路径；若又用参数调整抵消这些误差，最后得到的局地参数可能具有很弱的物理意义。精度应沿着完整资料链评价，而非只看最后网格有多小。

这并不支持永远使用简单模型。若独立观测表明某些子区响应差异稳定，而且这些差异影响任务输出，细分就有明确理由。关键是让复杂度跟随可辨认差异，而不是让观测被迫追随预设复杂度。可以先比较模型在哪类事件失败，再选择最可能有解释力的细分。

评价细化收益时，独立检验特别重要。粗模型和细模型应在相同信息条件下比较，不能让细模型多使用一批未来资料而仍把改善归于结构。除了平均性能，还应检查极端事件、不同季节与不同前期状态。若改善只发生在训练时段，细化可能主要吸收了偶然性。

一个没有改善的结果仍有知识价值。它可能说明现有任务不需要那种空间细节，也可能说明资料不足以约束细节。两种结论应区分：前者支持保持简化，后者提示需要新的观测。把两者都写成“细模型失败”，会错失下一步应当简化还是补测的判断。

格网加密怎样改变知识，可以沿 S、D、E 分别检查。S 所指的可用流域模型，应在任务误差上获得改善；D 增加的应是被观测区分的空间响应，而非插值生成的格点差别；E 展开的子区关系，则要让土壤、坡度与连通路径受到共同约束。若一个参数只用来抵消另一项输入偏差，即使地图更精细，也尚未辨认出新的局地过程。

模型选择还需要考虑运行代价。一个计算极慢的精细模型，可以用于事后研究，却未必在预警窗口内给出结果。若输出晚于行动所需时刻，物理细节的增加未必改善任务效果。这里应比较的是在相同时间和信息限制下的有效性能，而不是各模型在无限资源条件下的最佳表现。

模型集合提供一种折中：多个结构或参数方案表达不确定性，而不是只交付一条最优曲线。但集合成员如果共享同一输入偏差，集合范围仍可能过窄。多个模型不是自动独立的多个证人。需要了解它们在哪些假设上真正不同，在哪些地方继承同一资料或代码。

分辨率选择还应接受收益递减检验。若从粗到中等细度明显改善，而继续细化几乎无益，就应寻找限制信息的瓶颈：可能是降雨资料、参数不确定性或输出测量，而非网格本身。把资源投入真正限制判断的环节，比维持形式上的精细更能提高任务可靠性。

45.3 责任怎样进入验证

科学模型的误差与行动中的损失并非同一个量。预报洪峰高了一点，可能增加不必要的准备成本；预报低了一点，可能让行动不足。两者在统计上可以大小相同，在后果上却不对称。预警系统需要明确这种差别，但不能从水动力方程直接推出社会应怎样权衡。

自然科学负责说明事件发生的可能条件、预测不确定性与模型适用范围。行动阈值还涉及受影响人口、资源、行动成本与制度责任。将两者连接是必要的，但必须保留分工。某个概率本身不是一条行动命令，某个管理偏好也不能改变洪水传播的物理过程。

假设一个预警系统提高敏感度后漏报减少、误报增加，这不是单一“准确率”可以完整评价的变化。应分别报告不同结果，并说明在实际任务中怎样分配权重。权重由谁制定、依据何种公开程序，都属于可审查问题。研究者不应在技术模型内部悄悄设置社会取舍，再把输出呈现为纯自然结论。

验证时间也影响责任。模型开发时可以使用事后整理资料，真正运行时却受到数据延迟、缺测与计算时间限制。若测试忽略这些限制，便验证了一个现实中不可执行的系统。应把“在当时可得信息下能否作出判断”与“事后能否重建事件”分开。这不是降低科学标准，而是让标准与用途一致。

人工干预还可能改变模型所预测的对象。水库根据预报提前放水，居民和管理者采取行动后，观测到的结果不再是没有干预时的过程。若仅以最后灾害损失较小判断预警是误报，就可能否定了有效行动；若仅以损失较大判断模型失效，也可能忽略执行环节。需要区分危险过程、预报、行动与最终后果四条链。

这让“责任进入对象”具有精确含义。它不是说科学真值可以随价值偏好改变，而是说被验证的系统范围扩大了。若任务只是预测河流流量，可以聚焦自然与工程过程；若任务是支持预警，就还必须观察信息传递、行动可达性和执行时距。一个环节的成功不能代替整个系统的有效性。

预警链条要求各专业在交接处说明结果的用途。水文学给出流量与时间的不确定范围，决策分析比较行动方案的后果，制度研究判断信息是否及时到达有能力行动的人。若流量预报合格而发布延迟，失败发生在传递环节；若信息及时却没有可执行方案，失败又

在另一处。以整个系统作为 S，并非消除分工，而是让局部成功与最终任务之间的缺口能够被定位。

概率校准与区分能力也应分开。一个系统可以善于把高风险事件排在前面，却系统高估概率；也可以总体概率比例接近真实，却不能有效区分具体事件。前者需要校准，后者需要改善信息。单一准确率无法显示这两种差别，任务验证应让评价指标对应实际使用方式。

模型发布之后，使用者行为还会改变资料。某些被高度关注事件得到更密集记录，平常事件却缺测较多，后续训练资料便带有选择。研究者若只分析记录最完整的事件，可能高估常规运行条件下的可靠性。知识生产不仅发生在模型内部，也发生在什么事件值得记录、由谁记录的制度安排中。

◆ 45.4 可迁移结论与局地例外

一个模型在某流域有效，转到另一流域是否仍然有效，需要重新检查气候、地形、储留、人工干预与资料条件。参数可迁移、结构可迁移和任务可迁移是三个不同命题。某种结构能够使用，并不意味着原参数不必调整；某种预测关系仍有效，也不意味着原来的行动阈值适合新的地区。

迁移可以有三种结果。第一，原对象及主要关系保持适用，仅需常规资料更新。第二，核心结构可保留，但局地响应区、边界输入或时间尺度必须调整。第三，原对象遗漏了决定性过程，需要重建控制范围。研究者应让证据决定属于哪一种，而不应以“普适模型”的目标预先排除第三种可能。

局地例外的价值在于是否能够提前辨认。强人工调度改变出流规则，特殊地下连通改变边界交换，不同降水机制改变输入结构；这些条件若已知，就可以在迁移前决定保留什么、补测什么。相反，每次失败后才增设一个无法独立检验的局地参数，便无法区分真实例外与事后补偿。适用范围因此应由可观察条件划定，而非仅以模型在原流域拟合良好为依据。

跨地区比较还要避免资料条件造成的不公平。一处流域拥有密集测站，另一处只有粗略降水估计，性能差异未必来自自然结构。如果要讨论可迁移性，就应同时报告信息条

件。否则，“某地更难预测”可能实际意味着“某地测得更少”。对象的科学身份始终包括可以观测它的方式。

任务改变也未必导致自然对象重建。如果同一套流量概率已经足够，变化只发生在行动权重，那么应称之为决策层调整。只有任务要求新的时间尺度、边界或内部过程，且这些变化影响观测与解释关系时，才适合讨论对象改变。严格区分层级，能够避免把任何管理改动都夸大为科学革命。

对象卡因此不应只写地理名称与模型版本，还应说明验证过的时距、事件范围、资料延迟、主要失效模式和必要人工输入。它既帮助研究者迁移，也帮助使用者判断何时不能照搬。一个结论能否迁移，最终由这些条件能否在新场景中保持决定，而不是由模型名称是否先进决定。

跨地区可迁移的，常是区分错误层级的办法。先核对边界通量，判断异常是否由遗漏输入造成；再检查状态记忆，判断相同输入是否落在不同储留条件上；最后检查传播结构，判断峰形变化是否来自时延。这个次序并非所有流域的固定操作流程，但它能防止用一个局地参数包揽三类问题。具体阈值与参数仍由当地资料约束，诊断问题则可以保持共同语言。

这一编从边界、状态、输运到交换与任务，展示了水文知识怎样获得不同但相互可接的身份。它们的统一来自守恒和可追踪的过程，差异来自尺度与用途，和谐则来自接口清楚、责任明确。下一编将离开现代测站，把同样的问题带入过去：当没有当时的雨量计和温度计，树木、冰层与沉积物怎样成为气候证据？

第 10 编

古气候生命与地球系统

第 46 章 过去的气候怎样留在树轮、冰芯与海底

第 47 章 1976 年的轨道论证：多种记录怎样指向同一段过去

第 48 章 1967 年的辐射—对流模型：气候反馈怎样成为问题

第 49 章 氧化世界的形成：生命、岩石与大气如何共同进入地史

第 50 章 1988 年的地球系统纲领：一张整体图怎样组织研究

第 46 章 过去的气候怎样留在树轮、冰芯与海底

46.1 自然记录者的不同笔迹

一圈木纹、一层冰、一枚埋在海底的微体化石，都可能成为谈论过去气候的材料。但它们没有用同一种语言记录世界。树轮来自生长，冰芯来自降雪积累、压实和保存，海底沉积来自颗粒生产、输运、堆积与改造。要让它们承担温度或降水的证据责任，必须先理解各自怎样形成。

NOAA 对代理资料的说明列举树轮、冰芯、珊瑚、洞穴沉积及海湖沉积物等天然记录来源。[C04] 这一分类提供了入口，却不保证同一类记录具有相同响应。树轮宽度与木材密度不是同一指标，冰中水同位素与封存气体不是同一信息，沉积物中的生物组合与矿物成分也不应互换。自然记录者之间的差异，首先发生在被读出的具体特征。

树木在生长季形成组织，生长受水分、温度、光照、竞争、年龄和损伤等多种条件影响。一条宽年轮并不直接等于某年温暖，也可能反映水分更充足或生长空间改变。只有在特定地点与生理条件下，某个气候因素形成足够稳定的限制，年轮特征才可能用来重建相应变量。记录单位具有年度结构，不意味着记录的是全年平均气候。

冰芯则把不同信号放在同一根柱体中。冰本身的化学与同位素特征，和气泡中的大气成分，经过不同形成与封闭过程。同一深度处的冰和气体不必具有相同年龄。若把二者当作天然同步的记录，便可能制造虚假的领先与滞后。深度一致是空间事实，年代一致却需要另外论证。

海底记录的时间分辨率常受到沉积速率、采样厚度与生物扰动影响。一个样品即使被赋予单一年龄，也可能混合了一段时间内形成的材料。细小生物在不同季节或深度生活，壳体到达海底之前还可能被搬运。海底保存的特征因此可能代表某种加权环境，而不是采样地点上方某一瞬间的海水状态。

代理研究必须区分三个环节：天然过程留下特征，实验室操作读取特征，解释模型把读数转换成气候量。每一环都可能带来误差与选择。样品中某些组分被优先保存，实验室只测其中一部分，校准又只针对某个变量。若最后只展示一条温度曲线，读者很容易忘记曲线已经包含多次转换。

形成、读取和转换三个环节能够接受独立检查，天然材料才逐渐取得气候证据的身份。树轮宽度是否可测是一层问题，它是否主要响应夏季水分是另一层问题，这种关系能否延伸到仪器期以前又是一层问题。代理世界的形成，体现为这些层次逐渐接通：古代气候留下的痕迹获得明确目标、响应条件与时间范围，从而能够支持或排除关于过去的特定判断。

天然记录还具有选择性保存。树木只在能够生长的地方留下年轮，冰只在积累与保存条件合适的地方形成长芯，海底沉积记录则受侵蚀与溶解影响。记录的地理分布不是对地球表面的均匀抽样。若某一时期或地区缺乏记录，不能把空白解释成气候平静；它可能只是没有合适的保存窗口。

保存条件本身也可能随气候变化。更湿润的环境或许促进某类记录形成，却破坏另一类记录；海洋化学变化可能影响壳体保存。于是，样本数量的变化不完全独立于所研究气候。重建中若只看可用样本，可能忽略记录系统也在变化。对天然档案的研究，必须包括档案为何存在和为何缺失。

不同天然记录还具有不同的季节取样。植物主要在生长季留下组织，某些海洋生物具有特定繁盛季节，降雪积累也可能在一年中高度不均。即使年代被标成同一年，实际记录的时间权重仍然不同。若比较这些记录重建的“年温度”，必须说明季节偏重如何处理。年度标签统一，只解决了日历层面的对齐。

◆ 46.2 标定关系怎样成立

把代理读数转成气候变量，通常需要在现代观测重叠期建立关系。研究者同时拥有某种天然记录与仪器气候资料，比较二者是否存在稳定联系。然而，标定不是找到一个相关系数便告完成。必须先明确代理响应的季节、空间范围与机制，以及哪些非气候因素可能共同影响读数。

以树轮为教学例子，若生长主要受夏季水分限制，使用全年平均温度去拟合，可能得到不稳定或难解释的关系。即使全年温度与轮宽相关，也可能因为它与夏季干燥程度共同变化。选择目标变量应有生理与环境理由，不能只在许多候选变量中挑选相关最高的一项，再把它称为发现了自然规律。

样本选择同样影响结果。某些树木对水分敏感，另一些生长在不同微环境中，响应较弱。如果只保留符合预期的个体而不说明规则，关系会被人为加强。合理筛选应基于独立的保存、测量和生态标准，并报告被剔除样品的原因。研究者不能以“不像气候信号”作为删除反例的循环理由。

年龄趋势与生长几何也会影响木质记录。去除非气候趋势可以帮助提取目标信息，但处理方法可能同时削弱长期气候变化。于是，一个年代序列能够较好记录年际起伏，却未必同样可靠地记录长期均值漂移。不同频段需要分别评价，不能把短时拟合成功直接推广到千年尺度。

代理与仪器资料的空间代表性还可能不同。气象站位于谷地，树木生长在坡地；海表温度资料代表较大网格，壳体形成于特定水层。标定所连接的并非天然同一对象，需要解释空间和过程上的对应。若温度站迁移或周围环境变化，仪器记录本身也可能出现非气候断点，不能把全部不一致都归咎于代理。

有效验证应保留没有参与拟合的时间段。若关系在标定前半段建立，能否解释后半段？若换用不同气候背景，系数是否保持合理？这种检验可以发现条件依赖，但也受到短记录的限制。验证期过短，可能尚未经历过去所包含的极端状态；不能因现代几十年表现良好，就假定在所有古环境中无限有效。

从数学上看，代理读数可以概括为目标气候量、其他环境变量和保存过程的函数。反推气候时，若其他变量未知，问题可能没有唯一解。增加另一种代理有时能够缩小范围，但前提是它提供不同的响应信息。如果两种代理共享同一保存偏差或同一年代处理，简单增加曲线数量并不能成比例增加独立约束。

标定关系的失败可能具有建设性。某些记录在现代后期不再按原关系变化，提示生理限制、季节响应或环境条件改变。正确做法是调查变化，而不是强行调整曲线以恢复预期。一个代理适用范围的发现，与其成功范围的发现同样属于知识进步。

标定使读数、目标气候量与适用条件共同取得稳定身份。S 是有季节和空间范围的重建变量，D 是天然记录与仪器观测之间持续比较的差异，E 则把生长或沉积响应同保存、测量条件连接起来。若轮宽原先主要受水分限制，后来转受其他条件限制，变化的不只是回归系数；目标变量是否仍可辨认也成为问题。此时应先检查响应条件，再决定重标定、分段使用，还是停止对该变量的反演。

前向模型与反向校准提供两种互补路线。前向模型从气候与环境条件出发，描述某种代理特征如何形成；反向校准从读数出发估计气候。若前向关系是多对一，反向问题便不可能凭单一读数唯一解决。这个数学事实解释了为什么机制知识如此重要：它限制哪些环境组合合理，从而减少反演的多解性。

例如温度与水分都影响生长，某个轮宽值可能对应温暖偏干，也可能对应凉爽偏湿。增加第二种对水分响应不同的指标，有可能帮助区分两种情况。可是第二指标若受到同一非气候因素控制，信息增益就有限。多代理设计的关键是响应差异，而不仅是材料种类不同。

回归方向也会影响结果。以代理预测温度与以温度预测代理，并不在存在误差时简单互为代数倒数。两边都有测量与过程噪声，误差结构需要进入方法选择。把一条拟合线反过来使用而不检查条件，可能改变估计幅度，尤其在超出标定范围时更明显。

校准中的可重复性还包含程序透明。相同样品经不同处理得到不同结果时，需要能够追查筛选、去趋势与拟合步骤，而不能只比较最终曲线。共享原始读数与处理规则，使后续研究者能够区分真实材料差异和分析选择差异，代理知识也由此获得持续修订的能力。

◆ 46.3 从层位到时间序列

一根岩芯首先给研究者的是深度顺序，而不是日历时间。沉积速度可能变化，层序可能缺失，材料也可能被搬运。要把深度上的曲线转成年代曲线，需要时间锚点及其间的模型。这个转换不只是给横轴换标签，它会改变事件持续时间、变化速率以及频谱中的周期位置。

设两个可靠时间锚点之间相隔一米沉积和一万年。如果采用线性年龄—深度关系，就等于假定该段平均沉积速率足以代表内部变化。它可以是合理的初始近似，却不能凭空排

除短时快速沉积或沉积间断。若中间存在缺失，线性插值会把没有保存的时间分配到现存材料中，曲线看起来连续，历史却并不连续。

年轮交叉定年与简单数圈也不同。个别树木可能出现缺轮、异常生长或局部双重结构，多个样本的共同序列有助于辨认错误。这里重要的是对照结构，而非单个样本自称拥有多少个环。即使年度定位可靠，气候响应的季节范围仍需要另行确定。时间准确与变量解释准确是两种证据责任。

冰芯层位的时间转换还要考虑压缩与流动。越深的年层可能越薄，局地形变也可能改变层序关系。气泡封闭的时间又与冰形成时间存在区别。比较气体变化与冰中温度代理时，必须处理这两套时间，而不能因为它们来自同一根芯就视为天然同步。共同取样不保证共同年代。

海洋沉积物中的重沉积会造成另一种错误。较老材料被侵蚀后进入较年轻层位，实验室测出的年龄可以是真实的，却不代表该层最后沉积的时间。反过来，后期渗入或污染又可能使某些组分显得过年轻。年龄异常应结合沉积结构与材料身份解释，不宜一概删除，也不宜全部当成真实气候事件。

采样间隔限制能够看见的变化。若每个样品平均代表很长时段，快速事件可能被压低、拉宽或完全遗漏。对短周期的识别还受到采样规则影响，过稀采样可能把高频变化误呈为低频变化。无论曲线画得多平滑，插值都不能恢复原来没有观测到的全部变化。

年代误差与测量误差应一起传播。通常图上纵向误差条容易显示，横向年代不确定性却常被隐藏。但比较两条记录谁先谁后时，横向误差可能比读数误差更关键。若两个事件的时间差小于共同年代不确定范围，就不应以视觉上的先后断言因果方向。

更复杂的是年代误差之间可能相关。两条记录若借用了同一个时间标志或同一套调谐方法，其年龄误差便不完全独立。它们看起来高度同步，可能部分来自共同处理。研究者应画出时间约束的来源关系，区分独立锚点、共享锚点与根据目标曲线调整的时间轴。

时间序列的可靠性因而取决于哪些结论对年代误差稳健。长期趋势可能在多条可行时间轴上保持，而两个短促事件的先后却会反转；两者不应共享同一种置信判断。下一章讨论轨道与冰期时，这一区分尤其重要：依据目标周期安排时间轴，会改变频谱与相位的证据地位。独立时间锚点、采样分辨率与共同处理关系，必须随所要支持的结论一并说明。

采样厚度可以看成一种时间滤波。若一个样品混合了若干年的材料，实验室读数接近这段时间的加权平均；短促异常因此被摊薄，低频变化相对保留。不同代理的平滑程度不同，即使它们记录同一真实事件，曲线尖锐程度也可能不同。比较之前若没有统一有效分辨率，便可能把记录过程差异误判为气候响应差异。

年龄模型的不确定性还可能使尖峰在多记录平均中变宽。各记录中的事件真实同步，但时间定位存在误差，平均后峰值降低、持续时间延长。这种结果不能直接解释为气候事件缓慢发生。相反，若先强行对齐尖峰再平均，又可能人为增加同步性。两种处理各有用途，关键是明确它们改变了什么可检验主张。

年代集合提供一种表达方式：不只保留一条最佳时间轴，而是在独立约束允许范围内考虑多条时间轴，观察结论的分布。这样可以区分对时间选择稳健的趋势与高度敏感的先后关系。它不消除不确定性，却把不确定性从脚注变成结果的一部分。

层序中一处突变也可能来自沉积过程，而非气候突然变化。颗粒来源改变、搬运事件或侵蚀面，可以在很薄范围内制造显著读数跳跃。判断气候突变时，应检查该位置的沉积结构、粒度与其他组成是否同步改变。多个变量一起跳变有时支持真实环境转折，有时反而提示记录介质换了身份，必须具体分析。

反之，连续平滑曲线也不能证明气候变化平缓。采样平均、混合和后期扩散都可能抹平急变。自然记录相当于带有响应特性的测量系统，其输出形状需要结合系统本身的传递过程解释。恢复原事件速度时，不能只读最终曲线斜率。

对古气候数据的开放共享因此应包括原始深度、样品厚度和年代模型，而不只提供最后的时间—气候表。未来研究者若获得新的时间锚点，便能重算时间轴；若只剩最终曲线，许多处理选择将不可逆地固化。保存中间环节，是让知识能够更新的重要条件。

◆ 46.4 代理失效也是历史材料

成功的代理容易进入教材，失效的代理往往留在技术讨论中。但思想史若要理解知识怎样形成，就不能只保留最后被接受的转换关系。一次校准失效可能揭示研究者原先把什么当成恒定条件，也可能暴露观测范围曾经有多狭窄。

代理失效首先要定位发生的范围。若关系在相近气候条件下可重复、只在极端状态下偏离，可以收窄外推范围；若同一标定范围内也出现系统偏差，就要检查样品群体、仪器记录或处理程序是否改变。这个区分决定了修订方向：前者主要重划适用域，后者可能要求重建标定关系。某一树种或某一根冰芯属于哪种情况，需要依据其实际记录判断。

失效还可能发生在不同环节。天然记录本身受到新的非气候因素影响，是形成过程问题；样品后期被改造，是保存问题；实验室测量发生系统偏差，是读数问题；使用了不适当的回归关系，则是解释问题。把全部失败归为“代理不可靠”，会丢失可以修复的部分；把全部失败归为“个别噪声”，又会保护本应修正的模型。

可靠的重建应同时给出能解释什么与不能区分什么。例如某条同位素序列可以稳定显示环境变化，却暂时无法把温度与水源变化分开。此时报告混合信号比强行给出精确温度更诚实，也更便于后续研究选择第二种证据。未决状态不是没有知识，而是对已知与未知边界的准确刻画。

多代理一致性也应经过机制审查。不同记录若在独立年代与不同形成过程下呈现共同变化，会增强对某种广域事件的信心；若它们共享筛选、调年或训练资料，一致性就需要折减。不同代理不一致时，不能先假定其中一条错误，还要考虑它们是否响应不同季节、深度或地区。过去的气候本来就可能不是全球同相变化的单一数字。

代理失效也改变思想史的叙述单位。仅按材料排列“树轮史”“冰芯史”，容易掩盖同一种材料曾承担过不同目标变量；仅按成功曲线排列，又容易漏掉被放弃的季节、深度或响应假设。沿着目标变量、取样方式与标定关系的修订来写，才能看见研究者究竟缩小了哪一种主张，又保留了哪一种证据。失效由此成为对象分化的线索，而不只是成功史旁边的附注。

代理的失效还可能揭示目标变量需要重新定义。原先把某指标解释为年平均温度，进一步研究发现它主要响应特定季节，那么正确修订不是简单宣布指标无用，而是缩小它的证据身份。重建的世界因此变得更具体：不是一条笼统温度曲线，而是具有季节、深度或生态条件的环境变化。

这种缩小也可能增强跨代理对话。两条原先“不一致”的记录，一旦分别承认响应不同季节，就不再互相否定，而可以共同描述季节结构变化。和谐并非强迫曲线重合，而是让差异在适当对象定义下获得位置。这种思想同样适用于后面的轨道与气候反馈研究。

代理世界的共同性不来自天然材料完全相同，而来自不同材料能够围绕同一问题承担互补证据。某一记录擅长定位年度变化，另一记录擅长延伸时间长度，第三种记录提供不同物理机制的约束。把它们组织起来时，应保留各自分辨率与响应范围，统一的应是比较问题，不能强行统一它们的全部性质。

树轮、冰芯与海底材料由此形成不同但可相互约束的气候档案。形成过程限定记录什么，保存过程限定留下什么，测量、年代与标定则限定能够从中推出什么。比较这些档案时，保留响应差异与时间不确定性，才有可能把共同变化同处理造成的一致分开。接下来，我们将看一篇著名论文如何让海洋记录与天文计算相遇，并把这种相遇组织成关于冰期节律的论证。

第 47 章 1976 年的轨道论证
多种记录怎样指向同一段过去

47.1 一篇论文中的多条证据

海底记录与天文计算中都出现周期，提出了一个强有力的气候假说：地球轨道的变化参与组织冰期节律。要判断这一假说，首先需要知道海底曲线怎样取得时间轴，随后比较频率、相位与响应幅度，最后检验物理联系能否成立。每一步都可能增强或削弱下一步的推断。轨道理论的思想史意义，就在于天空与冰川之间的联系逐渐被拆成这些能够分别检验的问题。

海斯、英布里与沙克尔顿 1976 年的论文《地球轨道的变化：冰期的节拍器》，以海洋沉积记录与轨道变化的比较进入这一问题。[C05] 原文考察过去约五十万年的记录，讨论近十万年、四万一年及较短岁差频段的关系；结论把轨道变化视为冰期节律的重要调节因素，同时承认气候系统的非线性响应具有必要位置。论文的力量在于把一项长期假说转成多重约束，而不是宣布一个外部周期可以直接复制全部气候变化。

理解其论证，需要把四项工作分开。第一，样品提供随深度变化的观测序列。第二，年代模型把深度转换为时间。第三，统计方法比较气候记录与轨道计算中的频率和相位。第四，机制解释说明外部变化怎样经由气候系统产生响应。任何一项都不能由另一项自动代替。频谱峰再清楚，也无法补救错误的年代身份；年代再准确，也不能从相关直接推出唯一机制。

海洋记录之所以具有吸引力，是因为它可以在相对连续的沉积中保存长时间变化。但长记录并不等于无间断记录，连续曲线也不等于均匀采样。研究者必须根据沉积与测年条件判断哪些变化能够稳定识别。轨道计算则提供另一种信息来源：它的时间结构来自天体运动，而不是直接从气候曲线拟合出来。两种来源之间的差异，是比较具有证据价值的基础。

一项成功的思想转折往往改变什么被视为核心问题。在仅有零散地貌证据时，人们首先问冰川曾到达哪里；获得长序列后，可以进一步问变化是否具有稳定频段、不同记录是否保持相位关系，以及外部驱动与内部响应之间是否存在规律。对象从若干冰期事件，扩展为一个具有节律结构的气候历史。

但“节拍器”不是“全部乐曲”。节拍器规定时间组织，不决定每个音符的响度与内容。轨道变化能够组织季节与纬度上的日照差异，气候响应却还涉及冰体、海洋和大气。这个比喻只有在保持机制分工时才有用；若把它理解为气候被外部机械地逐项规定，反而抹去了论文所保留的内部动力问题。

沉积材料、年代约束、天文计算与统计比较接通后，冰期取得了新的可辨认属性。研究者可以问某一变化属于哪个频段、不同指标是否同相、响应是否随气候背景改变。这里发生的不仅是解释词汇的替换：采样需要满足时间分辨率，年代需要足以约束相位，物理模型需要同时面对节律与幅度。轨道—气候对象正由这些彼此牵制的操作共同成立。

轨道参数之间也具有不同物理意义。倾角改变地轴相对于轨道平面的几何关系，岁差改变季节与轨道位置的对应，偏心率改变轨道形状。它们并不是三个可以直接相加的“冷热旋钮”。要从天文参数走向气候输入，必须先计算特定纬度和季节接收的日照，再考虑气候系统怎样响应。思想史中的困难因此部分来自变量选择：用错误的日照汇总量比较，即使理论方向合理，也可能看不见关系。

这个问题说明，外部世界与记录世界之间需要构造共同问题空间。天文计算以轨道几何为对象，海洋记录以沉积特征为对象，冰期解释却关心积累、消融与热量分配。三者没有天然同一横纵轴。将它们接通，是一系列理论转换，而不是把现成曲线放在一起。转换越清楚，比较就越可能形成真正反证。

◆ 47.2 年代模型能否独立

对周期研究而言，时间轴是一项实质假说。若一段深度间隔被赋予两万年，它的波动会落在一种频段；若赋予三万年，频段就随之移动。纵轴数据没有改变，关于周期的结论却可能变化。因此，年代模型不应藏在方法附录中而被视为无关紧要的准备工作。

独立时间约束可以来自与待检验周期不同的证据来源，但独立性也有程度。某个地层标志的年龄可能依据其他记录标定，其他记录又可能使用轨道调谐；表面上不同的时间锚点，追溯下去可能共享同一假说。需要区分直接独立、部分共享和完全依赖，而不是只给资料贴上“独立”标签。

轨道调谐本身有合法用途。如果研究者已有充分理由相信某些气候变化与轨道关系稳定，可以借此细化时间框架，研究其他问题。困难在于不能随后把被调到该周期上的记录，当作对同一周期关系的全新独立证明。方法合理与验证独立是两个判断，否定重复计数并不等于否定调谐的全部价值。

可以用一个教学例子说明。假设深度曲线有十次起伏，研究者根据某个目标周期调整沉积速率，使各个峰尽量对齐天文峰。调整之后频谱在目标频段增强，是预期结果之一。它说明调谐程序实现了目标，却不能单独说明自然过程确实受到该周期驱动。若独立年代允许的范围很窄，而同样关系在未参与调谐的记录中也出现，论证才增加了新的约束。

相反，完全拒绝所有年代调整也不合理。沉积速度本来可能变化，使用恒定速度会制造另一种失真。关键是把允许的变化由地质与测年证据限定，而不是任意移动时间点以改善吻合。一个严格的年代敏感性研究，应在合理模型集合中考察结论是否稳定，并报告哪些结果依赖某种处理选择。

年龄误差还会影响相位。两条周期相同的曲线，如果时间轴略有偏移，领先与滞后的解释可能改变。对于因果链，时间方向很重要；对于周期存在性，影响可能不同。应分别评价“存在相近频段”与“某变量稳定领先另一变量”，不把前者的稳健性转借给后者。

共同时间锚点形成的依赖可以画成一张关系图：多个岩芯可能共享同一磁性界线、同一标准曲线或同一转换规则。图上曲线数多，并不表示独立时间证据同样多。这一事实直接影响置信程度。若十条记录都继承同一个时间处理中的系统误差，它们之间的一致不能把该误差平均消除。

在 SDE 分析中，年代模型参与规定差异序列 D：同一深度起伏经过不同时间转换，会被辨认为不同周期或持续时间。沉积速率、时间锚点与气候响应的特征纠缠 E，又限制这些转换的可行范围。由此显露的 S 应带着年代条件，例如“某频段在允许的时间轴集合

中持续存在”。这样的对象定义把结论与其依据放在一起，也明确指出新增时间锚点可能修订的是周期、相位还是二者。

年代误差还可以分成整体平移与局部伸缩。整体平移使所有事件一起提前或推后，可能改变与外部强迫的绝对相位，却不改变记录内部的周期；局部伸缩改变相邻事件的间隔，可能直接移动频谱峰。两类误差对结论的影响不同，不能用一个平均年龄误差统一代替。

如果某个时间锚点把长记录分为两段，两段沉积速率分别估计，那么锚点附近可能出现人为速率跳跃。对频谱而言，这种时间变形可能产生额外结构。研究者应检查频段是否在不同合理年龄模型中保持，而不是只展示最整齐的一条曲线。相同原始深度数据在不同时间转换下的比较，有助于把材料事实与年代假设的贡献分开。

独立性也不是非黑即白。一个时间模型可能在少数点使用独立测年，其余部分采用轨道约束；它对长周期存在性和对局部相位的证据地位可能不同。全面否定这类混合模型会丢失有用信息，全面当作独立验证又会夸大结论。适当做法是逐项说明哪个结论由哪些约束支持。

时间轴的证据地位应随使用问题说明。同一调谐记录可以适合研究某种相对变化，却不适合独立验证调谐目标。资料没有脱离用途的统一等级，它的解释资格由具体推断关系决定。

◆ 47.3 周期、相位与机制

周期接近是第一层证据。它说明某种时间尺度在两套序列中都出现，但相同周期可能来自共同驱动、内部振荡、随机过程中的偶然峰，或者资料处理。频谱分析把复杂曲线分解为频率成分，却不会自动标注每个峰的物理原因。研究者必须比较背景变化、记录长度和处理选择。

记录长度决定频率区分能力。若只观察少数几次长周期起伏，对其稳定性的判断会很有限；延长记录可以增加信息，却也可能跨越系统性质改变的阶段。一个频段在前后两段都有信号，未必意味着振幅与相位恒定。长时间平均可能把变化中的关系压成一个稳定峰，分段比较则可能揭示结构转折。

第二层证据是相位关系。假设外部变化与气候响应具有相近频率，还要问响应通常出现在外部变化之后多久，以及这个滞后能否由过程解释。冰体与海洋具有储留和惯性，响应不必即时出现。可是相位也可能随系统状态变化，不能强迫所有周期具有同一个固定时差。

第三层是幅度。某个轨道参数的变化可能清楚提供频率线索，却未必直接提供足够能量变化来解释观测到的气候幅度。此时需要考察季节与纬度分布、反照率、温室气体和海洋过程等放大或调制环节。外部输入小而响应大，不自动否定外部作用；它要求明确内部增益来自哪里。

一个简单教学模型可以分离储热与恢复的作用。以 T 表示温度相对参照状态的偏离，C 表示单位面积有效热容量，F 表示单位面积外部强迫，λ 表示恢复系数，有 $C\,dT/dt = F - \lambda T$。C 的单位为焦耳每平方米每开尔文，F 为瓦每平方米，λ 为瓦每平方米每开尔文。因而方程两侧都是能量通量。对周期强迫，响应还取决于强迫频率；即使外部周期不变，改变 C 或 λ 也会改变幅度与滞后。

真实冰期问题还包含非线性。阈值、冰体几何、反馈强度随状态变化，都可能使输出频率结构不同于简单线性响应。非线性并不是解释困难时随手加入的万能词，而需要具体说明哪一个关系改变了。例如某种积累与消融关系具有不对称性，就应产生可检查的波形或相位特征；若只说“系统复杂，所以什么都可能”，解释便失去约束。

轨道变化也不宜压缩为一个“太阳能量增加或减少”的全球平均命题。不同轨道要素影响日照的季节与纬度配置，气候系统对这些配置的敏感性不同。高纬夏季的融化条件与全球年平均辐射不是同一个变量。选择与机制相关的强迫量，是从曲线相似走向因果解释的重要步骤。

因此，周期、相位与幅度应分别报告。一个研究可以在周期层面很强，在幅度机制上仍未完成；也可以具有合理机制，却因年代不确定而难以判定相位。把多层论证合并成“已经证明”或“完全没有证明”，都会丢失实际知识结构。思想史的进展常常是不同层次逐步获得约束，而不是一篇论文同时关闭全部问题。

在教学层面，可以把离散序列看成不同频率正弦与余弦成分的组合。谱峰表示某些频率对序列方差贡献较大，却不告诉我们它们来自一个持续振荡器，还是来自若干不规则事

件形成的统计结构。有限长度记录的截断还会使一个频率的能量扩散到相邻频率，这就是为什么窗口选择会影响峰形。峰位稳定比峰形漂亮更值得重视。

显著性检验需要合适的背景模型。具有持续性的气候序列本来就可能在低频含有较大能量，不能总以完全独立白噪声作为比较基准。如果背景本身具有记忆，某个低频峰看起来突出，却未必超出合理随机变化范围。选择背景不是形式程序，而是对竞争解释的具体表达。

交叉谱或相干关系进一步比较两条序列在某一频段是否保持关联。它比单独比较两个谱峰更严格，因为两个序列各自有相近周期，不代表它们在时间上共同变化。但相干仍可能受共同处理影响，且在短记录中存在估计不确定性。统计工具增加约束，不会自动免除来源审查。

对于前述教学模型，令 $F=F_0\cos(\omega t)$，取 $C>0$、$\lambda>0$，并在初始过渡项衰减后考察稳定响应。此时温度振幅为 $F_0/\sqrt{(\lambda^2+\omega^2C^2)}$，相对强迫的滞后角满足 $\tan\phi=\omega C/\lambda$。较高频输入更易被热容量削弱；当 ωC 相对 λ 较小时，响应则更接近平衡变化。这个结果把频率、幅度与相位放到同一组参数之下，使不同频段的强弱不再只是一项图形特征。

这仍是单储库线性教学模型，但它给出了一项严格的比较要求：同一组 C 与 λ 必须同时解释幅度和相位。若只能用较大 C 解释滞后，却需另一组不相容参数解释振幅，模型结构就受到限制。转向多储库或非线性解释的理由，应来自这种联合失配，而非仅因其中一条曲线拟合不够漂亮。真实冰期是否需要相应过程，仍由代理、年代与物理约束共同判断。

非线性系统还可能产生阈值跨越。外部变化连续缓慢，内部状态却在条件达到某范围后迅速转变。这可以使冰期记录表现出不对称波形，而非规则正弦。要检验这种解释，需要寻找阈值相关变量、转变前后的状态依赖以及不同周期中是否重复出现相似条件。仅凭曲线不对称就宣布找到了阈值，仍然不够。

比较窗口的长度还影响“稳定关系”的含义。整段记录中的平均相干，可以掩盖某个频段只在部分时期较强。若分段分析显示关系随状态变化，研究者应考虑系统响应是否改变，而不只是寻找一个覆盖全段的平均系数。这里的变化可能来自真实机制，也可能来自记录质量差别，两种解释需要用独立信息区分。

轨道强迫与气候响应还存在空间投影问题。同一种日照重分配，对不同纬度和季节的影响方向可以不同。全球平均曲线可能压低区域强信号，某一局地记录却可能放大局部响应。选择一个海洋岩芯讨论全球冰期时，必须说明它通过何种变量连接到更广范围，而不是因为记录很长就自动具有全球代表性。

这也解释了多地点资料为何重要。若不同海区对同一外部节律表现出不同相位或振幅，这些差异可以帮助辨认输运与区域反馈。多样性不必削弱共同驱动假说；它可能是系统响应结构的一部分。但只有模型能够预期或解释这些差异，而不是事后任意接受，差异才成为理论资产。

◆ 47.4 怎样保留尚未解决的问题

1976 年的论文应按当时的资料与问题理解。它的重要性不要求后续每个细节都维持原样，更不要求现代研究只是在重复其结论。若后来的年代体系、代理解释或动力模型改变，恰恰说明它建立的比较方式能够容纳进一步检验。经典论文不是不可修改的终点，而是使问题获得更严格形式的节点。

研究中应保留代理之间不一致的位置。一个反映冰量与水温共同影响的指标，和一个更偏向区域海表条件的指标，未必在所有时刻同步。差异可能提供过程信息，而不是简单的坏数据。将它们强行对齐，可能删除海洋区域性、响应深度或时间滞后的真实内容。

同样，若某种处理改变后主要频段仍存在，而具体相位不稳定，结论应分层缩小。可以保留对周期关系的支持，同时暂停某条因果先后判断。这样的修订比整体宣布理论成功或失败更精确，也使后续观测有明确目标。

这里讨论的是原论文的论证结构，未对原始序列重新计算频谱。海洋材料确定被比较的信号，年代模型规定频率与相位的时间含义，统计比较检验关联，物理响应解释关联如何可能。阅读结果图可以辨认这些环节，却不足以独立重现数值显著性。因此，对经典的判断应保留两层尺度：其建立了何种可检验关系，以及各项数值结论在重新处理后能否维持。

轨道假说还会改变对记录的要求。识别长周期需要延长时间覆盖，区分相邻频段需要足够记录长度，估计相位则需要更可靠的时间锚点；提高采样密度只解决其中一部分。由

此形成的研究需求，使“更多资料”分化为几种不可互换的改进。具体航次或测年计划是否受到这些需求推动，要由出版、调查与方法记录连接，不能仅从后来曲线的吻合倒推出全部研究经过。

经典论证的后续生命，常表现为问题分工的变化。某一阶段主要争论是否存在轨道节律，随后焦点可能转向不同频段为何具有不同强度、响应为何随时间改变，以及区域记录怎样进入全球过程。理论并没有因为建立一个关系而停止发展，它把更细问题变成了可以研究的对象。

这种发展也提醒思想史写作者避免以今日知识嘲笑旧争论。早期研究者面对的记录长度、时间约束和计算能力不同，某些今天显而易见的比较当时并不可操作。重建当年的证据空间，才能理解为何一篇论文在其时代具有转折意义，而不是仅按今天的技术标准给它补做一道考试。

轨道论证使过去获得了一种新的时间结构，也使时间结构本身接受检验。它把此前似乎只属于描述的地层曲线，转化为可以同独立计算发生定量关系的证据。保留其未完成部分，不是削弱经典，而是准确呈现经典怎样为后来的问题提供可持续的研究空间。

轨道研究最终把外部节律与内部响应同时留在画面中。下一章将从另一篇承重论文进入内部过程：当气候受到扰动时，水汽与辐射怎样改变温度响应，固定什么、允许什么变化，又怎样决定我们所说的反馈？

第 48 章 1967 年的辐射—对流模型
气候反馈怎样成为问题

48.1 模型之前的问题边界

问明天是否下雨，与问大气成分改变后平均温度会怎样响应，并不是同一道题。前者需要追踪某个初始状态的演化，后者需要比较不同条件下系统的统计或平衡性质。两者都使用动力与热力知识，但对初始条件、边界条件和验证资料的要求不同。气候模型的形成，部分就是把这类问题从日常天气判断中分离出来。

真锅淑郎与韦瑟拉尔德 1967 年的论文，把给定相对湿度分布下的大气热平衡作为研究对象。[C06] 它采用辐射—对流框架，通过时间演化趋近平衡，比较控制条件改变后的响应。原文首部明确把湿度约束与敏感性问题联系起来。它不是完整模拟现实地球，而是以受控简化使一些难以在真实大气中单独改变的关系变得可比较。

垂直大气柱把辐射吸收、发射和温度廓线放进同一个可解问题，同时将水平输运、区域云变化与海洋差异留在规定条件之外。这样的边界使湿度约束的作用容易辨认，也决定了哪些结果可以直接解释。平均热响应属于其问题范围，各地区降水和环流则还需要其他结构。评价模型时，先辨认它所保留的自由度，才能判断误差来自近似范围，还是来自范围内部的关系表达。

辐射平衡本身也不等于所有高度温度相同。不同层接受和发出的能量不同，温度结构因此可以存在。但仅靠辐射过程得到的某些垂直温度梯度，可能不符合对流稳定条件。对流调整于是进入模型，使过强的不稳定分层受到重新分配。这里的“调整”是一种参数化表达，不等同于逐个模拟现实中的上升气流。

模型的一个重要优点是允许研究者控制比较条件。真实大气中二氧化碳、水汽、温度、云与环流同时变化，很难仅从观测相关性分离各自角色。模型可以规定一项外部变化，并明确其余变量如何响应。由此得到的不是无条件事实，而是“在这些规则成立时，会有这样的响应”。条件式结论并不低级，恰恰是机制解释的基本形式。

思想史上的转折在于，反馈从宽泛的相互作用变成可计算的差异。研究者不再只说水汽与温度有关，而要比较不同湿度约束下的温度响应，追问差异由哪一条关系产生。模型同时生成结果与反问：如果约束换了，结论会如何改变？这种可替换条件使模型成为实验装置的一种理论形式。

平衡与静止还需要区分。平衡模型可以包含持续的能量输入、输出与内部输送，只要在所研究平均状态下相互抵消。它并不描述一个没有辐射、没有运动的世界。这个区别使气候平衡成为开放系统中的统计问题：能量不断经过系统，状态却可以在一定条件下保持稳定。

初值演化在这里承担求解工具的角色。模型从某个起始温度廓线出发，反复计算净加热与调整，直到趋于相应条件下的平衡。使用时间积分不意味着研究目标就是预测那一条起始状态的真实天气过程。相同数学形式可以服务不同科学问题，必须通过目标与边界条件辨认其知识身份。

这种边界清楚的模型还有交流优势。不同研究者可以保持物理关系相同而改变一个约束，观察分歧究竟来自哪里。若只比较最终温度，分歧像是整体世界观冲突；若比较控制条件与计算路径，许多争论就能转成具体技术问题。可复核模型因此不仅生产数值，也组织研究共同体的对话，使不同解释有一块共同接受检验的地方。

◆ 48.2 固定什么，改变什么

相对湿度是实际水汽状态与相应温度下饱和状态之间的关系量。固定相对湿度，并不意味着固定水汽含量。温度改变时，饱和水汽压随之改变；若相对湿度按规定保持，水汽量就必须依相应关系变化。这个看似细小的区别，决定了模型允许哪条反馈路径存在。

相反，固定水汽含量的比较把温度引起的水汽变化暂时关闭。两套试验可以使用相近的辐射计算，却给出不同响应。差别不能简单归结为某套模型“多加了水汽”，因为两者都可能包含水汽吸收；关键在于水汽作为状态变量是否随温度而调整。静态存在与动态反馈是两件事。

可用一个局部热力关系说明方向。在适当近似下，饱和水汽压随温度上升而增加，这来自相平衡条件。若相对湿度固定，温度上升允许更高水汽含量；水汽增加再影响辐射传输。这里每一步都有前提：固定的是哪一层的相对湿度，如何处理压力与垂直分布，以及

模型是否允许其他过程改变湿度。不能从这一局部关系直接推导所有地区真实湿度必然不变。

原论文使用规定的相对湿度分布，并设置临界递减率等约束来获得辐射—对流平衡。[C06] 这些选择既不是任意装饰，也不是观测定律的无条件复述。它们让一组受控问题能够求解，同时限制结果的适用范围。研究者应问这些规定是否在相关尺度上合理，而不是因为它们是“假设”便否定模型，也不是因为结果合理便忘记假设。

对于一个外部参数的改变，至少可以设计两种比较。第一，暂时保持某些状态量不变，估计直接辐射影响。第二，允许温度与规定的反馈变量调整，求新的平衡。两者的差异提供反馈作用的线索。若没有说明参照状态，就可能把直接效应、快速调整与后续反馈混在同一个“增温”数字中。

固定条件还决定能否进行公平比较。若两套试验不仅湿度规则不同，云、反照率和边界能量输入也不同，那么响应差异不能全部归于水汽。受控试验需要一次改变明确的一组条件，并保留其余条件的对应关系。这与实验室控制变量的逻辑相通，只是模型中控制更容易执行，也更容易隐藏在代码里。

“达到平衡”的判据同样应明确。若系统仍有持续能量不平衡，某个时刻的温度只是过渡响应，不能直接当作最终平衡。不同储热能力影响趋近平衡的速度，因此相同运行时间未必使两个试验同样接近平衡。比较温度时应同时检查能量残差，而不只观察曲线是否看起来变平。

把固定条件与调整规则公开，模型试验便具有可重复的比较单位。另一位研究者可以保持同一辐射计算，只更换湿度规则，查看差异来自直接吸收还是随温度变化的水汽响应；也可以保留湿度规则而改变对流条件，检查温度廓线的作用。争论由此落在可替换的具体关系上。简化的价值正在于让修改位置可辨，而不是使模型免于接受真实大气的检验。

相对湿度的规定还可以从函数关系理解。以水汽压 e 和相同温度下的饱和水汽压 $e_s(T)$ 定义相对湿度 r，有 $e=r\,e_s(T)$；固定 r 时，温度改变便通过饱和关系改变 e。若进一步换算为比湿或其他水汽含量，还需给定压力并采用对应定义。水汽响应的大小因基础温度和压力而异，因而同样的温度扰动未必对应相同水汽增量。反馈必须连同参照状态一起说明。

把变量固定在一种坐标中，也不等于在另一种坐标中固定。固定相对湿度、固定比湿和固定水汽总量分别约束不同量。比较研究若只说“湿度固定”，读者无法知道哪条响应路径被关闭。模型语言的精确性在这里直接影响因果解释，而不仅是术语洁癖。

对流条件同样可以改变响应。若规定垂直递减率，当地表增暖时不同高度温度如何随之改变受到约束；若允许更复杂的稳定度调整，温度廓线可能不同，辐射响应也随之变化。水汽与递减率的作用因此在物理上相互联系，分项分析需要清楚说明如何拆分，而不能假定它们天然独立。

边界表面的热容量也影响过渡过程。若只关心最终平衡，在理想稳定条件下热容量主要影响达到平衡的时间；若关心某个有限年份的响应，热容量就是结果的一部分。把过渡响应与平衡响应并列比较而不说明这一点，会把系统速度差异误判为最终敏感性差异。

一个常见误读，是把模型中固定相对湿度理解为真实大气在每个地点、每个时刻都保持相对湿度不变。模型规定通常针对所研究平均结构，而实际天气存在明显波动。评价假设应问它对目标尺度上的响应是否提供合理近似，不能用任一局地瞬时偏离直接否定，也不能以平均合理为理由忽略区域差异。

同样，规定云量或反照率是一种控制方式，并不表示这些量在真实气候变化中永远不动。简化试验将它们暂时外置，使某条路径得以单独辨认。后续研究可以逐步打开这些自由度，比较新增响应。科学对象因此有层次：先定义受控响应，再研究更完整耦合，而不是要求最初模型一次容纳全部过程。

◆ 48.3 从响应到反馈机制

温度升高会增强向外辐射的恢复作用，这提供了一个稳定化方向；同时，温度变化还可能改变水汽等吸收条件，影响恢复强度。反馈分析关心的是这些随状态改变的过程，怎样改变外部扰动与最终响应的关系。它不是给某个物质永久贴上“正反馈”或“负反馈”的标签。

为说明概念，采用本书的教学模型：设净能量不平衡 $N = F - \lambda \Delta T$，其中 F 表示外部扰动，λ 为所定义参照下的净恢复系数。若 λ 为正，温度升高倾向于减少不平衡，平衡时 $\Delta T = F/\lambda$。若某种反馈削弱恢复，使 λ 变小，同样扰动会产生更大平衡响应。这里的符号约定已明确，不能与采用相反反馈符号的文献直接混用。

进一步把 λ 写成参照恢复与若干调整项的组合，可以比较不同过程的作用。但这些分项不必在所有状态下都是常数，也未必完全独立。水汽变化与温度垂直结构有关，云响应与环流有关，冰雪反照率又涉及空间分布。线性分解是一种有用的局部近似，不是对整个气候系统永恒结构的宣告。

水汽反馈的解释链可以表述为：外部扰动引起温度响应，规定湿度关系使水汽随之变化，水汽变化影响辐射传输，从而改变原温度响应。链条中每一箭头都有物理内容。若温度变化与水汽变化只是共同受第三因素驱动，而没有建立回到温度的作用，就不能仅凭相关性称为反馈。

辐射传输的作用也不宜用“热被完全困住”概括。大气与地表持续吸收和发射能量，系统仍向太空辐射。吸收条件改变后，在相同温度结构下的向外能量交换发生变化，系统通过温度和其他状态调整重新趋于平衡。这样表述既保留能量守恒，也避免把温室效应想象成一只密不透风的盒子。

温度廓线之所以重要，是因为辐射来自不同高度和波段，向外发射不只由地表温度一个数决定。若吸收增加改变了有效发射层次，相应高度的温度结构便影响能量收支。对流调整把垂向温度梯度带入问题，使辐射与热力结构不能各自孤立求解。模型中的 E 因此是吸收、温度、湿度与垂直稳定性的相互规定，而非多个独立因素的清单。

原论文将固定相对湿度与固定水汽条件下的敏感性比较，构成辨认水汽响应作用的早期节点。[C06] 在这类比较中，需要追踪的是被打开或关闭的水汽路径。若进一步允许云、冰雪或海洋随温度变化，响应对象便增加了新的自由度，必须另行界定与检验。这种分层让简化模型的成绩可以明确保留，同时使后续模型究竟增加了什么问题也更容易辨认。

反馈还与时间尺度有关。某些调整较快，某些储库变化较慢。一个短期试验可能已包含大气水汽响应，却未包含缓慢冰体或深海变化。说某模型的“气候敏感性”，必须说明所允许调整的过程和时距。否则不同研究之间数值不同，可能不是谁计算错误，而是它们定义的响应对象不同。

稳定性也不能只看一个反馈方向。局部正反馈不意味着系统必然无限增长，因为其他恢复过程可能占优势；局部负反馈也不保证所有模式稳定，因为空间和时间耦合可能产生不同演化。教学公式中的 λ 只是限定问题中的总效果，不能被用来替代完整系统分析。

受控比较也使 S、D、E 之间的关系具有清楚位置。E 是湿度、辐射与对流条件的相互限定，D 是按规定改变条件后获得的响应差异，S 则是特定意义上的平衡敏感性。所要辨认的 S 反过来决定试验该关闭哪条路径：研究水汽反馈，需要固定其余条件；研究总响应，则需说明还有哪些变量允许调整。这里的分析不另行计算温度，而是辨认不同计算结果为何可能属于不同对象。

反馈还可以用链式求导的形式表达。设某辐射量 R 同时依赖温度 T 和水汽 q，而 q 又随 T 变化，则温度对 R 的总影响包含直接项 $\partial R/\partial T$ 与间接项 $(\partial R/\partial q)(dq/dT)$。这段数学不是额外自然定律，而是清楚分配直接路径与间接路径的方法。固定 q 时，第二项被关闭；允许 q 按规定响应时，第二项进入。模型比较因此对应明确的路径差异。

若再允许云或地表反照率随温度变化，总导数会出现其他项。但增加项不代表这些项已经被可靠认识。每一项的符号、大小和状态依赖，都需要模型、实验或观测支持。把所有未知项合称“复杂反馈”，不能代替逐条估计；同样，忽略未知项也不意味着它们实际为零。

净恢复关系还可以从图像上理解。横轴是温度偏离，纵轴是顶层能量不平衡；曲线与零线的交点给出平衡。外部扰动使曲线位置改变，反馈使斜率或形状改变。一个近似线性的局部范围可以用 λ 概括，若曲线明显弯曲，则同样大小的扰动在不同基础状态下会产生不同响应。这种几何表达说明，所谓敏感性是关系性质，不是温度或二氧化碳单独拥有的固定标签。

对于稳定平衡，偏离之后的净能量趋势应促使系统返回相应状态。若局部恢复过弱，响应时间可能变长；若出现多个平衡，初始状态与路径也可能影响结果。但不能由一般数学可能性直接断言真实地球必然存在某个特定分岔。模型展示可能机制，观测与更完整物理约束决定它是否具有现实身份。

将反馈作为关系而非实体属性，也改变了因果语言。水汽是重要吸收物质，但在一种试验中它可以是外部规定，在另一种试验中则是响应变量。角色由问题与控制方式决定。相同物质进入不同研究设计，承担不同因果位置，这正是对象、操作与属性共同形成知识的一处清楚实例。

能量不平衡还提供对过渡响应的统一理解。若系统持续吸收多于释放的能量，多余部分必须进入储能或其他相应变化。温度不立即达到平衡，并不意味着辐射机制没有作用，

而可能意味着能量正在被较大热容量的部分吸收。把温度趋势与能量收支同时观察，比只看地表温度一个指标更能约束过程。

在教学模型中，加入有效热容量后有 $C\,d\Delta T/dt = F - \lambda\Delta T$。恒定扰动下，系统趋向 F/λ，时间尺度与 C/λ 有关。这个结果清楚分离了最终幅度与响应速度：改变 C 主要改变快慢，改变 λ 同时影响幅度与快慢。真实气候包含多个储库，可能出现多个时间尺度，但单储库推演已经足以防止把“尚未达到”误当成“最终不会达到”。

这些关系也说明参数估计需要覆盖相应时距。短记录可能主要约束快速响应，无法唯一识别慢储库；若强行用它估计全部平衡性质，结果会高度依赖模型结构。时间长度不仅影响统计精度，也影响哪些物理过程有机会在观测中显露。

反馈的识别由此要求路径清楚、参照明确，并允许在条件改变时重新评价其强弱与方向。

◆ 48.4 模型的成功与有限性

评价一个模型至少需要区分机理清楚、经验符合与预测可用。机理清楚意味着关键因果链和控制条件明确；经验符合意味着与独立观测之间存在合理一致；预测可用则要求在实际任务范围内提供足够可靠的信息。三者相互支持，却不能彼此替代。

一个模型可能具有清楚机制，却因空间结构简化而不适合区域预测；也可能在某组资料上拟合良好，却依靠补偿误差而缺乏可信机制。若两个错误正好抵消，输出正确并不意味着内部结构正确。需要借助不同观测变量和不同扰动情境，检查模型是否以同一组合理关系同时解释多个方面。

结构误差与观测误差也应分开。结构误差来自省略或错误表达过程，观测误差来自测量与代表性。扩大观测误差条不能修复一个明确缺失的机制，增加模型细节也不能解决错误资料。把两者分层处理，才能判断下一步应该改模型还是改观测。

简化模型的价值尤其在于让某些依赖关系可见。复杂模型可能同时包含许多过程，却难以追踪某项响应来自哪里；简单模型则可以通过有限控制明确揭示机制。两者的关系应是互相检查，而不是把复杂度当成唯一等级。一个好的简化解释能够指出复杂模型应呈现什么差异，一个好的复杂模型则能检验简化在真实条件下失去多少内容。

思想史据此可以区分问题的扩展与数值的更新。维持同一湿度和对流规则，只改进辐射计算，主要是在原问题内修订结果；改变允许响应的变量，则可能重新定义敏感性。两种变化都重要，却不能只按最终增温值的大小排列。模型版本、试验条件与结果之间的对应，能够说明争论究竟发生在计算精度、过程表达，还是研究对象本身。

这种区分也解释了为何两个结果不同却未必直接矛盾。一项结果让水汽随温度调整，另一项固定水汽，两者首先比较的是不同受控系统；只有控制条件对应后，数值分歧才需要由计算或过程表达来解释。跨案例分析的作用，是在概念比较之前恢复这些条件，避免把“敏感性”当作一个脱离试验设计的通用数字。统一语言的前提，是保留每项结果的生成方式。

模型检验还应包括极限情况。若把相关吸收成分减少、改变入射能量或关闭某条调整路径，响应方向是否符合基本能量与热力约束？极限试验不能证明现实精度，却能发现实现或推理中的明显矛盾。与复杂观测比较之外，这类内部一致性检查具有独立价值。

另一类检验关注参数补偿是否可迁移。如果某套参数只在一个平均状态附近有效，换一个季节、纬度或扰动范围就失效，说明其成功可能依赖局地补偿。机制清楚的简化模型能够帮助识别这种情况，因为它使参数究竟代表什么更可见。模型发展不只是增加分辨率，也包括让解释关系更透明。

因此，1967 年这个节点值得记住的并非一个脱离条件的数字，而是一种严格比较方式：把湿度规则、辐射计算与对流条件共同规定，再看扰动如何产生可归因的差异。这样的研究使反馈成为可以计算、拆分与反驳的对象，也为后来的复杂气候系统保留了一条追问机制的路径。

从固定条件到反馈路径，气候响应已经成为可以分解的关系：参照状态规定比较起点，允许调整的变量规定响应范围，在上述线性近似中，恢复强度限定平衡响应，热容量与恢复强度共同限定调整时标。阅读模型结果时，必须把这些条件与数值一起保留。下一章将把这种分析带到深时：生命、大气与岩石共同变化时，怎样区分产生信号、保存信号和解释氧化历史的不同机制？

第 49 章 氧化世界的形成
生命、岩石与大气如何共同进入地史

49.1 岩石为什么能谈论大气

从古老矿物中的读数推到早期大气状态，要经过形成、输运、沉积与保存等多个环节。岩石保存的是这条历史链留下的特征，其中某些同位素关系能够区分参与过的大气反应。矿物因而可以成为大气证据，但它所约束的是反应及其保存条件，不是一份原封不动留存至今的大气样本。早期氧化史的研究，正从这种间接而可检验的关系开始。

法夸尔、鲍惠铭与蒂门斯 2000 年的研究，以前寒武纪岩石中多种硫同位素关系为入口，指出早期硫循环的重要变化。[C07] 原文强调某些古老样品具有非质量相关的同位素特征，将其与大气反应联系，并讨论约 24.5 亿年至 20.9 亿年前这一宽时间范围内的转变。这个范围属于该论文当时的材料与判断，不能直接当作今日全部氧化史的最终边界。

同位素的意义不在于某一种硫“多一点”就自动代表氧多或氧少。不同过程会按一定关系分配不同质量的同位素。若观测偏离某类质量相关关系，偏离本身便成为需要解释的对象。研究者由此不只比较单一比值，还比较多个比值之间的结构，寻找能够区分过程的额外维度。

这里的思路类似于在二维平面上辨认一条关系线。若许多常规过程主要使数据沿某条关系移动，那么偏离该关系的方向可能提供新的信息。但基准关系、测量精度和转换公式必须明确。不能把任何偏离都叫作新机制，也不能因为一个异常很小就认为它没有意义；关键在于它是否超过分析与模型不确定性，并具有可重复结构。

从测量到大气解释至少需要三座桥。第一，读数确实属于所研究矿物，且实验室分析可靠。第二，矿物保留了与形成时期相关的信号，没有被后期过程完全重置。第三，能够产生该信号的机制受到限制，并能连接到大气条件。缺少其中任何一座桥，结论都会停留在更低层次。

年龄身份同样不可缺少。一颗矿物的年龄、围岩沉积年龄和后期变质年龄可以不同。若把它们混为一谈，信号变化便会被放到错误时间。深时研究中“古老”不是足够精确的标签，必须说明读数对应哪个事件。第十三章关于矿物时钟的区分，在这里重新成为化学历史的基础。

多同位素关系使原来分属不同领域的问题接通。大气反应解释异常怎样形成，海洋储库与输运解释它怎样分配，沉积矿物解释它怎样保存。若只保留任一环，都不足以确定岩石中的组成。早期硫循环由此成为跨圈层的研究对象：一种矿物差异可以限制大气假说，大气机制也会反过来要求研究者比较不同沉积环境中的保存方式。

同位素记号本身也体现了比较的构造。通常先把样品的同位素比值与标准物质比较，得到相对偏差；进一步再用多个同位素之间的预期关系，定义偏离质量相关趋势的量。这里有两层参照：标准物质提供计量参照，质量相关关系提供机制参照。一个样品相对于标准偏移，不等于已经偏离质量相关规律。

这种分层避免了“异常”一词的歧义。异常可以指高于平均值，也可以指偏离某条机制关系；二者的解释强度不同。前一种可能只显示程度差别，后一种可能要求不同过程。研究者必须说明异常相对于什么成立，才能判断它究竟打开了哪一个新问题。

分析精度也需要与自然差异相匹配。重复测量同一样品可以估计部分实验室误差，重复提取不同矿物部分则有助于检查样品非均一性。两者不是同一种重复。若实验室精度很高而样品内部差异大，报告一个极小仪器误差并不能代表整个地质解释同样精确。

◆ 49.2 不同机制争夺同一信号

生命研究容易赋予化学异常过强的故事性。某种同位素分布变化与生命活动相联系，于是岩石仿佛直接证明生物何时出现、怎样改变大气。然而化学信号首先记录过程，不自动记录唯一行为者。大气反应、生物代谢、无机反应、混合与后期改造，都可能影响最终组成，只是方式与范围不同。

多同位素分析的价值，在于让竞争机制面对更严格约束。某种机制能够解释一个比值，却不能解释另一个比值与它之间的关系，就需要修订或退出。法夸尔等人的论文正把

多个硫同位素之间的关系用于讨论气相反应，而非仅凭单一硫同位素跨度来判断早期生物作用。[C07] 这种证据结构比“异常出现，所以生命出现”更强，也更谨慎。

质量相关分馏并不是生物过程的专利。许多热力与动力过程都能产生相关变化，生物活动也在这些物理化学规则中发生。因此，识别质量相关关系本身不能唯一识别生命；识别非质量相关特征也不意味着排除生命参与后续输运与保存。不同机制可能在同一条历史链的不同时段承担不同角色。

应当把“产生信号”与“传递信号”分开。大气反应可以生成具有特定组成的硫物种，水体输运将它们带入沉积系统，微生物活动可以改变硫的化学形式，矿物形成再将部分组成保存下来。最后测到的异常可能主要来自第一步，却必须经过后续步骤才进入岩石。只把所有责任交给一个过程，会看不见保存选择的重要性。

混合尤其容易改变信号强度。两个具有不同异常方向的储库混合，最终异常可能减小甚至接近基准。于是，没有明显异常至少有两种一般可能：产生异常的过程不再有效，或者异常仍产生但被混合和循环削弱。两者对大气与海洋的解释不同，需要额外证据区分。缺失信号并不自动等于缺失过程。

后期变质与流体活动又带来另一层问题。它们可能改变矿物组合、迁移硫或促成交换，但影响程度要根据具体地质与矿物学证据判断。不能因为岩石经历变质就宣布所有信息消失，也不能因为某个比值看似合理就忽略改造。保存性是可以研究的性质，不应成为方便使用的全有或全无判断。

可用于区分机制的证据也不同。矿物内外的空间分布可以检验信号是否均一，岩相与沉积结构可以约束形成环境，多元素或多同位素关系可以限制反应类型，区域重复可以检验局地性。每种证据打开一个不同问题，不宜把它们都压成“更多样品”。样品数量增加只有在增加了新的约束时才显著增强归因。

实验室光化学研究能够帮助建立某种反应产生异常的可能性，但实验条件与早期大气条件之间还需要桥梁。波段、气体组成、压力和反应路径都会影响结果。实验再现某种特征，是支持机制的重要一步，却不直接证明该机制在真实地史中占主导。需要与地质记录中能够获得的独立条件相互限制。

同样，生物实验或现代环境观察可以说明代谢路径怎样影响分馏，却不能不经检验就外推到早期海洋。底物供应、反应速率、开放程度与生态结构不同，信号也可能变化。深时归因的难度不在于没有任何机制，而在于多个机制都具有部分解释能力。

这条信号链给 SDE 的特征纠缠 E 以具体内容：反应产物必须进入可分离的储库，随后迁移与保存不能把有关差异全部抹平。跨样品、跨年龄和跨矿物的比较构成 D，由此显露的 S 是受到这些约束的早期硫循环状态。若矿物之间的差异随后期裂隙而变化，保存环节便需要重新解释；若不同环境仍保留一致异常，则可进一步收紧对共同来源的判断。

把同位素关系看成几何结构，还能说明为什么多同位素比单一范围更有区分力。一个机制可能允许数据沿某条方向变化，另一个机制可能沿另一条方向变化。若观测同时包含两个维度，某些解释即使在单轴上可行，也会在联合关系中失败。这里新增的信息来自关系方向，而不只是多测了一列数字。

不过，混合会使这种几何变得复杂。两种来源的混合关系取决于各自硫含量与同位素组成，不能简单把所有图上直线都解释为同一反应。若样品是多个组分的总体分析，某些原本不同的微观信号会被平均。提高空间分辨率可以揭示组分差异，但也需要重新判断小尺度读数如何代表整块岩石与整个沉积环境。

氧化还原状态与同位素组成也不是同一个维度。化学价态反映电子转移条件，同位素关系记录分配与来源过程。二者可以相互关联，却不能互相代替。某种硫以氧化形态保存，不自动意味着当时整个大气富氧，因为局部化学路径也可能形成氧化产物。由矿物形态推及全球状态，需要额外反应与输运约束。

一个候选解释若诉诸大气光化学，应同时面对几个问题：异常怎样产生，产物怎样离开反应区，怎样进入不同储库，怎样避免在后续循环中完全抹平。缺少最后几步，即使实验室能够产生异常，也未必解释岩石中长期保存的分布。自然记录是整条路径的结果，而非仅仅反应起点的截图。

反过来，若解释主要诉诸后期流体改造，就应预期信号与裂隙、矿物世代或变质分带存在某种关系。若原生沉积结构与多地区重复限制了这种关系，后期改造假说的自由度便缩小。竞争机制不应只作为名词列在讨论末尾，而应给出彼此不同的空间和时间预期。

样品选择也决定竞争机制受到怎样的检验。若全部样品来自同一种沉积环境，就难以分辨信号是全球输入还是局地反应；若只追求跨地区覆盖，却忽略矿物身份一致性，又会把材料差异混入历史变化。合理设计应同时包含可比样品与有意改变条件的样品，使研究既能重复，也能区分原因。

同一层位内不同矿物之间的比较尤其有用。若它们记录相似来源却经历不同形成过程，差异可以帮助识别后续分馏；若不同矿物共同保留某种异常，则可能加强对更早来源的约束。这种推理并非任何一致都支持大气解释，而是根据矿物形成链，判断一致性排除了哪些替代路线。

◆ 49.3 氧化转折的时间怎样确定

一次全球性转变不必在每个地方留下同样的岩层，也不必以同一清晰瞬间被记录。大气可以广泛交换，海洋与沉积环境却具有区域差异；局地保存还受侵蚀、沉积间断和后期改造影响。将分散样品拼成全球氧化历史，需要同时处理自然异质性与记录不完整性。

最基本的资料单元应包含样品位置、地层身份、被测组分、年代依据、保存状况与读数误差。只把年龄和异常值放在散点图上，会隐藏很多差别：某个年龄可能来自直接测年，另一个来自上下地层约束；某个样品属于原生沉积矿物，另一个经历明显再结晶。它们不能完全等权地承担转折时间的证据责任。

设较老样品显示一类特征，较年轻样品显示另一类特征，可以得到一个转变区间。但区间宽度取决于样品分布与年代精度，不能把两点之间的中点当作事件发生时刻。若中间缺少记录，转变可能发生在其中任何满足约束的位置，也可能不是一次单调变化，而包含反复或区域不同步。

连续剖面具有优势，因为它保留相邻层位之间的关系；但连续外观也需要地质检验。沉积间断可能很薄，侵蚀面未必显眼。跨地区对比则可以检验某个转折是否局限于一个盆地，但不同地区的年代关联可能共享不确定性。局地连续性与全球覆盖各有优势，任何一种都不能完全代替另一种。

2000 年研究给出的宽时间区间，应放回当时样品分布与年代框架中理解。[C07] 它的重要作用是把变化安排在一个可继续检验的时间范围内。后来新增样品究竟缩小了区

间、揭示了反复，还是显示不同地区记录了不同过程，需要分别讨论。将后来的精确边界倒填到原论文，会遮蔽这项推进本身；保留原区间，才能看清哪些新增约束真正改变了氧化史的时间结构。

“氧化”本身还可能指向不同对象。大气氧含量增加、浅海局地含氧、深水氧化还原条件改变和某类氧化风化增强，并非完全同一事件。它们可以相互联系，却具有不同响应时间和空间范围。如果标题只写一个“大氧化”，正文就必须展开这些层次，否则读者会误以为全地球所有储库同时完成同一种变化。

一个局地含氧水体可以在整体大气氧仍较低时形成，反过来，大气氧增加也不保证全部海洋立即充分通气。将生命活动、地球化学循环与动力输运相接，才能理解这种不一致。深时系统没有义务按照人类命名的单一时代界线整齐换挡。

检测限也影响“消失”的判断。某种异常降到测量分辨率附近，不等于绝对为零。比较不同研究时，需要注意分析精度、标准与数据处理。若新仪器看见过去没有分辨出的微弱结构，这可能改变对转变过程的理解，却不意味着过去所有测量都错误；它意味着可辨差异的尺度改变了。

年代推理因此有三个层次：某样品何时形成，某地区何时改变，某类全球过程何时转折。由第一层走向第三层，需要越来越多的空间与机制约束。把一个精确样品年龄直接写成全球事件年龄，是常见但严重的越级。第十五章关于界线标准与自然事件的区别，在这里同样适用。

上下限的进一步收紧取决于中间样品能否增加独立约束。若新样品只重复已有端点附近的状态，覆盖更密却未必缩小转变范围；若它落在区间内部，且年代和保存身份可靠，才可能将范围分开。即使如此，单个新状态也只证明该状态曾经出现，持续建立还需邻近层位支持。这样，样品密度、时间精度与状态持续性便成为三个可以分别改进的方面。

如果不同地区给出的范围不完全一致，可以先问记录是否响应同一过程。某地区可能主要记录大气输入，另一地区更受局地水体循环影响。其次问年代是否存在系统差别，再问是否可能具有真实空间不同步。把所有差异立即平均成一个全球日期，会同时丢失机制与年代信息。

此外，转变可以包含起始、持续与完成三个不同时间。最早出现新特征，不等于新状态已经长期稳定；旧特征最后消失，也不等于所有地区同时结束旧过程。用一个年代代表全部转折，适合简化叙述，却可能不适合机制研究。思想史应说明科学家研究的是哪一种时间。

这种区分对生命因果尤其关键。生命活动可能在局地早已存在，全球储库却因消耗与循环保持旧状态；后来其他源汇关系改变，才出现可记录的全球转折。于是，“生命先于氧化”或“氧化先于某类生命扩展”都需要具体对象与时间定义，不能只比较两个宽泛标签。

跨年代样品的保存差异也可能不均匀。更古老岩石通常经历更长历史，但不能仅以年龄推定全部改造程度。应依据具体矿物与地质证据筛选，否则年龄与保存性被混同，可能制造虚假的时间趋势，并误导对转折的判断。

◆ 49.4 共同塑造如何成为可检验命题

生命改变环境，环境改变生命，这个命题太宽，几乎能容纳任何结果。要使它成为科学解释，必须选择一条具体链，并说明每个箭头的方向、载体与证据。例如，某类生命活动改变氧化还原通量，通量改变某储库的化学状态，化学状态又改变营养或代谢条件，从而影响生命活动。闭合链条不是画出一个圆就成立。

氧的积累尤其不能只由产生量解释。即使存在产氧过程，若氧被还原物质迅速消耗，大气中也未必持续积累。产生、消耗、储存与输运之间的差额才决定某个储库的变化。生命产氧的出现与大气显著增氧，因此在逻辑上可以不是同一时刻。这个区分能够阻止把地球历史写成“生物一出现，空气立即变样”的单线故事。

风化连接陆地与海洋。它改变物质输入，也受气候、暴露面积与岩性影响；海洋化学再影响沉积与生物可用资源。若要主张某条反馈是氧化转折的重要原因，就需要不仅证明各过程存在，还证明其方向和规模足以影响所讨论的转变。局地有效机制不自动具有全球主导地位。

有机物或还原物质的埋藏也需要与再氧化区分。某些过程能够把反应产物从快速循环中移走，使长期收支改变；若随后迅速返回循环，净效果又不同。地球化学思想的一项深

化，就是从“有没有反应”走向“反应产物最终去了哪里”。储库与时间尺度决定了同一化学反应的系统后果。

生命过程的可持续性依赖能量、底物与输运，同时又改变这些条件。例如代谢改变某种可反应物质的通量，储库状态随收支变化，再限制后续代谢能够采用的路径。这里可辨认的复合存在体，是能够持续维持或改变氧化还原状态的生命—化学—输运系统。判断这种关系是否成立，需要同时观察通量与储库；仅见生命存在或化学状态改变，都不足以闭合这条作用链。

竞争解释也应保留。若硫同位素资料可以由大气反应与保存变化解释，就不能凭同一资料唯一认定生物通量改变；若不同证据共同限制了大气条件，也仍需独立材料判断生命贡献。一个圈层的变化可以支持另一个圈层的假说，却不能免除后者自己的证据责任。

可以用简化收支说明积累：储库中可反应氧的变化率，等于在所定边界下的生产与输入，减去消耗与输出。若其中某项已按净值记账，就不能再重复扣除它包含的损失。该式没有给出各项数值，却限定了归因方式：生产增加而消耗同步增强，储量可能变化很小；生产近似不变而消耗下降，储量也可增加。因此，氧化增强首先约束的是净收支，尚需独立资料分解源与汇。

这种多源汇结构使深时研究必须比较替代路线。生命活动变化是一条路线，火山或变质气体的还原能力变化、岩石风化与埋藏格局变化也可能影响净收支。具体哪条路线在某次历史转折中重要，需要独立资料。整体解释的价值，是让这些路线在同一守恒框架中竞争，而不是用一个最有故事性的原因占满画面。

反馈的符号也可能随状态而改变。某种氧化过程初期可能增加可用底物并促进特定活动，随后底物耗尽或生态竞争改变，又可能限制同一活动。把“生命—环境反馈”永远写成促进关系，会忽略抑制、耗竭与转向。共同塑造包含互相限制，并不天然等于共同增长。

共同塑造包含这些促进、抑制与转向。局地生物的扩展可以改变其他生物的底物与生存条件，氧化环境扩大也可能压缩某些还原环境。长期状态由这些不对称作用共同维持，未必表现为各部分同步增长。把分化保留在氧化史中，才能解释为何同一次环境改变在不同生态与化学过程里具有不同后果，也避免把系统转折写成预定通向现代生命的阶梯。

从知识生成角度看，多同位素研究还改变了“生命证据”的门槛。过去可能凭一个同位素范围提出生物解释，增加大气机制后，同一范围需要接受更强竞争检验。这是一种知识的重组：旧数据没有消失，但它能够支持的主张缩小或改变。SDE 所谓回写，在此可以具体表现为新关系重新规定旧证据的发言资格。

对生命—地球共同塑造的研究还应避免目的论。氧化环境后来支持某些复杂生命，并不意味着早期过程以产生这些生命为目的。历史结果可以由当时局部过程与约束累积形成。把现代适宜性倒投为古代目标，会使自然变化看似沿着预设路线前进，遮蔽真实的分支、耗竭与偶然条件。

同样，今天的硫循环不能作为早期地球的完整模板。现代反应提供机制参照，但早期大气、海洋与生物条件不同，过程权重可能变化。思想史中真正的突破常来自发现现代直觉不再足够，而不是把今天的系统缩小后放回过去。

发生学由此能够提出一个严格问题：什么条件使某种循环结构开始能够维持，什么条件又使它失去稳定？这比简单问某种物质何时出现更深入，但仍必须回到样品、年代与收支。只有可观察差异能够区分候选历史时，发生解释才超出概念可能性。

多同位素关系使早期地球史增加了一种比较尺度。岩石中的细微组成差异，经由反应、保存与年代约束，能够排除一部分关于大气与硫循环的候选历史；生命的具体贡献则还要放进源汇关系中判断。研究对象因此扩展了，证据责任也分得更细。下一章将考察这种跨圈层研究怎样进一步进入统一计划：一个关于整个地球的纲领，如何转化为实际测量、模型接口和共同体分工？

第 50 章 1988 年的地球系统纲领
一张整体图怎样组织研究

50.1 纲领写作的历史现场

河流把陆地与海洋连接，水汽把水量与能量连接，硫循环把岩石、大气与生命连接。这些联系已经在不同研究中取得具体形式，却未必共享观测尺度、资料标准与计算接口。地球系统科学面对的任务，是使它们进入共同问题、持续观测和可运行模型：一个领域给出的结果怎样成为另一个领域可使用的约束，双方又怎样检查连接之后产生的整体变化？

1988 年的《地球系统科学：更近的观察》为这一转向提供了正式纲领节点。[C08] 它属于 NASA 顾问体系下地球系统科学委员会的报告背景。这里应将科学联系已经存在、统一纲领被提出、具体设施实际建成三种时间区分。报告发表不是所有跨圈层知识的诞生日，也不等于报告中的计划当年已经实现。

另一份 NASA 档案《地球系统科学：全球变化计划》登记出版于 1989 年，其摘要明确列出长期全球测量、地球历史与过程研究、系统模型、信息系统、机构协调和国际合作等内容。[C13] 它与 1988 年报告相关，但不是同一题录。把两份文件并列，可以看到纲领关心的不仅是理论，还包括使理论能够持续工作的基础设施。

整体图示在这种历史环境中具有特殊作用。它把学科内部熟悉的变量放到共同空间，让原本处于边缘的交换关系成为显眼问题。一位海洋研究者的边界条件，可能正是大气研究者需要解释的输出；一个生态过程的输入，可能来自地球化学循环。图示通过改变问题的排列方式，使接口从附属条件变成研究中心。

但图示也可能制造完成感。每个方框都画上，每条箭头都连通，读者容易以为一个整体模型已经存在。实际上，箭头可能只是已知联系，也可能是尚不能量化的假说。要从图示走向研究计划，需要为它指定变量、单位、测量方法、时间尺度和负责提供资料的机构。没有这些规定，整体图仍只是组织思想的工具。

这种转向还改变了成功标准。单项研究可以在局部变量上取得高精度，系统研究却必须检查不同部分拼接之后是否守恒、是否存在重复计算、是否隐藏相互矛盾的边界条件。整体不是各部分准确率的简单总和。一套大气模型与一套海洋模型分别表现良好，连接后仍可能产生漂移，因为两边对同一界面使用了不同定义。

因此，地球系统纲领的历史新意需要在两个层次评价。科学层次看是否形成了新的可检验问题，组织层次看是否建立持续获取和交换证据的能力。二者相关但不相同：机构建立不自动带来新发现，科学问题清楚也不保证资源与协作能够到位。把两层都写入思想史，才能解释为什么某些重要问题长期停留在愿景，而另一些成为常规研究。

纲领还改变研究时间的组织。一次航次或一篇论文通常有相对明确的结束点，全球变化研究却要求跨越项目周期的持续记录。短期资金安排与长期科学问题之间由此出现张力。要维护基线，必须有人承担看似没有即时新发现的连续观测；没有基线，后来显著变化又难以获得可信参照。

这使“全球”不仅是空间覆盖，也是一种时间承诺。观测一次整颗地球，可以得到全球图像；持续以可比较方式观测，才可能研究变化。两者在新闻传播中容易混同，在科学上却需要不同设施与责任。系统纲领的实质，应在是否建立这种持续能力上评价。

全球视野也不能抹除地方知识。卫星看到广域特征，地面调查可以辨认具体过程与误差来源。若只追求统一大图，局地异常可能被平滑；若只保留地方经验，又难以判断共同变化。系统科学需要的是两种尺度之间可检验的往返，而不是一种尺度独占解释权。

◆ 50.2 接口如何改变分工

大气与海洋相接的海面，是理解系统接口的一个清楚例子。大气需要海表温度、湿度交换、动量与热量通量等边界信息，海洋需要来自大气的风应力、能量与淡水交换。两边若只交换一个温度数字，许多反馈仍被切断；若交换所有变量但定义不一致，系统又可能在连接处产生虚假源汇。

首先要统一方向。大气向下的热通量，在海洋预算中应是输入；若两套程序采用不同正负号而没有转换，就可能把相同过程同时记为双方损失，或同时记为双方收益。错误并

不来自复杂物理，而来自接口约定。守恒检验能够让这种错误暴露，却要求研究者保留每一侧的独立账目。

其次是单位和面积。一个模型输出单位面积通量，另一个需要网格总量，就必须乘以相应面积并处理海陆覆盖。沿海网格尤其敏感：大气格点可能同时覆盖陆地与海洋，海洋格点则只覆盖水域。简单平均或插值若没有守恒要求，可能凭空增加或减少总交换。更平滑的图像并不保证更正确的预算。

第三是时间步长。大气过程与海洋过程使用不同计算步长时，交换量需要在一定间隔内积累或平均。若把瞬时值当作时段平均，快速变化可能引起误差；若交换过于稀疏，反馈又可能被延迟。时间接口不是纯软件选择，它参与决定耦合系统能够表达哪些变化。

第四是变量身份。海表温度可以指极薄表层，也可以指模型最上层的平均温度；风应力可以来自不同高度风速与交换公式；淡水通量可能包括降水、蒸发、河流与冰过程的不同组合。名称相同并不保证含义相同。每项交换量都应有明确元数据，说明它如何形成、覆盖何处和包含哪些过程。

可以用两个储库的教学模型说明接口守恒。若储库甲向乙输送通量 J，甲的方程含 $-J$，乙的方程含 $+J$，合并后内部交换消去。若甲使用 J_1 而乙使用 J_2，总系统就出现 J_2-J_1 的虚假源汇。这个简单代数说明，整体守恒依赖接口一致，而不是各部分分别写出了守恒方程便自动成立。

模型连接还会暴露参数补偿。一个独立海洋模型可能利用经验校正弥补边界输入偏差；与大气模型耦合后，原校正可能重复处理同一误差，或阻断本应出现的反馈。于是，耦合不只是把程序接线，还可能要求重新检验各部分过去成立的假设。原来被当作外部条件的变量，现在成为系统内部需要解释的状态。

组织分工也随之改变。谁负责通量定义，谁负责重网格，谁检查守恒，谁保存版本，必须明确。否则接口错误很容易成为双方都认为属于对方的问题。一个真正可运行的共同对象，需要科学责任与技术责任在边界上相接。协作不是简单增加会议，而是使相互依赖的输出有明确交付标准。

接口使各部分取得共同对象的身份。通量定义、尺度转换与守恒约束构成 E，跨模型交换、比较和诊断中形成的有序差异构成 D，由此形成的 S 才是能够共同运行和检验的系

统状态。例如，原先外置的海表温度进入双向耦合后，会随海洋热收支变化，并继续改变大气交换；它从给定边界变成内部响应变量。对象获得的新属性因此来自关系闭合，必须在接口一致的条件下通过对照试验辨认。

空间重映射可以进一步展示接口的困难。一个大气网格覆盖几个海洋网格时，交换量需要分配；反向汇总时，又要保证总量一致。若采用普通点值插值，局地数值可能平滑，却未必保持面积积分。守恒重映射关心的不只是每点相近，而是整体交换账目不因网格变化而改写。数值方法因此直接进入物理对象的成立条件。

陆面—大气接口具有类似问题，但还多出植被与土壤状态。大气提供降水和辐射，陆面计算水分与能量分配，再反馈温度、湿度与通量。若一边把降水当作均匀时段输入，另一边对雨强阈值敏感，平均化会改变产流与蒸发。时间分辨率不匹配可以在没有任何编码错误时产生实质过程差异。

接口还需要处理不同模型的有效尺度。某个参数化代表格网内部未解析过程，另一个模型可能已经显式解析其中一部分。耦合后若两边都计算同一效应，就会重复；若双方都以为对方负责，则会遗漏。职责表因此不仅是项目管理文件，也是模型结构的科学说明。

对诊断而言，可以先关闭部分反馈进行对照。单向驱动试验保留一侧影响，双向耦合试验允许回返，二者差异帮助识别反馈贡献。但这种差异仍要在兼容初始与边界条件下比较，不能把其他变化一起归入耦合效应。整体研究需要受控拆分，拆分不是背叛整体，而是使整体作用可以辨认。

模型漂移则是另一种系统性信号。没有外部变化时，长时间运行仍出现持续偏移，可能提示内部不平衡、参数化误差或未充分调整的初始状态。漂移不能一概当作真实自然趋势，也不能简单从结果中扣除后不再追究。它指出某些接口或内部关系尚未共同维持预期状态。

接口的质量还可以通过封闭试验检验。暂时取消外部源汇，让两个子系统只交换已知量，观察总质量或总能量是否保持；再逐一打开外部输入，检查整体变化是否等于输入积分。这样的试验没有现实全部复杂性，却能隔离连接错误。它展示了系统模型需要专门的耦合验证，不能只依赖各子模型原有测试。

另一个有效诊断是追踪守恒量在各尺度之间的去向。如果某个全球预算闭合，而区域预算存在大误差，可能是相互抵消；若局地交换合理而全球出现偏移，可能是重映射或边界遗漏。整体与局部应来回核查，单一尺度上的成功不能解除其他尺度的责任。

◆ 50.3 计划、设施与知识生产

系统研究需要连续记录，不能只依靠偶然成功的观测。一颗卫星、一艘调查船、一座测站，各自能够看到不同差异；长期气候问题还要求它们之间保持可比性。计划因此必须把仪器寿命、交叉标定、资料保存和后续接续纳入知识生产，而不是把这些工作留给发现之后的管理。

NASA 的 Terra 任务提供了后续设施的具体例子。官方资料记载，它于 1999 年 12 月 18 日发射，携带五台仪器，对大气、海洋、陆地、冰雪与能量收支提供相互关联的观测。[R50a] 多变量联合观测由此具有明确的平台形态。评价它与系统纲领的关系，需要继续区分平台能够观测什么、实际形成了哪些持续记录，以及这些记录支持了哪些系统问题；发射是这条建设链上的一个节点。

同一平台上的多仪器具有时间与空间配合的优势，却也带来共同误差。轨道、定位、平台状态或共同处理可能使多个产品彼此依赖。联合观测能够增强解释，但不能把五台仪器简单当作五份完全独立的证明。系统科学既增加了连接，也增加了理解证据依赖的责任。

资料系统同样不是附属仓库。NASA 关于 HDF-EOS 的历史说明指出，EOS 资料通过分布式档案中心组织，并采用共同数据格式以应对异质产品的交换问题。[R50b] 这类工作使数据能够被查找、读取和组合。若每台仪器都使用无法互通的格式，多圈层研究会在进入物理问题之前就耗尽精力。

然而，共同文件格式只解决一层接口。两个文件即使都能被同一软件打开，变量定义、质量标记与时间标准仍可能不同。可读性不等于可比性。真正的数据共同体需要同时保存算法版本、标定信息、缺测规则与不确定性，使使用者知道一个数值怎样成为这个数值。

长期记录尤其需要区分仪器变化与地球变化。旧传感器老化，新传感器接替，观测时刻或轨道条件变化，都可能在序列中留下断点。若不进行交叉检查，一个技术变化可能被误读成气候变化。反之，过度调整也可能削弱真实信号。记录连续性是一项持续研究，不是把若干文件按日期拼接即可获得。

“承诺—项目—记录—成果”可以作为考察计划实现的四层链。承诺说明要解决什么问题，项目说明建设什么能力，记录说明实际获得什么资料，成果则说明资料支持了什么结论。每一层都可能与前一层不同：预算与技术改变会调整项目，仪器性能会限制记录，科学分析又可能发现最初没有预见的问题。这种变化不是计划失败的同义词，但必须如实记录。

设施运行还涉及时间上的不对称。建设成本集中发生，知识收益可能多年后才显露；某些记录在当时没有明显用途，却在新的问题中成为不可替代的基线。系统研究因此需要保留尚未被当前模型使用的资料。若只保存经过某种解释处理的最终产品，未来研究者可能无法重新检查旧假设。

这也改变了科学家的工作形象。思想史常突出提出理论的人，而连续标定、资料管理和接口维护容易被视为辅助劳动。但在长期地球系统研究中，没有这些工作，理论比较所依赖的对象就无法稳定存在。新知识的发生既来自概念突破，也来自共同测量世界的持续维护。

1988 年报告与后续设施在问题上的对应，还应与机构历史上的因果联系区分。项目立项、技术论证和资源配置记录，才能说明某项能力怎样被选中，又怎样在实施中改变。一个符合早期愿景的平台可能同时继承多条研究线索。沿着这些档案辨认汇合点，可以理解纲领如何协调既有工作，而无需把所有后续成果归入单一报告的直接影响。

数据中心的作用还包括保存可追溯版本。一个卫星产品经过新算法重新处理，旧文件并不一定变成无意义垃圾：已发表研究可能依赖旧版本，比较新旧结果需要知道改变发生在哪里。版本管理让知识能够修正而不失去历史。没有它，用户可能以为同名产品始终代表同一种观测。

共同体因此需要区分原始观测、校准资料、反演产品与综合重建。越往后，产品越接近研究者直接需要的变量，同时也包含更多模型假设。便利性增加，不应使来源链消失。

一个全球地表变量可能经过大气校正、云筛选和空间合成，使用者若只看最终数值，就难以判断某种异常来自地表还是处理条件。

长期设施还会产生意外发现。一个为某类目标设计的仪器，可能为另一类问题提供关键线索；旧资料在新的理论下被重新处理，也可能显露此前没有被定义的特征。这意味着计划必须既有明确任务，又保留可重用原始信息。过早只保存任务专用结论，会关闭未来知识发生的路线。

设施的不同成绩还需要不同评价时距。入轨与仪器运行可以较早判断，跨传感器的连续性需要重叠观测和长期标定，某类慢变化能否被识别则要等待记录覆盖相应尺度。若用建设完成的日期替代知识成熟的日期，就会把这些时间压在一起。计划史应同时记录能力何时建立、资料何时可比，以及某项科学结论何时得到足够支持。

制度与技术之间还存在双向作用。共同资料规范能降低合作成本，合作中发现的新需求又会推动规范修订。若修订仅为方便某一部门而削弱其他研究用途，统一反而可能造成新的障碍。好的系统建设应让变更有公开理由、有映射关系，并保留对旧成果的追溯能力。

数据开放也改变了错误发现的分布。资料能够被不同团队重用，便有机会在原任务之外发现系统偏差或新用途。与此同时，使用者可能不熟悉原仪器限制，因此文档、质量标识与版本通知成为开放的必要组成。只把文件放到网上，并不等于形成了可负责使用的公共知识。

长期档案还要求保留失败记录。某些时段因云、仪器状态或轨道条件无法提供可靠观测，缺测标记应保持明确，而不是被插值掩盖。插值产品可以服务连续应用，但应与直接观测区分。若缺失与天气条件有关，缺测本身也可能带有选择偏差，不能把它当作完全随机。

这类细节使系统科学的制度史与认识史紧密连接。谁决定质量门槛，谁发布算法修订，谁承担长期维护，都会影响哪些差异能够进入共同研究。制度不只是科学之外的经费背景，它通过具体操作改变知识对象的稳定性与可修订性。

同一档案还可以承担教学与研究两种用途。教学产品追求易读，研究资料强调细节可追溯；两者可以共享入口，却应保留层级标识。否则便利的图像被误当原始观测，系统知识在传播中会失去其形成条件，也失去接受重新检验的具体路径。

◆ 50.4 系统科学的已有成绩与边界

从模型接口到长期档案，地球系统科学的整体观已经落实为多种研究操作：耦合检验连接两侧是否守恒，联合观测辨认跨圈层响应，资料系统保存这些比较所需的身份与版本。评价其历史成绩，应看这些操作使哪些原来孤立的问题变得可检验。相互作用、反馈与尺度也因此获得了不同技术含义，不能只用一张整体图概括全部推进。

共同变量的形成尤其值得追踪。不同机构最初可能用不同空间范围、平均窗口或处理等级表达同名量，后来通过校准与接口规范取得比较资格；模型检验又可能发现原先的统一掩盖了某类差异，需要重新分层。这个过程连起对象身份、操作程序与研究能力。具体何时形成共识、何种失配促成修订，仍需版本与档案确定，不能仅从最后一致的产品反推一条必然路线。

系统模型仍然不等于完整地球。任何模型都选择有限过程、分辨率与参数化，某些低频或小尺度变化只能近似表达。模型能够再现多个关键特征，不意味着所有未来状态都已被掌握。相反，一个成熟系统应当能够指出自身在哪些变量与尺度上更有约束，在哪些地方主要依赖假设。

复杂系统也可能因接口过多而难以诊断。一个误差在多圈层反馈中传播，最终在远离来源的位置显露。仅看最终输出难以定位原因，必须保留分层诊断与简化试验。整体研究与局部研究因此不是谁取代谁的关系：整体暴露相互作用，局部实验与理论帮助辨认作用链。

组织整合也存在边界。统一标准便利资料交换，但过早统一可能压缩尚未理解的差异。不同地区或学科使用不同变量，有时并非落后，而是面对不同对象。好的规范应允许映射与版本更新，而不是把所有记录强迫成一种格式后便宣称已经统一世界。

系统知识的统一性首先表现在交换量可以核算、时间可以追溯、变量可以对应；差异性则保留圈层机制、有效尺度与测量方式。二者通过接口获得协调：海洋不必采用大气的

全部网格与步长，但双方应对同一交换量承担兼容定义。这样的和谐是一项技术与组织成果，既允许局部研究保持自己的问题，也使其输出能够接受其他部分的限制。

整体系统的一个失败条件，是不同组成部分只能靠不断调整无法解释的补偿项维持一致。补偿在某些研究阶段可以作为诊断工具，但若没有界限地使用，系统就可能失去反证能力。任何观测都能通过新的补偿吸收，整体图越大，真实约束反而越少。复杂度必须与可辨认性共同增长。

另一个边界来自问题范围。研究气候变化的模型未必充分表达地震或地幔过程，研究深时地球演化的模型也未必能提供短期天气预报。名称中都有“地球系统”，不意味着对象尺度相同。系统科学不是一套模型包办一切，而是在明确时空范围内选择必须耦合的过程。

同样的分工也适用于跨案例比较。流域连通决定哪些输入属于同一控制体，海洋动力限定输运，代理校准限定从痕迹到气候的反推，大气反馈则依赖受控响应。它们可以共同接受边界、证据和修订条件的追问，却不能共用一套具体机制解释。地球系统的综合能力，正在于把这些不同结果接入适当位置，而非要求每个学科以同样方式生成知识。

地球系统的整体性还需要保留问题优先级。并非所有连接对每一项任务同样重要，研究计划必须判断哪些缺口最限制解释，哪些观测能够最大幅度减少竞争机制。平均向所有方向扩张，可能造成资源分散；只追逐已有强项，又可能让关键接口长期空缺。系统观念的实际价值，是帮助确定这些取舍，而不是取消取舍。

因此，一张整体图的成熟形式应当允许标出已知连接、弱约束连接与暂未纳入连接。知识边界可见，研究共同体才能据此协作。把未知藏在密集箭头中，会让图显得完整，却削弱它组织下一步研究的能力。

1988 年的纲领因而可以被理解为一个组织研究的历史节点，而非地球科学突然拥有全知视角的时刻。它把分散的联系转化为持续建设的任务，也让资料和制度成为思想史的一部分。下一章将把目光集中到遥感档案：当传感器与算法不断更新，同一个地表如何通过反复标定，保持足够稳定而又能够修订的科学身份？

第 11 编

遥感算法与行动中的地球知识

第 51 章 同一块地表如何跨越半个世纪的卫星记录

第 52 章 Destination Earth：数字孪生究竟复制了什么

第 53 章 AI 标出的地震，何时成为共同体承认的事件

第 54 章 雷道布特火山与航空预警：知识怎样进入行动

第 55 章 版本也是证据：从 Landsat 重处理到 AI 目录更新

第 51 章 同一块地表如何跨越半个世纪的卫星记录

51.1 像元取得位置与数值

一幅卫星影像很容易使人产生直接观看的感觉：蓝色是水，绿色是植被，白色是云，地表仿佛把自己的模样交给了机器。然而，真正能够支持长期地表研究的影像，并不止于一个可以观看的画面。它必须使今天的研究者知道，某一个数值由什么传感器、在怎样的观测条件下、针对哪一片空间范围产生。颜色可以随显示方案改变，科学对象的身份却不能随配色任意改变。影像的思想史，因而也是数值怎样取得空间位置和比较资格的历史。

传感器接收的是进入其响应范围的辐射。一个波段并非只对单一波长作出反应，而是在一定范围内以特定的响应函数加权。即使两颗卫星都提供被称为红光或近红外的波段，其响应范围也未必完全一致。当地表由水、裸土和植被组成时，这种差别可能影响由波段计算出来的指数。传感器的读数还要经过标定，才能与辐亮度或反射率等物理量建立规定的对应。数字相同，不足以证明物理量相同；数字不同，也不足以证明地表已经变化。

光从地表进入探测器，还要经过大气。气溶胶、云和水汽会改变传输。用于大气顶层描述的产品，与经过大气校正、指向地表反射特征的产品，不能只因名称中都有反射率便直接混合。处理等级不是文件管理中的装饰，而是告诉研究者，哪些影响已经按给定模型处理，哪些仍留在测量链中。研究水体面积、植被变化与地表温度，也不能把各自需要的处理条件合并成一种笼统的“清晰影像”。

位置同样不是相机自动赠送的确定性。卫星轨道、姿态、地形与投影共同影响像元怎样落到地图上。山坡上的一个亮斑若被配准到相邻谷地，前后两期影像相减就可能出现并不存在的变化。重采样还会改变一个输出像元综合哪些输入信号。因此，坐标一致只是比较的必要条件；有效空间分辨率、地形影响和重采样方法也会决定比较究竟发生在哪个尺度。

这里的像元不是天然地块的小型复制品。它是一种通过规定的响应、定位与处理操作建立的空间测量单位。同一像元中可能包含多种地物，像元边界也未必与田块、岸线或林缘重合。研究者把像元判为某种土地覆盖，是在测量之外增加了分类判断。把分类结果当成原始观测，会抹去这一判断发生的位置，也会使后来的误差难以定位。

长期研究至少要保留传感器、获取时间、场景标识、处理等级、产品版本、坐标参考、标定信息以及质量标识。它们并非可以在正文之外遗忘的附件。失去获取时间，季节变化可能冒充年代趋势；失去产品等级，处理差别可能冒充生态差别；失去版本，后来发现的系统误差就无法回到当初结论。所谓同一地表的连续记录，从开始就是一项需要维护的对应关系。

跨传感器比较还涉及观测几何。太阳照射方向、传感器观看方向与地表结构共同影响接收到的反射信号。对于具有明显方向性结构的植被或粗糙地表，相同材料在不同角度下可能表现不同。因而，所谓同一地物的光谱不是脱离观测互动而固定的一串数。研究者要判断哪些角度差异可以忽略，哪些需要校正，哪些已经超出当前比较的适用范围。

地表温度尤其显示变量身份的重要性。热红外产品所指向的地表辐射温度，与气象站在规定高度测得的空气温度具有不同物理含义。两者可以相关，但不能直接当作同一测量互相替换。若研究城市热环境，必须说明关心的是表面热状态、人体所处空气条件，还是某种暴露指标。遥感图上的高温斑块能够提出问题，却不能独自完成健康或气象意义上的判断。

时间合成也是一次对象转换。把一个月多景影像合成，可以减少云遮挡影响，但最终图像未必代表某个真实同时刻的完整地表。如果不同位置采用了不同日期的晴空像元，空间拼图中已经包含时间差。对缓慢变化的植被可能尚可接受，对迅速移动的洪水边界则可能构成严重混淆。合成方法必须与过程速度一起评价。

◆ 51.2 Collection 2 改变了什么

美国地质调查局在 2020 年 12 月公开提供 Landsat Collection 2，构成 Landsat 档案处理史中的明确节点。官方说明把它列为档案的再次大规模重处理，涉及几何精度、辐射标定、质量信息以及产品组织的改进。[C09][C14]这一事件并不意味着卫星重新飞过了

过去的地表，而是意味着保存的观测资料在新的参考与处理规则下获得了新的产品形态。观测发生时间与产品形成时间，在这里明显分离。

这一分离有重要的认识论后果。档案不是把过去永久封存之后就不再变化的仓库。新的控制资料、标定知识和算法，可以重新规定旧测量怎样与地表建立对应。产品值被修订，可能是对旧误差的修正，也可能来自产品定义、缩放方法或处理流程的调整。研究者必须询问改变发生在哪一层，不能一见新版本就认定过去的科学结论全部失效，也不能认为原始观测未变便无需重新评估衍生成果。

几何改进的意义，可以从边界对象理解。广阔均质区域中的轻微位移，未必显著改变区域平均值；狭窄河流、破碎林斑或城乡交界处的相同位移，却可能改变像元分类。这表明“更准确”不是一个脱离任务的单一形容词。定位质量必须与对象尺度相比较，才知道改进是否足以影响研究结论。一个适合大区域统计的产品，不因同属一个 Collection，就自动适合每条小河岸线的精细变化分析。

辐射标定的作用也要沿用途追查。跨年代比较通常关心信号的相对稳定，而不是单期画面的视觉悦目。如果一台仪器的响应随时间改变，未经适当处理的序列可能显示缓慢趋势。新的标定关系把部分趋势重新解释为测量系统的变化，这不是取消地表研究，而是使地表过程与仪器过程能够更明确地区分。真正有用的连续性，并非要求所有数值不变，而是要求数值改变能够被解释。

质量标识使另一个长期被图像直觉遮蔽的问题显露出来：没有被判为云的像元，不等于一定没有受云或阴影影响；存在质量字段，也不等于使用者已经正确解码。质量掩膜是一套规则化判断，它决定哪些像元进入统计。若不保存筛选方式，同一套公开影像仍可能支持不同的时间序列。一个团队只保留较严格的晴空观测，另一个团队保留更多边缘样本，差异可能来自筛选而非土地覆盖。

这些变化需要按任务分别评价。对河岸定位，验证材料应能分辨影像位移与实际岸线移动；对区域植被趋势，则应检查标定修订是否改变跨传感器的相对稳定性。官方发布说明记录了重处理的范围与目标，具体场景的改善幅度还取决于这类对照。档案生产者由此承担了一项持续责任：既提供共同产品，也让使用者能够识别哪些差异已获处理、哪些仍需局地验证。可比性在发布之后仍须经受使用中的检验。

这也改变了“统一档案”的含义。统一不是让各代仪器具有同样的物理响应，更不是擦去它们的历史差异。统一是在明确差异仍然存在的前提下，提供可以追踪和适当处理这些差异的共同规则。标准化因此既减少使用障碍，又不能替代研究者对具体波段、季节和地物条件的判断。可比较性是一项有边界的成就，而不是从此不再需要检验的属性。

还有一种看似细小的差别，是缺测怎样进入统计。研究者若只比较每年可用像元的平均值，某一年因云覆盖而缺少山区观测，平均值变化就可能部分来自样本空间组成改变。保留固定研究范围并不等于每年实际参与计算的区域相同。应同时报告覆盖率与有效样本分布，必要时限定共同可观测区域，或者明确建模缺测机制。

这些操作说明，长期连续性不是消灭所有差异，而是把差异分为可以校准、需要限制和暂时无法处理的部分。跨年代影像真正成为研究对象，是因为使用者能够对这些部分作出有记录的判断。一个不记录筛选与缺测的长序列，可能比一个较短但条件清楚的序列更难解释。

◆ 51.3 改变地表还是改变处理

假设研究者在两幅影像之间发现一片湿地面积减少。最直接的解释是地表发生了变化，但这只是候选解释。两个时点的季节是否相当，水位是否受短期天气影响，所用产品是否处在同一处理体系，质量筛选是否一致，都关系到“减少”究竟指什么。若这些条件没有整理，面积差值可以算得非常精确，所比较的对象却仍然含混。

一种清楚的研究设计，是把时间与版本分成两个比较方向。先固定原始获取时点，用旧处理与新处理比较同一场景，识别版本变化会产生何种空间模式；再固定处理版本，用不同时点比较地表候选变化。前一个方向控制自然时间，后一个方向控制处理规则。两者不能穷尽全部误差，却能阻止最基本的混淆：把档案重处理引起的数值修订算成自然变化。

这种交叉对照还可以发现交互作用。某种处理修订可能对平坦裸地影响很小，对山地阴影或浑浊水体影响较大。于是，不同时间的地表状态又会改变版本差异的大小。不能简单地用一个全景平均修正量消除所有偏差。研究者需要查看误差是否随坡向、水体浑浊程

度或观测几何而变。只有当误差结构进入分析，才能知道某种趋势是否超出了处理敏感性。

下面以湖岸季节性水域作研究设计示例。先选定公开可获取的同一轨道覆盖区域，保存获取日期和场景标识，再明确比较的是固定季节水域面积、年内最大范围，还是长期岸线迁移。三种对象需要不同的样本安排。固定季节比较应控制观测日期，年度最大范围需要说明全年覆盖，岸线迁移则要检查水位与定位误差。若从固定日期面积改成年度最大范围，单是影像数量增加就可能提高观测到的最大面积；这一采样效应应在解释自然变化之前估计。

随后应固定水体分类方法，同时保存连续光谱指标和最终类别。只保存二值地图会丢掉阈值附近的变化信息。若新旧产品主要使一批边缘像元跨过阈值，面积变化可能十分明显，但物理信号变化并不一定同样巨大。连续指标、阈值选择与空间分布结合起来，才能区分稳健变化与分类敏感性。分类器的决定，需要有机会被原始特征重新审查。

独立材料的作用，也要按其时空范围界定。现场水位站可以检验某个湖泊的时间变化，却不直接确认每一条岸线；高分辨率影像可以帮助核查边界，却可能与 Landsat 获取时间不同；历史地图具有价值，但其测绘尺度与制图目的未必相同。独立证据不是任何不同来源材料的集合，而是能够对当前争议环节施加另外约束的资料。

这个设计允许出现负结果。如果候选变化在固定版本后消失，就应把原先的地表解释撤回；如果变化只在一种阈值下存在，就应降低稳健性判断；如果多个独立约束都支持变化，才有理由继续讨论水文机制或土地利用。即便确认面积发生改变，也还不能由影像单独判定原因。降水、引水、季节和工程活动之间的因果分辨，需要另行组织证据。

还有一个尺度检验可以补充上述对照：把候选变化分别放在单像元、局部区域与整体范围内观察。若变化只在边缘像元存在，却在合理聚合后消失，应优先检查配准与分类敏感性；若多个尺度都保留一致方向，才更值得追踪自然过程。尺度检验不能替代独立资料，却能帮助选择最需要复核的位置，使有限的现场调查更有针对性。

51.4 新的地表连续性

遥感档案把长时间、广空间的地表比较变成可能，但这个新对象不是把原有地表照片按年份排成一列。它是带着共同定位、物理量定义、质量判断和版本记录的测量系统。研究者因而能够提出过去难以实施的问题：变化何时开始，是否跨地区同步，局部异常是否属于更广的模式。新的辨认能力、处理操作与地表特征的可比较规定，在这里共同发生。

用 SDE 语言说，跨传感器地表时序是显露 S；它依赖对波段、时间、位置和版本差异的组织 D，也依赖地表反射、大气传输、仪器响应与地图定位相互规定的特征纠缠 E。E 不是把四个因素列在一起。例如，地物光谱与波段响应共同规定所测数值，大气状态又改变这种对应，而位置误差决定该数值究竟被归给哪片地表。改变其中一项，其他特征对结果的意义也会改变。

时序对象一旦成立，又会反过来选择操作。对缓慢趋势的研究要求长期标定稳定，对灾变的研究要求足够密的时序与事件前后覆盖，对狭窄边界的研究要求更严格的位置核查。研究目标不是最后附加的用途，而是在选择哪些影像、接受何种质量和保留哪些元数据时，参与规定实际比较对象。发现系统性偏差之后，标定或处理规则再被修订，才构成能够追踪的回写。

回写范围取决于误差结构。若偏差集中在山地阴影，修订可能只涉及特定地形下的筛选；若多代传感器之间出现稳定偏差，则可能要求重新处理较长序列。前者未必改变区域趋势，后者却可能影响趋势起点和幅度。误差修正、能力扩展与对象重建因此可以沿同一份差异记录区分：检查哪些像元受到影响、哪些统计随之改变、哪些原来可比较的时段需要重新划界。

给一个长期地表对象制作身份说明，最少应回答：它覆盖哪些年代，使用哪些传感器与产品等级，在哪种尺度和筛选条件下可比较，哪些变化会触发重新处理或撤回结论。这份说明不是阻止使用，而是使他人能够接续使用。由此形成的连续性，不是没有裂缝的地球图像，而是能够说明裂缝、检查裂缝并在必要时重新建立对应的知识关系。下一章把这样的资料接入数字孪生，也必须继承这些边界。

第 52 章 Destination Earth 数字孪生究竟复制了什么

52.1 从项目承诺读出对象

如果一张影像可以使人误以为自己直接看到了地表，一个能够旋转、放大并显示未来变化的数字地球，就更容易令人相信自然世界已经获得了完整替身。数字孪生这一名称尤其具有吸引力：孪生似乎意味着对应完整、同步可靠，并且可以在计算机中试做现实不能试做的事。但名称并不规定科学能力。要判断一个平台是什么，首先要问它承担什么任务、包含哪些状态变量，以及怎样把模拟结果与观测和用途连接。

Destination Earth 为这一问题提供了具体研究对象。其气候变化适应数字孪生旨在提供具有全球一致性的高分辨率气候及影响领域信息，项目说明涉及多年代际模拟、应用需求和共同设计。[C10]本书核读的官方资料报告，项目于 2026 年 6 月进入第三阶段。[R52a]这一节点说明工程已经进入持续建设与交付过程。对其成果的判断则须落实到具体产品：已有模拟资料的交付、应用接口的建立、适应措施的实际效果，分别需要数据说明、使用记录和实施后评价，三类证据不能共用一项项目进度声明。

第二代气候数字孪生数据说明提供了更具体的边界：新一组模拟自 2026 年 5 月提供，所列时间范围为 1990 年 1 月至 2049 年 12 月，数据类型是全球气候模拟。[R52b]这是一项具有明确时间范围和产品身份的交付。它支持我们说平台已有可供研究的模拟资料，不支持我们说平台准确预报了未来每一天的天气。模拟时间轴覆盖未来，与获得未来事实不是一回事。

范围审查可以从四种不同对象展开。大气或海洋状态变量，是模型内部运算和输出的对象；极端天气事件，是在特定空间与时间条件下组织这些变量所得的对象；某种能源、洪水或热暴露指标，是经过影响模型与需求定义加工的对象；一项适应政策是否有效，则还包含实施、成本和人类行为。前一层能够为后一层提供约束，却不能自动承担后一层全部证据责任。

因此，平台所声称模拟的，是在选定方程、分辨率、边界条件和情景下的地球系统行为。它没有因为名称而包含每个街区的全部工程设施、每个家庭的行动选择，或每一种生态反馈。某些过程能够显式表达，某些需要参数化，某些暂未进入任务。范围表最有用的内容，往往是明确写出这些差别。一个认真限制边界的平台，比一个宣称什么都已复制的平台，更容易成为可检验的科学对象。

同样需要区分的是项目整体与具体产品。平台可能包含多种模型、试验和访问方式；研究者不能只引用项目名称，就让他人复原自己使用的数据。产品代际、运行配置、情景、变量与时间范围应落到明确的记录。数字地球越具有整体外观，越需要在内部保存这些可区分的身份。否则，整体性会从研究能力变成含混的来源。

数字孪生还包含三个容易混淆的时间。第一是被模拟系统的时间，例如某个历史时期或未来情景；第二是模型实际运行的时间；第三是数据被修订和交付的时间。一次 2026 年运行可以产生 1990 年代的模拟，也可以产生 2040 年代的情景，两者都不是 2026 年的直接观测。把这三条时间线保留，才能避免把新发布误写为新发生，或把未来标签误写为未来事实。

不同用户对时效的要求也不同。气候适应可能关心多年尺度的分布变化，某次极端事件的准备则需要短时间内更新的信息。一个平台同时服务这些任务，并不意味着所有输出采用同一更新机制。平台架构中的共同接口，应让差异可见，而不是以同一种图标掩盖预测与情景的不同身份。

◆ 52.2 气候情景不等于逐时预报

气候情景与短期天气预报都使用物理模型，但它们向世界提出的问题不同。短期预报关心给定初始状态以后，一次实际大气过程将怎样演化；气候模拟关心在给定外部条件与情景下，状态分布、变率与极端事件特征可能怎样改变。两者都不能脱离初始与边界条件，但其评价重点和可解释时间尺度不同。把长时间模拟理解为延长了的逐日天气预报，会把模型放到错误的检验位置。

想象一组未来夏季的模拟。某一天出现强降水，不等于作者声称现实世界在同一日期必然发生这场雨。研究者可能关心的是强降水频率、持续时间、空间关联或与热浪的共同

出现方式。这些统计对象需要足够样本与明确比较基准。若把模型某个格点的单日输出直接当作某条街道的未来天气，错误不必发生在方程求解中，也可能发生在对象从模拟到使用的转换中。

观测在这些任务中也具有不同角色。它可以用来建立或约束初始状态，可以帮助估计参数，可以检验历史时期模拟，还可以为影响模型提供地面参照。不能把所有这些用途都叫作实时同化。尤其是跨年代模拟，一次观测支持的模型评估，并不意味着未来每个时刻都由新观测同步纠正。使用者应知道观测在哪里进入、多久进入一次、约束了哪些变量，而不是只获得“数据驱动”的笼统保证。

故事线试验进一步说明这一点。官方第三阶段说明提到，在不同气候条件下重新模拟特定热浪或洪水一类事件。[R52a]这类试验能够提出有价值的反事实问题：若某些大尺度条件变化，一个事件的物理表现可能如何不同。但它的答案取决于被保持和被改变的条件。把事件放入较暖背景，并不自动包括所有未来社会变化，也不等于计算事件在现实未来发生的全部概率。

反事实模拟的解释应从控制条件开始。哪些动力结构被保留，哪些热力条件被调整，模型怎样满足物理一致性，初始与边界之间是否存在不协调，这些问题决定比较的含义。如果控制条件不清楚，两个逼真的动画只能说明计算产生了两种结果，不能说明差异由哪一项改变造成。可视化提高了理解效率，却不减少控制比较的责任。

分辨率也需要拆开理解。网格变细，可以允许某些结构得到更直接表达，但不保证所有过程都随之准确。观测误差、物理参数化、耦合接口和边界输入仍可能限制结果。一项高分辨率输出在空间上很精细，若其系统偏差尚未评估，精细并不等于可信度同幅提高。评价必须针对变量、区域和用途，而不是把网格尺寸单独作为质量判决。

多模型或多成员模拟可以帮助呈现不同不确定性，但成员数量不自动代表所有可能未来。若模型共享关键结构或输入，其误差仍可能相互依赖。情景之间的差别也不是由模型投票决定哪一种必然发生。气候数字对象的可靠性，来自把这些不同层次的不确定性明确保留，而不是通过统一界面把它们压成一个看似精确的未来。

还可以用一个教学反例说明尺度转换的问题。设某个网格平均降水量被输入一条很小的城市排水沟模型，若网格内部降水高度不均，平均值可能无法反映真正触发积水的局部

强度。把网格继续插值成更细的彩色地图，并没有增加原始物理信息。只有引入有根据的降尺度关系、局地观测或更适合的模拟，才能使细尺度对象获得新的约束。

这一区分同样适用于极端概率。一个模型模拟出罕见事件，不等于已经准确估计其重现水平。尾部统计需要足够样本，也需要检查事件生成过程是否可信。若模拟长度有限，估计的不确定性可能很大；若不同试验共享关键结构，增加成员也未必按直觉增加独立信息。数字平台可以使分析更方便，但不能通过界面自动消除抽样限制。

◆ 52.3 应用如何反向规定模型

为了理解应用共同设计，可以选取一个明确标为本书教学设计的任务：某地区希望研究夏季持续高温对电力系统的影响。这个任务不能只向平台索取一幅未来平均气温图。它需要知道连续高温持续多久、夜间是否降温、高温与低风速是否同时出现，以及需求增长和设备运行条件怎样进入影响估计。应用问题由此反过来选择模型变量、输出频率与组合方式。

如果只比较月平均温度，一连串高温日可能被平滑；如果只看单点最高温，区域范围和供需共同变化又可能被忽略。用户所需对象不是原始温度变量本身，而是温度、风、用电需求和系统能力之间具有时间关系的组合。因此，共同设计的实质不是让用户选择图表颜色，而是让使用任务参与规定哪些信息必须保留，哪些统计会损失行动相关特征。

这并不授权用户任意要求模型输出确定答案。一个任务可能超出当前模拟可信尺度。研究团队必须能够说，区域气候资料可以支持某种风险范围，却不足以直接决定单台设备的改造参数。此时的合理回写，可能是增加局地资料或影响模型，也可能是缩小问题，而非无限扩大平台承诺。任务对模型的约束，与模型对任务的限制，应当同时发生。

评价指标也会因用途改变。全球平均误差较小，不代表某地区高温持续时间误差较小；一般天气状态表现良好，不代表极端尾部已被充分检验。对应用而言，最有价值的评价往往是针对会改变决策的特征。例如，若两种适应方案只在连续多日复合事件下表现不同，那么对这种事件结构的检验就比一项笼统平均指标更有解释力。

从评价指标到行动效果，还隔着使用与实施两步。以持续高温影响电力系统的教学任务为例，保存需求版本可以判断指标为何改变，保存模型输出可以恢复当时可选方案，记

录设备或调度措施才能说明输出如何进入行动，实施后的观察才涉及效果。若某项改变止于增加连续高温指标，它改善的是评价设计；若使用者据此改变方案，才多出一段应用关系。这样记录，能够把共同设计的贡献定位到实际完成的环节。

这里还能看到主体能力的变化。使用者需要学习怎样理解情景、误差和条件性；研究者则要理解哪些输出差异在现实任务中具有意义。双方并不是在一个已经完全确定的对象旁边交换信息，而是在共同操作中形成新的可用对象。这个对象既不同于模型内部的状态场，也不同于原先没有气候细节的规划指标。其成立依赖两种辨认能力在接口上的接合。

用 SDE 语言分析，S 是任务限定的模拟—应用对象，D 包含情景、模型、观测与使用方案之间的差异组织，E 则表现为变量定义、尺度和影响关系的相互规定。若热暴露指标要求夜间最低温，输出设计就不能只提供日均值；若模型的可靠尺度限制局地判断，应用单位也必须调整。这些具体的相互限制，才是特征纠缠的内容。只把机构、计算机和用户罗列在 E 中，不能解释对象怎样成立。

共同设计可以把问题从单变量准确转向关系是否足够准确。对于电力任务，温度误差与风速误差分别不大，仍可能因两者同步关系错误而产生较大的供需风险误判。对于洪水任务，降水总量接近观测，也可能因为时序与空间落点错误而得到不同流量。应用对象往往依赖联合结构，因而要求评价从各变量分别合格，推进到变量怎样一起变化。

因而，评价可以具体检查联合特征是否被保留。先分别验证温度与风速的边际分布，再检查二者同现的频率、持续时间和空间覆盖；最后把这些差异传入同一供需估计，比较哪一项真正改变风险判断。若现有评估已包含这些步骤，SDE 在此承担的是跨层说明；若遗漏联合结构，补做检验才可能增加任务所需的信息。特征纠缠在这里落实为可分别检查、又必须联合使用的变量关系。

应用反馈也可能改变数据保存方式。若用户需要研究连续多日过程，平台只保存时间平均就不足以接续分析；若用户关心空间相邻关系，单独导出点值可能丢失关键结构。输出选择因此不是最后一步的格式问题，而会决定未来哪些问题还能被提出。模型生产者与使用者共同保留必要信息，才使数字对象具有持续研究的能力。

在适应任务中，还需要区分情景敏感性与措施有效性。模拟显示某地区在更暖条件下风险上升，可以支持研究适应需求；但哪一种措施最有效，需要把措施怎样改变暴露或脆

弱性纳入模型。仅把未来气候输入保持不变，再比较两个没有明确实施机制的标签，不能形成措施评价。行动变量必须具有可解释的作用路径。

如果两种措施在当前评价下表现接近，合理结果可能是无法区分，而不是强制给出排名。平台可以帮助识别需要补充哪些地方资料，或者在哪些极端条件下差别才显现。一个允许输出未定结论的数字工具，比总能给出唯一推荐的界面更符合科学使用。

这种能力改变了应用者的提问方式：从要求平台告诉自己唯一未来，转向检验哪些决定对多种合理未来仍然适用。数字对象的价值，由此部分体现在帮助保留选择空间，而不是制造超出证据的确定感。

这种提问方式仍然需要外部验证。多种情景下表现一致的方案，也可能共同依赖一个错误模型假设；稳健性必须说明它相对于哪些变化成立，不能被理解为没有条件的可靠。

◆ 52.4 数字对象的失效边界

数字孪生的失败可能发生在不同层。输入错误会让模拟从不适当的条件出发；模型结构不足会使重要过程被遗漏或失真；耦合接口的问题可能破坏变量之间应有的一致；情景被误解，则会把条件性结果当成无条件预言；使用环节还可能把空间尺度不匹配的输出直接套到具体地点。相同的错误图像，背后未必是相同的失败机制。

因此，退出使用条件应随任务规定。若某区域关键变量缺乏必要的验证，便不宜把该输出作为唯一决策依据；若输入版本与评价版本不一致，已有验证不应被自动继承；若结果对合理的情景或参数扰动高度敏感，就应保留多种方案而不是只展示最平滑的一条路径。退出不是否定平台价值，而是防止能力边界被界面越过。

模型与现实的差异也可能成为新的研究对象。当某类复合事件持续偏离观测，问题不一定是再加一层展示，而可能是过程表示、观测误差或尺度转换需要重新审查。只有当这种差异真的改变了模型接口、评价流程或资料安排，才有证据说反馈回写了平台。单纯收集用户意见，或在报告中列出未来计划，还没有完成这条因果链。

同样，不能因为系统规模庞大，就直接认定发生了新的科学对象。如果与既有地球系统模拟相比，改变的只是访问方式和图形界面，那么科学意义可能主要表现为可用性提高。只有当新的任务、持续更新和独立检验关系，使此前不能被稳定提出或检验的问题获

得身份时，才有理由作更强的发生学判断。应让具体记录决定层级，而不是由技术名称预先决定。

Destination Earth 的交付，使模型、观测与应用之间的接口成为可以持续研究的工程对象。判断它的历史意义，可以比较同一任务的两种工作方式：用户直接取得气候场以后自行转换指标，或在共同设计中把所需的持续时间、空间关联与评价条件写入产品要求。区别落在任务信息是否进入生产环节，以及改变以后是否仍可追溯。本章依据产品与项目说明界定这些联系；模拟性能和措施效果，仍须在具体任务中分别检验。

下一章将把尺度缩小到一条地震波形。宏大的数字地球与局部的人工智能识别，看似距离很远，却共享一个根本问题：计算输出怎样取得对象身份。一个模型可以生成丰富状态，一个网络可以标出大量候选；从输出到知识，仍然要经过外部约束、身份确认和允许撤回的程序。

第 53 章 AI 标出的地震，何时成为共同体承认的事件

53.1 从连续波形到候选信号

地震台站不停记录地面运动，但连续记录并不天然分割成一个个清楚的地震。仪器故障、环境噪声、局部活动与真实地震波可以出现在相近的时间窗口。即使信号来自地震，不同震相的起始位置也可能模糊。一个研究者看到波形以后，需要辨认是否存在事件信号、某种震相何时到达，以及这些判断能够承受多大误差。人工智能进入这一过程，首先改变的是辨认与处理的能力。

Mousavi 等人在 2020 年发表的 EQTransformer 研究，把地震检测与 P、S 震相拾取放入一个联合学习任务。论文明确针对单台局地地震记录，并使用注意力机制组织模型。[C11]“检测”回答时间窗口内是否有地震信号，“拾取”则估计特定震相到达的位置。两者相关，却不是同一个输出。一个窗口可能被判断含有地震，但震相起始仍不可靠；一个貌似清楚的局部到时，也可能来自非地震干扰。

网络学习依赖带标签的样本。标签使连续波形与任务定义连接，也把已有辨认实践带入模型。人工拾取不是无误的自然刻度，而是由信号条件、操作规范和专家判断形成的参考。不同信噪比下，起始时间的可辨程度不同；同一套标注准则在不同仪器或地区也可能面临不同困难。因此，模型对标签的拟合能力，不能直接等同于对自然事件全部属性的把握。

信号窗口的安排同样参与对象形成。短窗口有利于聚焦局部变化，却可能丢失较长时序关系；较长窗口保留上下文，又可能包含多个信号。归一化、采样率、通道组合和缺测处理，决定网络实际接收的资料是什么。模型名称描述了一种计算组织方式，但若这些输入条件不清楚，读者仍然不知道模型所说的地震候选具有什么身份。

概率式输出也容易被误解。网络给出的数值是训练任务和模型条件下的评分，不应未经校准就被解释为某一现实事件存在的普遍概率。研究者通常还要设定阈值和窗口规则，

把连续评分转成候选清单。阈值降低可能增加检出，也可能增加误报；阈值提高可能使清单更保守，却遗漏弱事件。模型和决策规则共同规定了最后目录的入口。

模型比较的含义取决于数据划分、阈值与评价方式。固定数值阈值可以比较输出差异，但若不同模型的评分未经共同校准，相同阈值并不代表相同的候选筛选强度；允许各模型选择阈值，则同时比较模型与运行策略。跨区使用还会改变噪声、仪器和地震类型，使原数据集上的结果失去直接外推的条件。EQTransformer 把检测与拾取联系起来，其科学使用价值应沿这两项输出分别追踪：新增了哪些可靠候选，到时误差在哪些信号条件下降低，两者又怎样影响后续关联。

这个阶段产生的是候选信号与候选到时，而不是完整的地震世界。它为后续研究提供了更多入口，却也把错误传播的可能带入后续链条。若把所有候选直接叫作“新发现的地震”，计算能力与事件证据之间的区别就被抹掉了。

联合任务的一个值得分析之处，是检测与拾取之间的约束。只有局部尖峰而缺少合理的后续波形结构，可能提示噪声；具有较长地震信号特征的窗口，又可能帮助辨认较弱震相。模型试图从这些联系中学习，而不是把每个时间点完全孤立判断。这里可以讨论计算结构怎样利用时序，但不应把注意力权重直接解释成完整的地震物理机制。网络内部强调某段输入，与证明那段输入具有唯一因果作用，是不同层次。

标签误差还具有非均匀性。强而清楚的到时容易获得一致判断，弱信号和重叠事件则可能出现较大分歧。若评价只用一个平均误差，容易被大量容易样本主导。按信噪比、距离或事件环境分层，能够看见模型在哪些条件下真正改善、在哪些条件下仍然不稳。科学使用往往最关心边界条件，而不只是总体平均表现。

◆ 53.2 一台看见不等于事件成立

同一次地震通常会在不同台站留下受传播路径影响的信号。要把单台候选提升为事件，研究者需要判断哪些到时来自同一震源，并估计震源位置与起始时刻。这是事件关联和定位的问题，与单台波形分类具有不同约束。一个候选在单台上很像地震，仍可能无法与其他台站形成合理的时空关系。

最基本的定位思想，可以用一个教学表达说明：某台站的到达时间等于震源发生时间加上波从震源传播到该台的时间。传播时间取决于位置和速度结构。多个台站提供多条约束，研究者寻找能够同时解释这些到时的震源。这个过程并非简单地数一数多少台站出现峰值，因为不同震相、路径与位置误差都会改变对应关系。

关联错误与拾取错误应分别保存。拾取错误发生在某个到时被放错位置，关联错误则可能把两个事件的到时拼成一个事件，或把同一个事件拆成多个。网络能够提高单台拾取能力，并不自动消除关联问题。在密集余震序列中，事件重叠还可能使单台上看似合理的候选组合，在多台约束下不再成立。后续流程不是附加手续，而是赋予事件身份的组成部分。

定位残差也不能单独作为真假的万能判据。若采用不合适的速度模型，真实事件可能表现出系统残差；若可调参数太多，错误组合也可能被拟合得较好。台站几何决定了某些方向是否受到充分约束，少量集中在一侧的台站，不能与环绕震源的台网简单等同。一个位置数值输出得很精细，不代表误差范围相同精细。

在联合检测与拾取之后，还需要一套候选复核流程。这里提出的后续研究设计，首先保留原始波形和模型评分，使到时判断能够回查；其次标明候选进入关联的阈值与规则；随后检查跨台一致性、定位约束与不确定性；最后对新增候选开展独立人工或替代方法复核。拒绝原因也应随候选保存：波形不支持、跨台关联失败与位置约束不足，分别指向不同的改进环节，不能合并成一个无效标签。

独立人工复核还涉及审查者所见材料的范围。只展示系统标出的短窗口，容易把候选入口和判断结果同时交给模型。盲化部分标签、混入对照窗口、保留审查者分歧，可以减轻这种提示效应；从完整连续记录中另行抽样，则可检查模型从未提交的信号。前者主要约束候选确认，后者触及漏检。两类样本同时保留，才能说明人工复核究竟校正了目录的哪一面。

还应注意，不能把无法定位的候选一律判为非地震。台网覆盖不足或信号太弱，可能使真实事件暂时无法获得稳健位置。此时合理的身份是未充分确认的候选，而不是为了清单整齐强行二分。科学目录允许不同可信等级，前提是等级含义明确，使用者不能把弱候选与已充分约束事件不加区分地汇总。

事件成立因而是一个共同约束的结果：波形特征支持地震来源，震相时序支持传播解释，台站几何支持位置估计，复核规则支持目录可信度。这里没有任何单项完全代替其他项。AI 扩大了第一道辨认能力，真正的新事件对象则在这些约束之间获得身份。

到时与位置之间的关系，还能解释为何更精确拾取未必按同一比例改善定位。假设某台站的到时误差减小，但速度结构仍有系统偏差，定位可能依旧受后者限制；如果台站全部位于震源同一侧，某个方向的不确定性也可能较大。误差链中最弱的环节会限制整体收益。这个教学推理说明，算法单项指标不能未经传播分析就转换为科学对象的精度。

相反，一些单台到时略有误差的候选，仍可能在多台联合约束下得到稳健事件身份。共同体的判断不要求每个输入完美，而要求错误范围与约束结构相容。多个资料能够互相限制，也是复合对象超过单一输出的地方。事件目录的可靠性因而不是把每个模型评分简单相加，而是检查它们是否能在同一物理解释中共同成立。

还要防止对照资料被当成无误真值。人工目录可能漏掉弱事件，新方法相对于旧目录的新增项并不全是误报。评估应对新增项作独立复核，同时对旧目录中的事件也检查是否能够重现。把旧目录当作最终裁判，会低估真正的发现；把新目录全部当作真值，则会高估模型能力。双向检查使比较保持开放。

◆ 53.3 更多目录与更多自然地震

当新方法检出更多事件，最容易出现的叙述是地震活动突然变得密集。然而，目录数量同时受自然活动与观测能力影响。台站增加、噪声下降、采样方式改变、阈值降低和算法升级，都可能使原来没有进入记录的事件被纳入。观察到的数量变化，需要先解释检测条件，再讨论自然过程。

一种教学模型可以帮助把问题拆开：记录事件数近似取决于真实事件的分布与不同条件下的检出概率。这个表达不是地震活动的完整统计模型，也不意味着真实事件数可以直接从记录反推。它的作用是提醒读者，同一记录增长，可以由更多自然事件、较高检出能力，或两者共同造成。没有对检测条件的约束，数量增加本身无法区分这些解释。

因此，新旧目录的比较应尽量在相同时间范围、相同原始台站资料上开展。若新算法处理了更多台站或更长记录，却把全部增长归于模型能力，比较对象已经改变。若旧方法

只处理较强信号，而新方法使用不同阈值，研究者也应说明自己是在比较操作策略，还是比较固定策略下的模型表现。不同问题可以都合理，但不能互相替代。

在比较自然事件率时，还需要限定目录的完整性范围。对于某个给定时期和区域，较小事件可能更容易遗漏。若前期只能稳定记录较强事件，后期加入大量弱事件，总数曲线就会产生观测系统造成的转折。稳健比较可以限定在共同可靠的范围，也可以明确建模检测概率，但不能把观测能力的改善隐去之后宣称活动增强。

迁移检验则触及另一种边界。模型在一种地区或台网训练，在另一地区运行，波形、噪声与传播条件可能不同。若训练与测试资料共享相似事件、台站或时间片段，表面良好的表现未必代表真正的跨区能力。评价划分应与使用目标一致：准备移植到新地区，就需要新地区的独立检验；准备长期运行，就要检查时间漂移和仪器变化。

新增候选被否定的记录尤其重要。只展示确认成功的事件，会使读者无法知道系统带来了多少复核负担。某种模型可能检出更多真实弱震，同时也产生大量难以排除的噪声候选；另一种模型可能更保守，适合某些业务任务。能力评价应同时讨论漏检、误检与人工成本，而不是把任何数量增长当成单向进步。

目录升级还会改变科学问题。原来稀疏的记录可能只支持区域级活动描述，较密的事件位置可能允许研究断层内部的时空聚集、余震迁移或震群结构。但这种潜力不意味着已经发现具体机制。事件间距离需要连同定位误差解释；看似迁移的模式可能受时变检出能力影响；聚集关系还可能对关联规则敏感。更细目录把问题打开，也同时提出更细的误差责任。

在这一点上，新科学世界的发生可以作严格限定：研究者、台网与算法共同获得了以前不足以稳定区分的事件关系，使它们成为可比较、可反驳的对象。古老地震并非因为模型运行才自然发生，发生的是这些事件获得新的科学身份与联系的能力。若新增关系在独立检验中消失，就应撤回这一对象判断，而不以“AI 时代”作为免检理由。

数量关系可以通过一个明确的教学例子说明。假设某地区实际活动不变，而检测系统对弱事件的检出比例提高，记录总数仍会增长。若研究者随后用总数增长解释应力状态变化，就把测量能力变化放进了自然机制。反过来，若检测条件稳定而较大事件与弱事件都

表现出相容变化，才更有理由继续研究活动变化。例子不提供真实地区的数值，只揭示两种解释在何处需要不同证据。

更细事件目录还可能改变空间聚集的观感。旧目录只显示少数较强事件，新的弱事件填补其间空白，原来分离的两团活动可能变成连续带。这个形态变化可能有地质意义，也可能受位置误差与选择范围影响。检验时应在相同定位方法下比较，并评估把弱事件删除或移动到误差范围内以后，结构是否仍然存在。

对震群迁移的研究尤其需要时间一致性。若台网在序列中途增加台站，远处弱事件开始被检出，目录可能呈现向外扩展的形态；若处理者逐步降低阈值，也可能产生类似效果。因此，传播速度一类量不能只由点图拟合得到，还应检查检测覆盖与定位误差是否随时间变化。新对象越精细，对观察过程的审查也越细。

对新增事件的命名也有技术后果。若每次处理都给事件重新编号，却不保存新旧对应，后续引用会难以追踪。一个事件在新版本被拆成两个，或两个候选被合并，应保留这种谱系，而不是简单删除旧条目。目录身份的连续性，使机制研究能够判断结果变化来自自然资料增加还是对象重组。

这种谱系还可以帮助发现系统偏差。如果被反复合并或撤回的事件集中在某些台站、时段或信号条件，问题可能不在单个候选，而在共同处理环节。误差分布因此成为新的研究材料。AI 系统所产生的失败记录，可以像成功候选一样参与知识形成，前提是它们被保存而非隐藏。

此外，工程速度的提高也需要放在完整流程中评价。检测很快而复核积压，未必缩短可信目录的形成时间；快速输出如果缺少必要元数据，可能增加后来追查成本。真正有意义的效率，是从连续记录到可使用知识的整体效率，而不是某个网络单次运行的速度。这个尺度将算法性能与共同体使用联系起来。

目录质量最终还要与科学用途相接。同一组事件可能足以描述区域活动，却不足以研究细小迁移；能够检出事件，也不代表能够可靠估计其震级或机制。使用者应根据自己的推断链重新检查目录条件，而不是继承生产者的一句总体评价。对象的使用资格随任务变化，这正是工程输出进入知识时必须保留的边界。

未定候选也应拥有稳定标识和待补条件。例如，波形支持事件而位置不足的候选，可以等待新增台站约束；关联存在多解的候选，则应保留各组到时组合。这样，后来资料进入时能够针对原先缺口继续判断，而无须从头猜测旧目录为何排除了它。

◆ 53.4 从工程改善到科学问题

应当把三种成就分开。第一种是工程能力改善：在规定输入与评价条件下，更有效地处理连续波形。第二种是事件发现：一些此前未进入可信目录的候选，经外部约束取得身份。第三种是机制发现：新的事件关系排除了某些解释，并支持一个关于致震过程的命题。它们可以相互促进，但没有任何一层保证下一层自动完成。

例如，一套新目录可能显示更细的活动带。要解释为特定断层结构，需要检验位置不确定性与空间分布；要进一步解释活动迁移，需要考虑时间完整性和不同速度模型；要归因于某种驱动机制，还需要另外的物理或地质证据。把这些步骤压缩成“算法发现了机制”，既夸大算法，也隐藏了真正产生科学价值的联合工作。

SDE 分析在这里应追踪三种方向。S 式询问事件目录依赖哪些差异与特征约束，防止单台评分冒充完整事件。D 式询问新的事件对象怎样选择后续复查，例如是否要求重新查看邻近台站或某段连续记录。E 式则询问这些复查是否迫使研究者修改标签、传播假设或台站配置。只有相应记录存在，才能说三个方向在同一研究链中发生了联动。

三向追踪使用的仍是台站几何、误差传播和独立测试。其区别在于检查这些技术判断的交接：拾取误差是否完整传入定位，定位不确定性是否进入迁移分析，机制解释又是否要求重看被阈值排除的连续记录。若每一层各自评价良好而交接遗漏了不确定性，目录的总体结论仍可能过强。SDE 能够组织这类跨层问题，具体修订则由相应地震学方法承担；已被常规流程处理的交接，不再算作新增发现。

一个有价值的后续结果可能是“没有新机制”。如果模型使目录更完整，但新增资料仍支持既有解释，这同样是科学成果：它扩大了原解释受到检验的范围。科学进步并不要求每次技术改进都推翻理论。把稳定化也纳入历史，才能理解仪器与算法怎样在长期运行中使对象更可靠，而不把全部价值集中到少数革命性叙述。

对于人工智能史而言，这种区分尤其必要。自动处理减少的是某些操作成本，新增知识则取决于节省的能力被怎样投入复核、比较和新问题。模型只是复合研究系统的一部分。只有它与人的判断、台网资料和物理约束形成可纠错联系，工程能力才会转化为共同体能够继承的知识。

最后，目录的开放与版本记录使这条证据链能够成为共同体知识。每一个事件最好都能回到所用波形、拾取、关联规则和定位配置。模型更新以后，事件可以合并、拆分或撤回，旧结论也能据此重新检查。共同体承认的事件，不是永远不再修改的编号，而是在明确条件下可被他人追踪与质疑的对象。下一章中的火山预警，将把这种科学身份进一步放到行动责任之中。

第 54 章 雷道布特火山与航空预警知识怎样进入行动

54.1 1989—1990 年的具体难题

火山喷发进入航空系统以后，问题不再只是火山是否活跃。灰云可能沿大气输送进入航路，飞行活动又具有自己的时间窗口、空间高度与通知要求。对地面研究者有意义的观测描述，未必能够直接转化为航空使用者所需的信息。危险过程与暴露方式相遇，使一种新的知识对象出现：不是孤立的火山状态，而是需要及时传达、能够影响航空行动的火山危害状态。

Guffanti 与 Miller 在 2013 年的回顾研究中指出，阿拉斯加火山观测机构在 1989—1990 年雷道布特喷发期间设计了面向航空的活动警报制度。[C12]这一历史锚点说明，警报体系并不是把一套早已完整的知识简单贴上颜色。它是在具体危险、观察能力与信息需求之间形成的制度化对应。研究思想史时，需要把这项对应的建立与火山物理机制的认识分别追踪。

时间压力是其中的关键条件。科研论文可以在资料整理后提出解释，航空通知却往往需要在信息仍不完整时发出。等待所有细节明确，可能失去行动价值；过早作出过强判断，也可能造成不必要的扰动。预警因而承担一种特殊任务：在不确定性尚未消失时，说明当前证据支持什么，以及接收者应如何理解这份信息的时效与边界。

这并不意味着预警可以跳过证据。恰恰相反，越需要迅速传递，越需要清楚区分实际观测、推断与尚未确认的信息。一个灰云报告来自直接观察、仪器记录还是间接推测，会影响其可信度和下一步核查。若通知只有一个等级而没有必要的时间、地点与解释，接收者可能无法判断它是否适用于当前航路。等级压缩信息的能力，必须与可恢复的解释配合。

1989—1990 年的制度形成，应与后来全国或国际范围的标准化分开叙述。回顾材料支持航空警报制度形成这一历史节点；具体时刻的判断，则需要对应当时的原始通报和观

测记录。现行制度适合帮助辨认地面与航空任务的区别，不能反推早期已经具备同样成熟的术语与分工。历史问题因此落在规则的形成过程：哪些信息先被纳入通知，哪些责任后来才获得明确安排。

航空预警的空间对象也不同于传统地质图上的火山位置。火山口是一个重要来源位置，灰云却会离开来源区，在一定高度与气象条件下移动。把危险限制在火山周围的固定圆圈，可能忽略远距离输送；只知道火山位置而不知道灰云时间与高度，也不足以判断航路关系。危害对象因此从一个地点，扩展为具有来源、传播和暴露联系的动态结构。

这使火山学与气象学之间的接口成为知识形成的一部分。火山观测说明喷发与排放状况，大气资料帮助判断输送，航空系统再把空间信息转成使用条件。每个环节有自己的尺度和不确定性。接口若只传递一个等级，就可能丢失下一环节所需的信息；若传递过量未经组织的技术细节，又可能降低使用效率。通知制度的形成，包含了对这种压缩与保留关系的探索。

◆ 54.2 危险状态与行动阈值

危险是自然过程可能造成损害的能力，风险还涉及谁在何处暴露以及可能承受什么后果。火山活动的同一证据，对地面居民与航空使用者可以具有不同意义。地面危害可能关心熔岩、碎屑流或近火山区域的影响，航空危害则特别关心空中灰云及其传播。两个对象并不矛盾，它们在不同的互动条件中获得规定。

USGS 对火山状态与航空颜色代码的说明，明确区分地面状态信息与空中灰云相关信息。[R54a]本章引用这一现行说明，是为了理解对象区分，而不是把现行细节全部当作雷道布特喷发初期的历史事实。它告诉我们，预警体系不应只用一个抽象的危险大小概括所有使用场景。相同物理过程通过不同暴露关系，形成不同的通知任务。

行动阈值也不等于一个传感器的读数阈值。仪器记录可以支持活动增强的判断，但是否需要改变航路，还涉及灰云位置、传播、可用替代路线和负责机构的程序。科学判断与行动决定之间存在连接，却没有简单同一。若将两者混合，科学机构可能被要求承担超出其信息权限的决定，行动部门也可能把自己的选择错误归为自然事实的直接命令。

可以用一个不指向具体事故的教学情景说明这种差别。两个航班面对同一份火山灰信息，一个预计穿过相关空域，另一个远离该区域。危险信息相同，暴露关系不同，行动需求便不同。若天气变化使灰云移动，即使火山活动强度没有明显增加，某条航路的风险也可能上升。因此，警报对象必须保留空间与时间条件，不能只表达火山自身的单一状态。

不确定性也应按层传递。活动是否增强、是否已有显著灰云、灰云会怎样移动、某条航路何时可能受影响，分别依赖不同资料和推断。一个环节不确定，不代表整条链完全无知；一个环节确定，也不代表其余环节都确定。良好的信息产品能够让使用者识别这种分布，而不是把所有不确定性压成“可能危险”四个字。

这里的 S 是面向特定暴露与责任关系的警报对象，D 是活动变化、观测更新和通知时刻之间的差异序列，E 是火山过程、输送条件、使用场景与信息责任的相互规定。火山灰的物理特征规定何种传播和探测值得关注，航空运行又规定通知需要怎样的时间与空间分辨。这种纠缠不意味着管理需要可以改变物理事实，而是意味着可行动的危害对象不能由物理读数独自完成。

阈值选择可以通过决策后果来理解，但不能由后果直接替代自然判断。若漏报的潜在损失很高，行动系统可能采用较保守的策略；这并不等于科学上可以把不确定信号说成已经确认。更合理的做法是分别表达证据强度与行动倾向，使接收者知道为何在有限信息下仍需采取预防性措施。

这种分层能够避免两种相反误用。一种是要求科学通知只有绝对确定才发布，使信息失去时效；另一种是把任何担忧都表达为确定危险，使等级逐渐失去辨别能力。制度需要在证据更新与行动需求之间建立可解释的对应，并通过后续记录检查这种对应是否适合实际使用。它不是一次规定以后永远不变的翻译表。

预警的撤销或降级同样是一种知识行动。它需要说明证据为何支持危险减弱，或者原先判断为何不再适用。若只重视升级而不保留降级依据，历史记录就会偏向危机时刻，难以评价整套制度的稳定运行。没有灾害的时期，也提供关于监测噪声、阈值与通知负担的重要资料。

54.3 误报、漏报和迟报

如果一份预警最终没有对应严重后果，就能称为误报吗？不能直接这样判断。行动可能已经减少了暴露，危险也可能在不确定范围内没有发展到最严重状态。评价预警需要回到发布时可用证据与当时规则，而不是只凭事后结果。否则，成功避免损失的预警反而可能因为没有事故而被判为无用。

反过来，发生严重后果也不自动证明科学人员事先已经掌握充分证据。研究者应区分在当时信息下可合理预见的风险，与事后资料完善以后才清楚的过程。这种区分不是免除责任，而是使责任可以具体讨论：是没有观察到关键变化，还是观察到了却没有解释出来；是解释已经形成但未及时传递，还是信息已经送达但没有被正确使用。

因此，误报、漏报与迟报应分别建账。误报涉及信号被错误解释或等级过强，漏报涉及应当辨认的危险没有进入通知，迟报则可能发生在探测、分析、审批、传输或接收任何一段。它们可能同时出现，也可能彼此独立。笼统说预警失败，会把改进方向混在一起；明确失败位置，才知道应增加传感器、改变解释规则，还是修订通知流程。

一份可复核记录应至少保留观测时间、资料到达时间、初步解释时间、通知发布时间与接收确认。它们不是为了事后制造精确的问责故事，而是为了辨认延迟在哪里产生。若只有通知最终版本，没有此前变化，就无法知道判断是否随证据及时更新。若只有科学记录，没有传递记录，也无法断言用户何时获得了信息。

还应把语言误解作为单独的问题。一个等级在科学部门内部含义明确，接收者却可能把它理解为更强或更弱的行动指令。术语的共同性需要培训、解释与使用反馈来维持。颜色或等级并不会自己携带完整语义。预警信息在发送者和使用者之间取得稳定含义，也是一项需要持续检验的互动成果。

公开档案可以帮助检验这些问题，但材料缺口必须留在结论中。缺少接收记录时，不能仅凭发送时间推断所有人已经知道；缺少当时观测覆盖时，不能用后来更完善的监测能力评价早期判断；缺少行动档案时，也不应把后果直接归于某个部门。保留未知，使历史分析有能力提出下一步取证，而不把推测写成裁决。

评价一套预警制度，还需要合适的比较单位。可以比较单次通知是否准确，也可以比较一段时期内系统对活动变化的响应。前者适合分析具体错误，后者有助于看见日常运行

中的重复问题。若只选择著名事故，容易忽略大量平稳通知；若只统计总体比例，又可能掩盖少数后果极大的失败。两种尺度应相互补充。

接收者行为还会改变后来可见的结果。航路调整使飞机避开灰云以后，事故数量减少，不能据此反推灰云原本没有危险。评价应区分危险过程、暴露与损失，并尽量利用独立的灰云和运行记录恢复关系。这与临时把所有结果归功于预警不同：其他措施和天气变化也可能参与作用，因果归属需要比较。

信息延迟则可以被拆成更具体的时间差。观测到资料可用的间隔，反映测量和传输；资料可用到判断形成的间隔，反映分析；判断到通知的间隔，涉及组织程序；通知到行动的间隔，又属于使用系统。把这些间隔分开，才可能发现某项技术升级是否真的缩短了关键环节，还是只把延迟转移到下一处。

若一套制度长期没有出现严重事件，也不意味着无需检验。可以通过历史回放或程序演练检查信息是否能按规定流动，但这种演练检验的是流程能力，不等于真实喷发条件下的全部表现。演练资料与自然事件资料应分别标记，避免把模拟顺利当成实战效果。

预警知识还具有有效期。关于活动或灰云的判断可能随新观测迅速改变，旧通知如果继续在传播，就可能脱离原来的条件。通知中明确发布时间、更新关系与失效状态，能够使接收者知道自己使用的是否仍是当前判断。版本责任在这里具有直接的行动意义。

由此，知识进入行动不是一个单次传递动作，而是持续更新的关系。发送者需要解释变化，接收者需要识别新旧，行动系统需要把信息与自身暴露相接。任何一个环节失去时间身份，都可能使本来正确的科学内容在错误条件下被使用。

知识行动的评价因此应同时问两件事：通知是否忠实表达当时证据，通知是否以接收者能够正确使用的形式到达。前者防止夸大自然判断，后者防止把责任停在发送按钮。两者相互依赖，却需要不同记录。把这种双重要求保存下来，能够使制度改进有明确方向，也使科学不确定性与组织责任得到合理区分。

更新关系还应说明旧通知被替代还是仅被补充，因为两者对接收者继续使用旧信息的意义不同。

54.4 事故如何回写监测

复盘的价值不在于报告写了多少条建议，而在于哪些建议改变了实际运行。增加观测点、修订术语、调整通知模板、缩短信息传递环节，分别对应不同层的回写。要证明某次事件推动了这些变化，需要连接复盘意见、决策记录与实施证据。只知道某项措施后来出现，并不足以证明它完全由该事件造成。

监测安排的改变，可以使新的危险特征进入可见范围；通知模板的改变，可以使原来模糊的时间与空间条件得到明确表达；责任分工的改变，则可以减少信息在机构之间停滞。三者都重要，但不能一概称作自然机制发现。某些改进主要是知识使用关系的重组，某些才涉及科学对象或解释的修订。

制度修订是否进入科学工作，可以看它怎样改变观测和解释。仅调整表格顺序，通常属于程序维护；若通知要求新增灰云高度或更新时间，监测任务就可能随之改变；若新记录又暴露原分类不能区分某类航空危害，分类规则也需要重审。这里呈现的是一条可追查的联系：使用所需的区分进入观测，观测结果再检验通知中的区分。各环节是否实际发生，应由相应记录确认。

本章据制度史料提出的分析链，由活动观测、危害判断、有时效的通知和后续修订组成。原始通报可以说明某项判断何时公开，接收记录可以说明它何时进入使用系统，行动档案与修订文件才把后续变化接回这份信息。现有材料确认制度形成的节点，尚不足以逐次恢复这些连接。保留各类记录的不同作用，可以使后续研究具体区分：知识在传输中延迟，在解释中失真，还是在行动条件改变后已经失效。

由此看，预警中的统一性不是要求所有使用者采取相同行动，而是让他们共享来源清楚、含义稳定、能够更新的危险信息。多样性来自不同暴露和责任，和谐性则来自接口能够承认这些差别并相互协调。若以同一等级强行替代所有具体判断，表面的统一反而可能破坏信息使用。真正有效的共同对象，需要既稳定又保留任务差异。

雷道布特案例由此把本书的问题推进了一层。对象的可靠性不仅需要自然约束，也需要信息身份与责任边界。当一项知识被不断更新并反复使用，旧判断与新判断之间的关系就必须保存。下一章讨论版本，正是把这种责任从预警扩展到遥感档案与人工智能目录。

第 55 章 版本也是证据
从 Landsat 重处理到 AI 目录更新

◆ 55.1 版本号背后的不同改变

一个研究结果被更新以后，旧结论究竟发生了什么？答案取决于更新改动了哪一层。增加新的观测，与修正旧观测的标定不同；替换算法，与修改对象分类标准不同；撤回因果解释，又不同于保留解释但缩小适用范围。如果这些变化都只表现为文件名从第一版变成第二版，知识的实际变动就被一个编号遮蔽。

Landsat Collection 2 展示了档案重处理的情形：自然观测时点未变，产品却可以因参考资料、标定和处理改进而改变。[C09]AI 地震目录展示另一种情形：相同连续波形经过新模型和规则以后，可以产生更多、较少或重新关联的候选。两者都涉及版本，但前者的像元产品与后者的事件身份具有不同结构，不能用“算法升级”把它们混成同一种过程。

原始测量修正，应保留改动对象和依据。例如，若仪器响应的认识发生修订，就需要说明哪些时期或通道受到影响。处理算法更换，应保留输入、参数和软件配置。分类标准改变，应说明新旧类别怎样对应，哪些对象被重新归类。解释撤回，则必须指出原先哪一步推断失去支持。不同层的证据字段，正是让版本变化可以被理解的最低条件。

一个数据产品标识，因此具有科学功能。它帮助使用者知道自己引用的是哪一次处理结果。只有产品名称而没有处理身份，后来重做研究的人可能得到不同数值，却不知道差异来自新资料还是新方法。版本不是附属于知识的技术杂务，而是知识获得可追溯身份的组成部分。

但版本号本身并不证明进步。某次更新可能只改善文件组织，另一次可能修复错误，也可能引入新的偏差。判断应回到明确任务与独立证据。流程更加可追踪，值得肯定为治理改善；解释更加可靠，则需要另外检验。把每次发布都写成知识跃升，会使版本史变成产品宣传史。

版本变化还可以按可逆程度区分。有些处理保留原始资料与完整参数，后来者能够重新生成旧结果；有些分类只留下最终类别，原来的连续评分已经丢失；还有一些人工判断没有保存理由，无法精确复原。可追溯性不是有无版本号的二分，而是沿证据链分布的能力。一个环节的信息丢失，可能限制整条结论的恢复。

这也说明原始资料并不意味着完全未经规定的自然。仪器记录已经带着采样、响应和时间标定。保存原始资料的意义，是保留尽可能靠近测量的层次，使未来能够重新处理；它并不使后续分析摆脱所有前提。版本治理应同时记录测量身份与处理身份，避免把原始二字当成无需解释的保证。

◆ 55.2 同一输入的前后对照

固定输入，是识别处理变化的基本策略。对于遥感，可以选择同一获取时刻的原始资料，比较两套处理产品；对于地震，可以选择同一组连续波形，比较两套检测、拾取和关联配置。这样仍然不能自动证明哪一套更好，但至少可以把自然事件与资料覆盖的差异暂时控制住，让处理本身成为比较对象。

接下来应记录差异发生在哪里。遥感产品可以比较连续数值、空间位置、质量掩膜与最终分类。AI 目录则可以比较候选增加与删除、到时变化、事件合并与拆分，以及位置不确定性。只比较一个总数，会隐藏很多不同机制。两个目录事件数相同，可能包含完全不同的事件；两幅分类图面积相同，也可能在空间上大幅错位。

差异的解释还需要外部参照。新产品更接近旧产品，不必然说明更可靠，因为旧产品可能存在系统偏差；新目录增加了候选，也不必然说明更完整，因为其中可能有误报。独立的地面资料、另一路观测或盲化复核，能够针对具体争议施加约束。比较不能把“新旧一致”与“接近真实”视为同一评价。

同输入比较可以保存一组相互对应的材料：固定输入清单、两套配置、差异结果、外部检查与解释修订。输入与配置使运行能够恢复，差异记录保留新增和删除，外部检查判断变化是否修复实际错误，解释修订再指出受影响的科学命题。譬如，关联规则减少了重复事件，但主震后的活动率变化仍然保留，那么目录计数需要更新，活动变化的方向则可能继续有效。这个条件性例子说明，处理改进和结论改动应分别记录。

还有一项容易忽略的控制：算力与人工投入。新流程若使用更多人工复核，结果改善不能全部归于模型结构；旧流程若被粗糙复现，也不能作为公平基线。比较双方应尽量获得相当的资料权限和合理调参机会。否则，所谓算法胜利可能只是投入差异的结果。

对于无法完全恢复的旧版本，研究者应说明限制。可能缺少当时依赖的软件，可能原始数据已被修订，也可能人工决策没有记录。此时仍可比较现存结果，但不能宣称完成严格的同输入实验。复原失败本身具有思想史意义：它揭示知识共同体曾把哪些条件视为无须保存，而这些条件后来如何成为检验瓶颈。

对照还可以分为结果回归与科学检验。结果回归检查新流程是否保留应当不变的行为，例如同一输入的某些稳定输出；科学检验则检查结果是否更好满足外部约束。新流程可能通过回归而仍然保留旧偏差，也可能因为修正旧错误而有意改变结果。二者不能混同，否则团队会为了保持一致而阻止必要修订，或为了追求变化而忽略意外损坏。

一种清楚的差异说明应区分预期改变与非预期改变。标定修订预计影响的区域和变量，可以预先列出；比较后出现的其他差异，则成为调查对象。这样能够避免事后把任何结果都解释为改进。负结果也有价值：如果某项修订没有影响预期对象，就需要检查假设、实现或资料范围。

版本比较还应注意尺度。像元层的局部差异在区域平均中可能抵消，事件级的合并拆分在总数中也可能相互抵消。因而，粗尺度一致不能证明细尺度对象未变。反过来，许多微小差异也未必改变研究结论。评价需要同时说明对象变化与命题变化，防止把文件差异夸大为理论转折。

◆ 55.3 标签回写与循环验证

AI 目录更新提出一个更尖锐的问题：模型输出能否成为下一代模型的训练标签？答案不能简单地是能或不能。自动输出可以帮助发现值得复查的材料，减少人工搜索负担；但若未经外部确认就当作真实标签，模型的错误也可能被写回训练集。下一代模型再与这些标签一致，表面性能提高，实际却可能只是更熟练地重复前一代偏差。

这种循环的危险，在弱事件和边缘样本中尤其明显。它们本来就是最需要外部证据的部分。如果模型把噪声误判为某类事件，而人工只确认容易的案例，未确认候选又被默认

为真，标签库就会发生选择性积累。后来的评估若仍使用同源标签，错误不容易显露。资料数量增长与外部约束增强，在这里可能背离。

应当区分自动建议、人工确认与独立观测支持。自动建议记录模型看到了什么，人工确认记录人在规定条件下作出了什么判断，独立观测则可能通过不同信号或物理约束检验事件身份。三者可以结合，但不应在数据库中被压成同一个没有来源的标签。来源等级保留得越清楚，越容易检查新模型是否把不确定材料当作确定事实。

人工确认也不是独立性的自动保证。若确认者只看到模型选中的窗口，没有接触被模型遗漏的部分，整个过程仍然受到选择偏差影响。要估计漏检，需要从连续记录中抽取不依赖模型筛选的样本；要估计误检，需要保存被拒绝的候选及理由。被排除的资料不是垃圾，它们是衡量边界和错误结构的重要证据。

数据划分也必须尊重事件关系。来自同一事件、相邻时间窗口或同一台站环境的资料，若同时进入训练和测试，可能使模型利用高度相似的信息。形式上不同的文件，不一定是独立样本。研究者应依据实际用途选择划分单位，并说明共享因素。若目标是跨地区应用，地区独立就比随机窗口独立更有意义。

两种反馈可以从标签的证据来源区分。外部复核否定一个候选后，标注规范据此修改，下一轮训练便获得了与旧输出不同的约束；模型输出若直接成为新标签，则可能只扩大原有选择。二者都有数据回流，只有前者在这一环节增加纠错依据。SDE 对 E 的追踪，需要落到波形特征与类别资格的对应是否改变：哪些特征原先被当作地震，何种跨台证据使这一对应失效，修订以后怎样处理相似信号。

循环验证还会改变研究者对错误的可见性。当模型只推荐它最有把握的样本，人工看到的资料会逐渐集中于模型擅长的类型。难以识别的区域被排除在工作流之外，错误反而不再进入反馈。一个看似高效的系统，可能通过减少遇见困难的机会维持高成绩。独立抽样的作用，就是重新打开系统不主动提供的差异。

另一个风险是类别逐渐收缩。如果新模型的建议反复成为标准，原来具有争议但真实存在的边缘类型可能被重新标为噪声。下一轮训练再把这些类型排除，模型会越来越自洽，却越来越缺乏发现能力。保存争议标签、允许多名审查者分歧和追踪被删除类别，能够使这种收缩进入可见范围。

因此，标签治理不只是提高标签准确率，还涉及保持外部纠错与新类型进入的通道。一个研究系统需要稳定标准才能比较，也需要允许标准在新证据下改变。完全固定会阻碍发现，任意变化又会破坏连续性。版本记录使这两种需求能够共存：改变可以发生，同时旧判断的条件仍然可以被恢复。

◆ 55.4 更新后谁仍能复原旧结论

一个可负责的修订说明，应让读者知道哪些结论继续有效，哪些需要重新计算，哪些已经撤回。例如，标定变化可能影响绝对量值，却不改变某些大尺度排序；事件关联规则改变可能影响小事件数，却不改变某次较大地震的身份。更新不是把旧知识全盘清零，而是重新划定有效部分。

这要求引用同时包含数据与处理版本。只引用论文题名，未必足以复原使用的目录；只记录软件版本，也可能遗漏模型权重、阈值和速度结构。记录粒度应与结论敏感性相称。某个结论若依赖细小差异，就需要更细的处理身份。保存并不是越多越好地堆积文件，而是保留能够改变判断的条件。

版本更新还可以改变结论的比较单位。原来按整景统计的遥感结果，若改为按有效像元统计，两次平均值的分母已经不同；原来按候选信号统计的地震清单，若改为按关联事件统计，数量也不再直接对应。修订说明应首先指出单位是否改变，再谈数值增加或减少。否则，精确的百分比会建立在不可比对象上。

相同原则适用于不确定性。新版本可能不仅改变中心估计，也改变误差表达。较宽的误差范围未必表示方法退步，可能是原先遗漏的不确定性终于进入计算。评价不能只奖励看起来更精确的数字，而应检查范围是否更诚实地覆盖实际误差来源。

这些变化说明，版本治理既保存计算，也保存解释权限。后来者应知道一个结果允许支持什么，不允许支持什么，以及这些权限为何改变。它把知识的边界变成可以讨论的记录，而不是依赖作者临时记忆。

当修订没有改变核心结论时，也应说明稳定来自何处。可能是误差远小于效应，可能是不同环节的变化抵消，也可能是结论本身较粗。只有说明稳定条件，才能判断它是否能够迁移到其他任务。

版本引用还应保留访问与获取的时间，因为在线资料可能在同一入口下继续更新。固定的标识、可保存的副本和明确的修订关系，共同减少这种漂移。对思想史作者而言，这使引文不只是指向一个网站，而是指向曾经支持特定判断的一份材料。

当后来者发现引用已经失效，恢复材料的困难也应记录。它可能提示公共档案需要更稳定的保存制度，而不是把责任全部交给个别作者的记忆。

保存撤回原因尤其重要，它使后来者能够区分资料错误、处理错误和解释范围改变，避免重复同一种误用。

版本史还是一种思想史材料。它让我们看到，哪些错误被共同体承认，哪些对象被重分类，哪些研究路线因此改变。与事后平滑的成功叙事相比，修订记录能够保留真实的犹豫、局部有效和失败。一个旧结论为何曾经合理，又为何后来需要缩小范围，常常只有在版本关系中才能看清。

用三式复审，S 是特定版本下成立的知识对象，D 是新旧结果与外部资料的比较，E 是标定、标签、规则和对象身份之间的相互规定。对象会选择比较的尺度，比较又可能改变规则；但并非每个版本都完成这三向关系。只有文件组织改变时，应限于维护；独立证据改善解释时，才可以讨论认识进展；对象定义与操作安排同时改变时，才进入更强的发生学判断。

知识版本还具有传播后果。某篇论文引用了旧目录，后续研究又引用该论文的结论，原始数据更新并不会自动传遍整条引用链。修订说明应指明受影响的命题与使用条件，使后来者能够判断是否需要回查。对于未受影响的结果，也应明确保留，避免一次技术更新引发没有依据的全盘怀疑。

这种传播关系使档案维护成为共同体记忆的一部分。一个结论的生命，不只取决于它最初如何产生，也取决于修订能否到达使用它的人。SDE 所说知识在互动中传播与变化，在这里可以具体落实为引用、版本通知、复查和再解释的链条。链条是否畅通，可以由实际记录研究，而不必停留在抽象宣言。

因此，保存旧版本的目的不是阻止修订，而是使修订有来处。真正开放的共同知识，允许后来者恢复当时条件，理解旧判断，并据新材料改变它。版本制度把“我们如今知道什么”与“我们怎样改变所知”联系起来。最后一编将以这种可追溯性为基础，检验本书自己的解释框架。

第 12 编

发生学综合与研究边界

第 56 章 不能重演的过去，怎样接受可以重做的检验

第 57 章 三方程的增量在哪里：从提示框架到可败解释

第 58 章 六路径与三原理：比较转折，不替历史排队

第 59 章 两项承重论证：样品年龄与板块对象的形成

第 60 章 开放的地球：下一轮问题从哪里发生

第 56 章 不能重演的过去，怎样接受可以重做的检验

56.1 重做的是什么

地球科学研究的许多事件已经过去。古老大气的变化不能重新发生在同一个地球上，某次沉积过程也不能在原地点按原条件完整重演。如果把科学检验等同于把对象原样再做一遍，地史研究似乎天然处于不利地位。但这个要求混淆了事件与检验。能够重做的，是测量、取样、模型运行和证据比较；被检验的，则是关于不可重演事件的解释。

第 47 章讨论海洋记录与轨道变化，第 48 章讨论辐射—对流模型，第 49 章讨论硫同位素与早期大气。三个案例中，重复操作的对象不同。研究者可以重新测量保存样品的同位素组成，可以重新取得沉积记录，可以在明确条件下运行模型。它们都不等于重新制造一个冰期或一段古老大气史，却能够检查原解释所依赖的联系是否成立。

重复测量主要检验测量链与样品特征的稳定性。不同实验室得到相近结果，会提高对某种样品属性的信心；但若样品后来发生过改变，一致测量仍然可能稳定地描述了一个后期改造结果。重新取样可以检查空间分布，却又受到地层对比与保存条件影响。模型重复运行可以检查计算是否可复原，但同一模型得到同一结果，并不独立证明模型适合历史对象。

因此，重复性必须接到具体命题上。若争议是数值误读，重测可能有决定意义；若争议是痕迹形成机制，单纯提高测量精度未必有帮助；若争议是年代对应，更多同类曲线也不能替代年代约束。历史科学并不缺少可检验操作，它要求更清楚地说明每一种操作能够约束哪一步推断。[N03]

可重复操作还具有不同的独立程度。对同一样品重复测量，主要检验分析稳定性；换取同层位的另一件样品，增加局部代表性约束；换到另一剖面，则同时检验空间对比，但也引入新的地层关联问题。三种重复不是越远越好，而是分别回答不同问题。研究设计应让重复位置对应推断链中的薄弱环节。

模型重复也可以分层。完全相同代码与参数的运行检查可复原性，改变参数的运行检查敏感性，采用不同结构的模型则检验解释是否过度依赖某一表示。它们都具有价值，但第三种比较仍可能共享相同物理假设。把模型数量当成独立证据数量，会再次遇到后文讨论的依赖问题。

◆ 56.2 机制可能与历史实际

硫同位素案例提供了一条重要分界：某种机制能够产生特定信号，与它确实造成了某个古老样品中的信号，是不同命题。实验与理论可以表明某些大气化学过程具有产生特定同位素关系的能力，但要把这种能力连接到地史，还需要样品年代、保存状态与分布证据。[C07]机制的可行性提供一条通路，历史实际则要求证明这条通路适用于眼前材料。

这一点可以通过逆向推断理解。若机制甲能够产生痕迹乙，观察到乙并不能仅凭这一关系就唯一确认甲，因为机制丙也可能产生相似结果。需要寻找能够区分甲与丙的其他痕迹，或者证明丙在相关历史条件下不能成立。解释强度来自替代路径受到限制，而不是来自一个机制故事越来越生动。

保存过程尤其重要。古老信号从形成到今天测量，中间经历了沉积、埋藏、成岩或后期改造。某种特征可能被保存，也可能被混合或重置。研究者必须区分原始形成条件与后来的记录条件。若测得的信号来自不同年代材料的混合，把它直接解释为某一时刻的全球状态，就跨过了尚未建立的对应。

空间分布可以增加约束，但也不能自动消除歧义。不同地点出现相似信号，可能支持广域过程；也可能受共同的保存选择或年代对比影响。要判断其意义，需要知道各地点是否独立定年、样品是否来自相近环境，以及缺乏信号的地区是否具有同样保存机会。缺席并不总是过程不存在的证据，出现也不总是全球同步的证据。

替代解释受到限制的方式，应进入历史结论本身。若一种机制与信号方向相容，却不能解释其地层分布，它保留的是局部可能性；若另一机制同时解释方向、层位与保存条件，它获得的支持更强。两者差别来自覆盖了哪些独立约束，而非哪一个故事更完整。新增材料的价值也据此判断：它是否把原先同样可行的历史路径分开，还是只重复确认双方都能解释的特征。

海洋记录与辐射模型的关系，显示机制与历史推断可以互补而不互相吞并。轨道或外部条件提供某种变化的可能驱动，沉积记录提供地史响应的痕迹，气候机制说明能量与反馈如何使响应发生。若只有机制，可以产生许多可能轨迹；若只有曲线对应，因果方向与放大过程可能仍不清楚。二者结合，才能对具体历史提出更有约束的解释。

但这种结合不应把模型当作万能填补。资料空白不能仅靠一条平滑模拟曲线变成已知历史。模型在缺乏观测的区间可以给出条件性推断，表达的是在假设成立时会怎样，而不是直接记录了曾经怎样。历史叙述应保留这一证据身份，使读者知道哪些部分由痕迹约束，哪些主要依赖模型。

◆ 56.3 独立证据到底独立在哪里

多条曲线看起来相互支持，并不意味着它们提供了同样数量的独立证据。两条记录可能共享同一个年代模型，几个指标可能来自同一批样品，多项模型评估可能依赖同一训练资料。如果共享环节存在偏差，表面的多重一致就可能部分来自同一条件。第 47 章轨道与海洋记录的比较尤其需要检查这种依赖。

假如某条记录的时间轴经过对齐，以便与特定周期结构相符，后来再把这种相符全部作为该周期假说的独立证据，就可能重复使用同一个约束。这并不意味着调谐方法没有价值，而是意味着它的结果不能在论证中获得双重身份。用于建立年代的资料，与用于独立检验年代后关系的资料，应当明确区分。

共享样品也有类似问题。同一层位的多种指标能够提供不同物理信息，但它们可能共同受到保存或混合影响。独立性不是全有或全无，而是分层的：测量方法可以独立，样品来源可以不独立；年代约束可以独立，气候解释模型可以共享假设。把依赖关系画清楚，比笼统宣布多证据一致更有价值。

辐射模型的检验同样如此。若参数通过某组观测调整，再用同一组观测评价吻合，得到的是拟合能力，尚不等于独立预测能力。另一个时期、区域或变量的资料，可能提供新的约束，但也要检查它是否仍共享关键输入。模型与观测的关系越复杂，越需要记录资料在建立与检验中分别扮演了什么角色。

评价证据独立性，不是追求所有资料完全没有联系。地球系统本来就存在真实联系，完全切断关系反而可能失去研究对象。需要避免的是把来源依赖误算成额外支持。真实的物理耦合可以成为解释内容，方法上的重复使用则应在论证强度中扣除。两种联系必须区分。

这一原则在第 47 至 49 章形成三种不同检查。对海洋周期，先区分用于建立年代的约束与用来检验周期的资料；对辐射模型，区分设定条件与留待检验的响应；对硫同位素，区分实验中的形成机制与古老样品的保存历史。三案可以共享依赖分析，却没有可直接相加的证据总分。联合权重还需要原始资料与统计模型另行计算。本章能够比较的，是新增材料进入哪一步推断，以及共享环节改变后哪些结论必须一起重审。

可以把证据关系理解为一组具有共享节点的推断链，而不是一叠互不相关的证明。某个年代模型是多条曲线共同经过的节点，某种保存假设是多个指标共同依赖的节点。若这个节点改变，多条结论可能同时变化。识别这样的承重点，能够把复核资源集中到最能改变判断的地方，比不断增加同源曲线更有效。

这种做法也使负证据具有位置。某项新测量没有支持预期关系，可能反驳机制，也可能表明样品保存条件不同。应先检查它击中了哪一个节点，而不是立刻排除异常。异常只有在保留其来源和条件时，才能成为修订历史解释的力量。

在这种意义上，历史解释的可靠性具有结构，而不只是一个总分。年代关系可以较稳，具体机制仍有竞争；机制可能受到强约束，事件范围却不清楚。分层表达能够让下一项检验直接进入薄弱部分，也避免一处不确定把全部知识抹去。科学的有限性并非任意性，而是每项判断都有自己的有效范围与改动入口。

这种分层也保护了跨学科合作：实验者知道需要检验哪种机制，野外研究者知道需要寻找哪类痕迹，模型研究者知道哪些条件仍未受到充分约束。

◆ 56.4 怎样形成有限可靠的过去

一个成熟的地史解释，应当把替代方案分成几种状态。有些已与明确证据冲突，有些仍可竞争，有些目前缺少可实施的判别材料。这种分层比给所有解释贴上真或假更准确。

它让研究者知道，下一步工作应当针对测量误差、机制差别，还是资料保存的原则性限制。

SDE 在这里追踪历史对象与检验操作的相互改变。显露 S 是被限定的历史过程，D 是按年代、剖面或实验条件组织的差异序列，E 则在信号响应、保存状态和年代归属之间展开。例如，同位素差异只有与样品保存条件相接，才有资格约束古大气；若保存解释改变，同一差异就可能转为后期改造的证据。对象判断因此会选择下一轮取样位置，新的空间比较又可能改变保存解释。这样的回写应落实为取样与解释版本之间的对应。

若实验只使测量更精确，而未改变解释边界，应记录为稳定化。若新痕迹排除了一个历史路径，应记录为解释收缩。若原先被视为同一事件的材料必须分开，才涉及对象身份的修订。三种进展都可以有价值，却不能用同一个革命词汇替代。

过去不可重演，不意味着过去不可追问；能够重做一个实验，也不意味着历史已经被唯一证明。地球科学的力量，恰在于把不能重复的事件与可以重复、可以更换、可以彼此限制的检验联系起来。这种有限可靠性已经足以构成严肃知识，同时把下一轮问题保留下来。下一章将据此检验本书自己的三方程分析究竟增加了什么。

第 57 章 三方程的增量在哪里
从提示框架到可败解释

57.1 给竞争解释充分表达空间

一套框架能够解释很多案例，并不自动说明它比已有解释增加了知识。若它只是把原有概念重新命名，覆盖面越广，也可能越缺乏区分力。三方程进入地球科学思想史以后，必须面对同样的问题：在相同资料条件下，它究竟帮助研究者看见了哪一个原先没有被充分提出的差别？这个问题不能靠每章都填入 S、D、E 来回答。

本章选择水文状态、辐射模型与档案版本三个案例，分别考察输入状态、条件化响应和可追溯对象。水文问题需要分辨相似降雨为何产生不同流量；辐射问题需要分辨模型响应在哪些条件下成立；版本问题需要分辨旧资料重处理之后究竟改变了什么。共同框架的价值，应由它对这三种不同判断分别做出的贡献来衡量。跨案例的统一提问，也必须留下各案可能没有新增贡献的位置。

成熟学科解释构成比较的起点。水文学已经把土壤状态、入渗、产流和空间差异联系起来；辐射—对流模型以过程关系、边界条件和反馈规定计算；遥感与数据治理已经使用标定、质量控制及可复原流程。因而，关系是否存在并非双方的分歧。值得比较的是：某一具体解释保留了哪些关系，在哪个操作或使用环节压缩了它们，压缩是否改变了结论。只有把基线解释写到这个层次，SDE 的新增问题才有确定参照。

公平比较应先用各学科自己的语言重建最强说明，再问 SDE 增加了什么。可以暂时隐去框架标签，让评阅者比较问题覆盖、证据对应和操作建议。篇幅、资料权限与分析粒度也应尽量相当。若一方获得更多原始资料和更长论证空间，结果无法仅归于理论框架。

本章先完成文本层面的成对比较：对每个案例保留同一问题、同一已知条件，分别写出学科说明与 SDE 说明，再标明重合处和新增要求。这可以检查两种表达是否实际不同；要进一步判断它们是否改善研究者的工作，则还需独立评阅与操作记录。两级评价的

对象不同：前一级比较论证内容，后一级比较使用效果。本章结论限于前一级，后者留作可执行的研究。

比较对象还要避免不对称的理论层级。不能拿 SDE 的一般本体论与某个学科教材的简短段落比较，再把教材没有展开的内容算作学科缺陷。也不能拿一个具体水文模型的数值预测能力，与 SDE 的跨学科组织能力作同一维度排名。双方要在共同问题上接受评价：是否说明对象身份，是否指出可改变结论的证据，是否形成合理的下一步操作。

这种分层允许框架具有不同价值。一种理论可能不增加预测精度，却改善研究者识别对象偷换的能力；也可能帮助跨学科团队沟通，却不产生新的自然解释。把这些价值分开，并不贬低它们，而是防止某一种收益被扩大为所有方面的优势。

◆ 57.2 列出真正新增的判断

判断增量可以分为四层：是否提出新增问题，是否要求新增证据，是否改变研究步骤，是否修订结论。四层不必同时出现，也不能互相冒充。一套框架帮助整理问题，可能很有教学价值；但只有当整理使研究者增加了可独立检验的判断，才能提出更强的解释贡献。

在水文案例中，三方程可以提示研究者检查“同一场雨”是否真的构成相同输入：前期土壤状态、降雨强度过程与空间分布如何共同规定产流对象？然而，这些问题也是成熟水文学的核心内容。如果 SDE 只把已有状态依赖重新命名，就没有新增机制。可能的增量应进一步落到跨层联系，例如某种洪水对象定义是否改变了采样时间尺度，而这一改变又是否使原先的边界划分失效。即便如此，也必须拿出相应操作记录。

在辐射模型案例中，SDE 可以帮助把“模型响应”与“自然地球未来轨迹”分开，追问固定相对湿度等条件怎样规定可解释对象。但物理建模本来就要求说明假设。若所有重要条件在原论文中已经清楚，本书不能把自己再次强调这些条件当作首创。更有价值的工作，是检查某个后来的传播或使用环节是否删掉了条件，从而改变了结论身份。

档案版本案例则可能显示框架的跨学科组织作用。遥感像元与地震事件具有不同技术结构，却都需要追踪输入、处理、分类和解释的变化。三方程能够为二者建立一套共同提

问方式，使研究者检查对象判断如何选择复查、复查如何修改规则。这种共同性可以降低比较成本，但仍不能说明存在一个新的自然机制。组织增量与机制增量应分别评价。

新增证据也不能只指多引一篇论文。它应当能够改变竞争解释之间的判断。若某份操作日志证明对象定义改变以后采样计划实际改变，这会约束仅有措辞变化的解释；若一份版本记录证明新规则没有影响结论，则可能削弱对象重建的强主张。资料的价值来自它约束了什么，而不是来自数量或权威名称。

比较记录应把新增问题与已有答案分开。例如，询问洪峰目标是否改变采样，是一个问题；找到采样方案的修订，是一项材料；证明修订排除了原先的混淆，才形成判断增量。若只有第一项，结果栏就应保留未定；若基线已完成三项，则记录两种解释的重合。这样的区分使后续研究能够从缺失的一步接续，也防止把同一个启发在问题、证据和结论三栏重复计数。

可以具体比较一个水文判断。基线解释说，前期湿润程度改变可用储水空间与产流条件，所以相似降雨不必产生相同流量。SDE 解释说，洪水显露依赖降雨过程与土壤状态的共同规定，而且洪峰研究目标会要求较细时间采样。两段解释的前半部分高度重合，新增部分在于研究对象对采样的反向规定。但这部分是否具有历史或操作增量，仍取决于基线研究是否已经明确包含它。

再比较一个辐射判断。基线解释说，给定相对湿度分布等条件的平衡计算，得到的是这些条件下的温度响应。SDE 解释说，响应对象的身份由约束共同规定，后来的观测差异可能迫使改变约束。第一层同样重合，第二层只是提出后续研究方向，除非有实际修改记录。这个成对比较说明，框架最容易夸大的地方，是把合理的后续可能性写成已经增加的经验解释。

档案案例略有不同。基线治理说明要求保留版本、输入和参数，SDE 比较则把对象身份、操作选择与规则修订放在同一张关系图中。它可能帮助发现两个技术领域在回写责任上的共同问题，但共同图式并不代替各领域验证。实际已经得到的是可比较性，尚待检验的是这种可比较性是否改善研究判断。

57.3 三式怎样彼此约束

S=F(D,E)提醒研究者，显露对象不能脱离差异组织与特征纠缠而得到解释；D=G(S,E)进一步追问，已经成立的对象怎样选择后续操作；E=H(S,D)则要求检查操作结果是否反过来修改特征之间的规定。这里的操作与规则是本书寻找 D、E 变化的证据入口，并非二者的全部定义：D 仍指有序展开的差异序列，E 仍指具体特征如何相互规定。只有操作改变确实重组了所比较的差异，或规则修订改变了特征的相互规定，才能据此讨论相应维度的变化。三式作为分析关系，并不因为写成等式就自动获得地学定律的地位。[M01]其经验价值，要看在具体链条中能否找到对应的限制。

三式之间的约束，可以通过资料需求辨认。说明一个目录依赖波形与台站几何，只支持对象如何成立；要说明目录反过来选择复查，还要看到哪些位置或时段因新对象而被重新检查；要说明复查改变了特征关系，则要看到标签、传播解释或分类资格怎样随证据修订。三个方向需要接续的记录。某一方向缺少材料时，链条在该处保持开放，其余已经建立的联系仍可分别讨论。

水文案例可以说明第二式的要求。若研究者把对象从全年平均流量改为短时洪峰过程，采样频率与模型时间步长可能随之改变。但要把这种可能写成历史事实，需要看到研究设计确实发生了改变，且改变与对象转换有关。单凭两个时期采用不同仪器，不能排除经费、技术更新或其他目标的影响。

版本案例可以说明第三式的要求。如果新增候选暴露出标签错误，研究团队修订了标注规范，便可能形成可追踪回写。但若模型输出未经确认就加入训练，只能证明数据循环发生，不能证明外部纠错成功。E 式必须能够区分规则被证据改变与规则被自身输出强化，否则“回写”会成为任何反馈都适用的空泛名称。

三式还应互相限制概念伸缩。如果 S 一会儿指自然事件，一会儿指测量数值，一会儿又指公共制度，作者就可能在不同层之间移动，制造貌似完整的链条。分析前应固定单位与时间范围。必要时可以分层建模，但每一次跨层转换都应解释。不能拿某年事件对象的形成、几十年后的仪器变化和另一项目的制度改革拼成一个闭合证明。

联动的局部中断也能提供解释。对象区分已经改变而采样仍旧，可能使新问题受到旧资料分辨率限制；操作更新而分类规则未动，可能只是提高同一对象的测量精度；规则重

组后误差反而扩大，则需要检查新对应是否失去原有约束。这些结果分别指向资料瓶颈、稳定化和修订失败。它们不宜被合并成一种未完成状态，因为各自要求的下一步工作不同。

若要解释中断的原因，还需把时间顺序与行动理由接起来。采样升级早于新对象定义，便不能直接归为对象引导；两者同时改变，可能共享仪器升级或新的研究任务；对象定义先变而操作随后调整，也需要计划或使用记录说明连接。规则回写同样如此：修订文件应能指出何种新差异触发改变，以及改变怎样影响旧分类。时间线确定候选关系，行动材料再帮助区分各条因果路径。

特别需要注意的是共同原因。技术进步可能同时改变对象分辨率与操作方式，二者的相关变化未必意味着对象单向驱动操作。SDE 强调相互规定，有助于提醒这一点，但不能以相互作用之名回避因果识别。历史解释仍需比较时间顺序、行动理由和替代原因。

还可以检验框架的使用成本。若为了编码每个案例，研究者花费大量篇幅重复定义，却没有改善问题选择，那么统一性收益可能不足以抵消负担。相反，若简短的三向检查能够及时发现对象偷换或证据循环，即使不增加自然机制，也可能具有实际方法价值。评价应保留这种成本与收益关系。

因此，后续研究可以记录使用框架前后提出的问题，而不只比较最终成稿。最终文章经过润色以后，往往难以看出哪一项判断真正由框架引发。保留草稿、修改理由与被放弃建议，有助于追踪方法在实际思考中的作用。这与把作者事后感受当成证明不同，它为判断增量提供过程材料。

框架也可能因使用者不同而表现不同。领域专家本已熟悉的联系，对跨学科读者可能是重要提醒。教学效用与专家研究增量因此应分开测量，不能由其中一种成功代替另一种。

上述比较可以保留一项明确结果：水文状态依赖和辐射条件化响应，与成熟学科解释高度重合；版本案例则让两个技术领域的对象身份、操作选择与规则修订获得可比表达。前者为框架划定已有知识的边界，后者给出组织作用的实际位置。继续研究应集中于这些位置，而不把每个案例重新扩写为完整的三式叙事。

组织作用也有可检验的后果：它是否让跨领域读者更准确地定位错误，是否帮助研究者减少无关补证，是否让修订沿依赖关系更快到达使用者。三项效果分别需要理解任务、研究过程和版本传播的记录，不能由文本看起来整齐直接推定。

◆ 57.4 撤回怎样使框架继续发展

三方程在本章形成的作用，是把对象成立、后续操作选择与关系修订放到同一比较中。它把已有学科解释之间的接口显现出来，也暴露自身容易混淆的两件事：一个关系在材料中存在，和一个分析框架首次辨认了它。前者可以支持历史解释，后者才涉及框架增量。撤回后一主张，不会自动抹去前一关系；两者分开，才能保留有效分析而修正过强评价。

后续评估可以作三种结论。若框架提出了常规说明遗漏的、经独立材料支持的具体判别，可在该案例中保留局部增量。若它只是使原有知识更清楚地组织，可保留方法或教学价值，并撤回额外机制主张。若它使对象层次混乱、增加不可检验解释，应修订或放弃该用法。三种结果都能推动框架发展。

撤回不是把此前工作宣布为毫无价值。它可以把过大的命题缩小为更精确的命题。例如，从“所有科学变化都完成三式联动”，缩小为“某些对象转换通过可追踪操作影响规则修订”；从“关系分析是 SDE 独有”，缩小为“SDE 提供一种跨领域比较这些关系的方式”。精确的缩小，常比无边界的扩张更能支持后续研究。

如果同资料、同篇幅的成熟解释已经覆盖全部关键变化，本书应承认没有证明额外机制。这个条件不是附加礼貌，而是比较框架的失败入口。只有允许这种结果，下一章关于六路径与相位的讨论，才不会沦为把任何历史都排列进预设格子的工作。

第 58 章 六路径与三原理
比较转折，不替历史排队

58.1 路径不是人物心理

研究一项科学转折，可以从已经出现的对象进入，也可以从操作差异或特征关系进入。六路径提供的是这些入口与换路的组合方式。[M03][M04]但研究者今天选择的讲述顺序，不等于历史人物当时的思考顺序。若把两者混为一谈，框架就会把自己的叙述投射成历史事实。

一篇论文首先写理论、再写资料，可能只是论证组织；一份项目计划先列目标、再列设施，也不证明目标先于所有技术设想产生；一次警报修订的正式日期，更不等于问题第一次进入工作人员意识的时刻。公开文献能够确认某些表达与行为的时间，却很少完整保存全部心理活动。历史解释应尊重这一区别。

路径复审中，论文发表、测量执行与规则发布可以标定公开行为；作者回忆可以提示某个问题何时受到重视，却需要与当时记录互相对照；后来的分析则负责说明这些材料之间可建立何种联系。三类材料的作用不同。例如，修订通报能够证明用语已经改变，回忆可能解释为何认为旧用语不够，二者仍需使用记录才能说明改变是否解决了接收问题。区分材料身份，使动机、行动和效果各有对应。

海洋动力论文、地球系统科学纲领与航空警报制度，恰好提供三种不同的材料结构。论文通常集中呈现一个论证，纲领表达未来组织目标，警报制度则包含反复使用和修订。不能把它们都强行恢复成同一条“发现—验证—胜利”链。六路径若有价值，应帮助研究者看见不同材料的限制，而不是把限制抹平。

发表日期之所以重要，是因为它规定某种公开主张在何时能够被共同体回应；它并不囊括发现的全部时间。测量可能早于发表多年，概念可能经历多次私下讨论，正式规则也可能总结先前实践。思想史可以在资料允许时区分这些时间，但不能用后来回忆填满所有空白。保持时间层次，能避免把一个复杂转折写成某天突然完成的思想事件。

对路径而言，这意味着至少应保存材料时间、行为时间与叙述时间。材料时间指观测或事件发生，行为时间指研究者处理与提出主张，叙述时间则是后来的历史重建。三者经常不同。六路径是在第三条时间线上选择入口，然后用前两条时间线约束叙述；它不能反过来替前两条时间线安排顺序。

◆ 58.2 同一材料的两次进入

以第 43 章海洋动力问题为例，可以先从观测差异进入：不同位置的流速、温度或盐度怎样形成需要解释的分布？这个入口容易凸显采样稀疏、空间代表性与分类的问题。随后转向动力关系，研究者才会问哪些约束能够把分散记录组织成环流对象。D→E→S 在这里是本书的复审路线，不是对原作者心理过程的断言。

也可以先从动力约束进入。研究者考察在某些尺度与假设下，风应力、地球自转和海洋运动之间应当满足什么关系，再询问需要哪些观测量检验它。这个入口容易凸显方程适用范围、边界与尺度选择。它可能忽视测量记录本身的分类和误差，因而需要再回到实际资料。两条路线各自照亮问题的一面，也各自具有盲点。[C03]

若两次进入最终得到相同对象边界，说明该边界受到不同方向的约束。若结果不同，不能直接称为框架的多样性优势。应先检查两次分析是否使用相同材料、是否在相同尺度讨论、是否把理论变量与测量变量混为一谈。例如，一个入口讨论深度积分输运，另一个入口只看表层流速，两者得出不同结论并不构成真正冲突，因为对象已发生移动。

这类比较能够揭示常见的叙述偷换。作者从一幅直观图像开始，读者以为讨论的是局部可见流动；推导随后转到一个积分量，最后又把结论投回图像，仿佛两者完全同一。路径复审的价值，是让这些转换显露出来。若转换有明确模型与条件，它可以合理；若没有，漂亮的路线反而掩盖了推断缺口。

因此，同一材料的两次进入应保留各自首先提出的问题、使用的资料和最后成立的命题。比较的重点不是哪条路线更顺口，而是哪条路线暴露了对方遗漏的约束。六路径不应被用来增加六倍篇幅，而应被用来减少一个入口造成的盲区。

从动力进入海洋问题，还可以显露方程变量与可测量量之间的距离。理论关系可能涉及垂向积分或区域平均，观测却在离散深度和位置取得。研究者需要插值、积分或其他转

换，把资料连接到理论对象。从测量进入时，这些转换容易成为焦点；从方程进入时，它们可能被简写为数据处理。两次进入的互补，正体现于把被另一条路线压缩的步骤重新展开。

反过来，单从测量分布进入，容易把每一种可见差异都视为同等重要。动力约束则能够指出哪些组合与所研究机制有关，哪些可能只是局部扰动。它不是否定测量多样性，而是选择解释尺度。路径比较的目标应是让这种选择显露，并检验它是否损失了关键过程。

因此，最后的对象边界不是两条路线的平均值。若理论假设明显不适用于所选尺度，应修订理论应用；若测量不足以约束理论变量，应增加或改变观测。和谐性来自相互纠错，而不是强迫两条叙述在语言上达成一致。

◆ 58.3 相位可以没有发生

本书用 P1 标记 D 与 E 的张力推动 S 改变，用 P2 标记 S 与 E 的张力推动 D 改变，用 P3 标记 S 与 D 的张力推动 E 改变；对象、路线与关系的变化只是各相位的观察入口，还须追踪变化如何回写另外两维。若只有变化结果而缺少张力与回写证据，应保留为候选相位。这些分析代号不应与特征律、自由律、幸福律直接混同，也不是每个案例必须依次通过的三个历史阶段。某个对象可以长期稳定而测量路线改善；某种规则可以改变而新对象尚未形成；一项纲领也可能提出了关系重组，却迟迟未能落实。

地球系统科学纲领说明这一点。报告能够明确提出跨圈层、跨尺度的研究任务与组织方向。[C08][C13]这支持我们分析一种新的整体研究议程，却不能仅凭报告存在，就认定所有机构、数据与模型已经成功协同。P3 作为关系重组的计划可以被识别，P3 作为已经发生的稳定运行，则需要实施材料。计划与实现必须分开编码。

航空警报制度提供另一种情形。通知模板或责任分工改变，可能主要改善信息路线和使用关系。若自然危险分类没有发生实质改变，就不必强行赋予它一次 P1 对象革命。反过来，若某种此前没有被清楚区分的航空危害获得了专门定义，也不能因制度后来未立即完善而否定对象区分的意义。不同变化可以不同步。

同一事件中还可能存在重叠。例如，新的灰云危害分类要求新的观测与通知安排，操作变化又推动责任关系调整。此时可以讨论多个相位，但应分别提供证据，而不是由其中一个变化推断其余必然发生。重叠不等于不分层，联动也不等于同时。

停顿可以显示相位转换究竟卡在哪一环。资料格式不兼容，限制的是既定变量之间的交换；一个新分类只在局部可靠，限制的是对象的迁移范围；通知规则已经统一而接收者仍采用不同含义，限制的则是公共使用条件。三者都可能呈现为项目没有完成，却要求不同修订。按其具体失效位置比较，才能判断下一步需要接口工程、分类检验，还是语义与责任安排。

相位判断还受分析单位影响。从一个机构内部看是规则重组，从整个学科看可能只是局部维护。作者应先限定讨论尺度，不能为了提高创新性而在段落之间放大单位。一个局部变化有自己的价值，不需要借助全球革命的措辞才能成立。

在系统纲领中，可以把“提出共同研究对象”与“建立共同资料接口”分开。前者允许不同学科围绕整体地球系统提问，后者才使某些资料和模型能够实际接续。两者之间可能存在很长的工程与组织过程。只有当接口、变量定义和评价方式得到落实，整体性才从目标转化为运行能力。一个宏大的概念并不会自动克服数据格式、尺度和责任分工的差异。

航空预警则表现出另一种顺序：实际行动压力可能先推动通知路线改善，随后才使危害分类与研究任务更明确。这里的变化不必以一篇理论论文为起点。若按学术发表顺序独自讲述，就可能漏掉日常运行中的知识形成。不同材料类型因此要求不同证据入口，六路径的灵活性应服务于这种差异。

三相位的主次也可以随着时间改变。初期可能主要是建立一个对象，后期重点转向稳定操作，再后来才出现制度重组。分析者可以划分阶段，但划分依据应是实际变化，而不是预定周期。若某一阶段没有足够材料，应保留不确定范围，而非为叙述流畅补上一段必然过渡。

路径复审还需要警惕资料可得性造成的偏向。正式论文容易保存，日常维护、失败试验与口头协作较少留下记录。如果只依据易得材料，历史就可能过度表现为概念驱动，低

估操作与组织的作用。E 入口可以提醒研究者追查制度与接口材料，D 入口可以追查实际测量记录，但最终能说多强仍受档案限制。

同样，某种资料大量存在也不意味着它在因果上最重要。一个机构留下许多会议记录，不能据此认定会议是转折的主要动力。数量与作用必须分别判断。研究者需要寻找能够连接变化前后、排除替代解释的承重材料，而不是让档案丰富程度代替解释。

改变入口的收益，可以用原先未见的约束衡量。从对象进入时，研究者容易看见分类结果；转向 D，可以追查分类所用样本是否覆盖边缘情形；转向 E，则要问区分这些类别的特征在何种条件下还能共同成立。若路线停在资料缺口，缺口本身应被标出：缺的是样本、时间次序，还是特征对应。这样，路径提供了可接续的取证方向，而非一份必须填满的历史顺序表。

若不同入口持续得到互相矛盾的对象边界，最先需要修订的可能是分析单位，而不是历史材料。缩小到一个具体转折，再恢复其前后关系，往往比扩大框架解释范围更有效。路径的灵活性应服务于澄清，而非无限吸收矛盾。

异议记录应指出分歧发生在哪一项判据。若一方把标签名称改变计作对象改变，另一方要求可辨属性改变，两者首先需要统一对象判据；若判据相同而对材料解读不同，争议才进入历史取证。把这两类分歧分开，可以避免增加史料以后仍在使用不同规则，也避免仅靠修改术语掩盖真正的证据冲突。

◆ 58.4 比较表怎样接受他人复核

一项相位判断至少应附带变化前后材料、连接两者的证据，以及可能推翻判断的观察。若说对象发生了变化，就要说明新旧对象在哪些可辨属性或操作上不同；若说路线改变，就要有研究计划或实际操作的变化；若说关系重组，就要说明哪些特征或规则的相互规定不同于过去。只给抽象标签，别人无法复核。

不同编码者的分歧应被保存。某人可能把一次变化理解为对象重建，另一人认为只是精度改善。解决办法不是用理论权威压过分歧，而是比较双方采用的单位与判据。若争议来自尺度不同，可以分层表达；若来自材料不足，就保留未定；若来自概念边界模糊，则应修订框架。

还应事先规定哪些案例无法归类。比如只有一篇宣言，没有实施记录，就不能判定完成回写；只有结果差异，没有过程资料，就不能确定路线为何改变；只有术语变化，没有对象边界变化，就不能认定 P1。不能归类并非研究失败，它防止框架把未知吸收成支持。

本章的路径比较属于作者重构，独立复核可以检验它是否稳定。复核者应先固定分析单位，再分别记录对象属性、差异组织和特征关系的变化；无法确认的方向保持未定。若两位分析者都只能确认名称变化，就没有足够材料判为对象重建；若能共同指出操作怎样改变了可辨属性，才形成较强的相位依据。这样的复核把分歧落在具体材料上，也给本书的分类留下可修改的位置。

复核还可以采用交换入口的办法：让两位分析者先分别从对象与操作进入同一材料，再公开比较他们遗漏的部分。若最终差异来自资料选择，就补齐资料；若来自概念尺度，就统一或分层；若仍然存在实质分歧，就把争议保留为结果。这种设计使多路径成为发现盲区的工具，而非把所有人引向同一个预设结论。

统一判据以后，结论仍可不同。海洋案例需要比较测量量与积分量的对应，航空案例需要比较危害信息与使用条件，系统纲领则需要比较目标与实际接口。相同的相位语言不能取代这三种证据任务。六路径的收获在于让不同入口彼此查漏，分析相位的收获在于说明变化实际落在哪一层。下一章把这种分层比较集中到样品年龄与板块对象两项承重论证上。

第 59 章 两项承重论证
样品年龄与板块对象的形成

59.1 年龄案先分清事件身份

一个测得的年龄究竟是哪一个事件的年龄？这是煤山案例的承重问题。样品具有采集位置，矿物具有形成与后期经历，地层具有沉积次序，地质事件则需要另外定义。把这些层次直接合并，会让数值精度掩盖对象身份的含混。第 13 章和第 15 章已经分别讨论矿物时钟与地层界线，本章不重讲技术故事，而把它们放到解释比较中。

煤山 GSSP 的界线身份，与一次灭绝过程的起止范围不是同一对象。国际地层标准规定了一个可对比的界线位置，定年研究则通过样品、层位与模型估计相关事件的时间。[K01][K02]标准的位置不能由某一个数值直接替代，灭绝区间也不能因为邻近标准界线就被压成同一瞬间。只有先分清事件，年龄差异才有可解释的对象。

常规定年学已经区分多种年龄不一致：矿物可能具有不同形成经历，封闭系统可能受到后期扰动，结晶与沉积之间可能存在时间差，分析误差与模型假设也会影响估计。每一解释指向不同检查：前两项涉及颗粒与保存状态，第三项涉及地层关系，后两项涉及测量和推断程序。发生学比较应继承这些区分，继续追问某种修订是否改变目标事件、取样位置或使用年龄的方式。

SDE 可以进一步组织一个问题：某个年龄身份的修订，是否改变了后续采样与事件定义？例如，研究者是否因为发现样品代表的不是目标事件，而改变层位选择或分类？这一问题有价值，但要证明它增加了判断，需要研究计划、样品记录或原论文中的具体说明。仅仅知道年龄后来更精确，不足以证明对象身份发生重建。

Burgess 等人的研究及其勘误，使版本关系直接进入年龄判断。[K02][K03]引用修订后的结果，需要说明改动作用于哪些数值或推断，不能把两个版本当作两项独立支持。本章沿用原研究讨论其论证结构，不给出新的单颗粒复算结果。比较重点是三个连接：同位素数据如何支持矿物年龄，矿物年龄如何约束层位，层位又在何种沉积与事件定义下约束持续时间。任何一处连接改变，都应沿链检查其影响范围。

补证可以沿这三个连接分流。若修订仅来自分析误差，先检查数值区间怎样改变；若矿物不再代表预期沉积时刻，便要重审样品与层位的对应；若目标由界线位置改为灭绝持续过程，则取样设计本身也可能需要调整。三种情况分别改变精度、时间归属和研究对象。把修订定位以后，才能判断 SDE 提出的是已有流程中的说明，还是需要另外追踪的对象—操作联系。

可以先写出年龄案的基线短论证。火山灰层中的锆石提供结晶时间的约束，样品处理与同位素测量决定年龄估计的可靠性；层位关系使这些年龄能够约束沉积序列；层间事件的年龄还依赖沉积解释与插值条件。因此，更精确的矿物年龄可以缩小某些时间范围，但不能单独消除层位与事件对应的歧义。这个解释已经同时包含对象、操作与条件，并非把年龄当作矿物内部自带的标签。

再写出 SDE 短论证。目标事件作为 S，依赖矿物测量与层位比较组成的 D，以及结晶、沉积、保存与事件定义相互规定的 E。若目标从地层界线改变为灭绝持续过程，研究者需要选择能够约束过程起止的层位，而不能只围绕一个标准点取样；若新增资料显示沉积存在凝缩或不连续，时间插值的条件就应改变。这段解释的实质，是把对象选择如何改变操作、操作结果如何限制时间关系放在同一链条中。

两段论证的共同部分很大：都拒绝矿物年龄直接等于事件年龄，都要求处理层位与沉积条件，都允许新资料修改模型。SDE 在本书目前的证据范围内，并没有给出不同的年龄数值或额外排除一个地质机制。它增加的是比较焦点，使读者检查对象、操作与条件的双向连接是否在叙述中被保留。只有找到基线遗漏而 SDE 实际指出的具体连接，才能再讨论解释增量。

这个结论具有建设性。它告诉后续研究者，不必把精力用于证明常规定年学忽视关系，而应查找关系在哪个使用或传播环节被简化。例如，一个条件化区间在转述中是否被压成单一时间点，标准界线与灭绝峰值是否被误当同一事件。这些对象偷换可以被具体文本与数据检查，远比泛称旧本体论更容易获得判别。

◆ 59.2 板块案先恢复证据到达次序

板块对象的形成不能写成一张今天已经完整的图在历史上逐步显影。1963 年磁异常解释、1965 年转换断层概念与 1967 年震源机制检验，在发表时间上具有先后。[K04]

[K05][K06]后来的联合约束不能倒投为早期假说已经拥有的证据。恢复这种到达次序，是理解风险、分歧与创造性的前提。

Vine 与 Matthews 把洋底扩张与地磁倒转联系起来，为磁异常提供了一种可以产生空间模式的解释。但磁信号还受磁化、几何与测线条件影响，不能把条带外观直接当作完整板块运动证明。早期假说的力量，在于它把若干差异组织成能够继续检验的关系，而不是当时已经消除了所有其他可能。

Wilson 的转换断层概念进一步改变了对错列洋脊之间断裂关系的理解。关键不只是给一条断裂换名字，而是提出不同的运动几何与可检验预期。Sykes 随后利用地震机制与空间分布检验这种预期，为转换运动提供独立约束。[K06]思想史应把预测与后来的检验分别放在各自日期，才能显示一项新关系如何承担失败风险。

在这里，对象的规定逐渐超出单一海底地貌。磁异常、断层运动、地震分布与相对运动几何相互限制，使一种可操作的运动体系获得身份。所谓新板块世界的形成，不是古老岩石圈在二十世纪才开始运动，而是这些特征在研究共同体的测量、推演与检验中，形成新的可辨认聚合体系统。主体能力与对象规定共同改变，才使新的问题能够被稳定提出。

成熟科学史同样以假说、预测、多证据收敛和研究实践解释板块转折。SDE 比较的切入点，应是某个特征怎样改变了其他特征的解释资格：错列地貌本身提供形态信息，加入相对运动几何后，断裂的活动区段与滑动方向获得特定预期；震源机制随后检验这些预期。若标准叙述已经完整表达这一转换，双方就在该处重合。比较重点由有无关系，转为关系何时获得了可检验内容。

因此，研究应追查对象形成之后是否改变测线选择、震源分析或边界分类，并区分这类改变的原因。后来的研究更密集，可能来自理论引导，也可能来自仪器发展或研究资源增加。要建立回写链，需要证明具体对象判断与具体操作之间的联系，而不是把时间先后直接视为因果。

本章重建的是发表节点与原始论证之间的连接，范围限于已核材料。进一步检验磁异常解释，需要给定磁化、几何与测线条件进行正演；进一步检验震源机制，则需要逐事件处理方向约定、位置和解的不确定性。两类数据工作尚未在本书重新执行，它们可能修正局部判断，也可能保留原结论。历史重建因而可以说明不同证据为何相关，独立数值验证仍须分别沿各自的操作链推进。

板块案的基线短论证可以这样表述：扩张与倒转的结合为海底磁异常提供空间预期，转换断层几何为运动方向和活动位置提供不同于简单错断的预期，震源机制与分布则对这些预期施加检验。几类证据并非重复描述同一图像，而是限制同一个运动体系的不同部分。体系因此获得强于单一地貌类比的支持。

SDE 短论证把同一过程写成对象规定的共同形成。条带的年代含义依赖扩张与磁化关系，断裂的类型身份依赖边界之间的运动联系，地震机制又使某种运动方向能够被检验。新的运动对象一旦稳定，研究者可以据它重新选择测线、分类断裂并寻找不符合预期的事件。这里 S 不是一幅静态地图，而是具有可操作约束的体系；D 与 E 也不是地图之外的附加条件。

两种解释再次具有明显重叠。多证据约束、预测与检验、对象对后续研究的引导，都不属于 SDE 独占内容。实际差别在于叙述重心：基线强调假说支持如何增强，SDE 强调可辨认对象与操作能力如何共同获得规定。重心差别可以改善某类思想史问题的表达，却尚不足以证明新的经验机制。若要提升主张，必须展示后者确实找到前者未覆盖的可检验联系。

◆ 59.3 让两种解释面对同一问题

两项案例应共同接受同一组具体问题。第一，当前对象究竟是什么？年龄案是某矿物、某层位还是某事件的年龄；板块案是地貌条带、断裂类型还是相对运动体系。第二，什么新资料会迫使对象修订？第三，修订会怎样改变下一步操作？第四，哪些后续结果会否定预期联系？这些问题不依赖 SDE 标签，因而能够用于公平比较。

对于年龄案，可以让常规定年学解释与 SDE 解释分别提出下一步所需资料。双方都可能要求检查单颗粒数据、样品处理、层位关系与勘误。若建议完全重合，就应承认没有显示操作增量。若 SDE 另外指出一个未被注意的事件身份错置，则应检验这一错置是否真实存在，而不能因为建议措辞新颖就给予贡献。

对于板块案，可以比较双方如何处理反常资料。若某组震源机制不符合预测，常规解释会检查资料质量、几何假设与替代构造。SDE 解释也必须说明何时应修改对象，而不是把所有反例都解释为复杂纠缠。一个框架如果不能指明哪一种结果会迫使撤回，就没有比成熟方法承担更多风险。

比较还应控制后见之明。向评阅者提供 1963 年证据包时，不应混入 1967 年的检验；评价 2001 年标准界线时，也不能默认后来高精度定年已经完成。分别冻结当时可得材料，可以检查框架是否真的帮助选择下一步，而不是在知道结局后重排故事。这样的设计更接近思想史中的实际不确定性。

此外，解释长度应受到约束。一个框架用数倍篇幅列出所有可能联系，可能比简短叙述看起来全面，但这不能说明它更有区分力。应要求双方在相当篇幅内回答同样问题，并标出每个判断的证据。评估者才有机会区分实质覆盖与语言扩张。

独立比较可把各组建议去掉框架名称，再检查其材料对象与预期结果。两组都要求核对单颗粒数据，属于操作重合；只有一组提出检查引用中事件身份的变化，属于候选新增任务；这项任务若发现实质错置，才获得结果支持。板块案例也可同样记录新增检查是否改变异常的解释。这里提出的是评价设计，尚无独立评阅结果。它以可追查的判断差异为单位，使篇幅增加与实际增量能够分开。

两案还有一个值得保留的差异。年龄案的对象修订经常表现为把原先混在一起的时间身份拆开，板块案则更多表现为把原先分散的特征组织成共同运动约束。前者通过区分获得可靠性，后者通过联结获得解释力。若统一框架只强调聚合，就会低估年龄案中的分离；若只强调分类，又难以表达板块案的跨特征结合。发生学必须同时容纳拆分与重组，不能把新世界发生简化为联系越来越多。

这一差异也影响后续操作。年龄案优先追查哪个样品代表哪个事件，板块案优先追查不同信号是否服从同一几何约束。相同的三式语言在两案中指向不同工作。统一性的价值在于保留可比问题，多样性的价值在于不抹去答案的结构差异。

还可以从失败后果比较两案。年龄模型中一项插值条件失效，可能主要改变区间估计；板块运动几何的一项关键预测失效，则可能要求检查边界分类或整个局部运动关系。两者的修订半径不同，不能把所有反例都赋予相同理论重量。SDE 若强调回写，应进一步问回写扩散到哪些层次、在哪些条件下停止。

这种修订半径可以由依赖关系具体判断。哪些结论共享同一层位假设，哪些运动解释共享同一边界条件，决定了一个错误会影响多少后续判断。比较由此获得实质内容：不仅问知识有没有改变，还问改变怎样传播、哪些旧成果能够保留。它连接两案，又不消除它们在证据结构上的差别。

两案比较得到的直接结果，是修订方向与传播范围可以分开考察。年龄案常从混合的时间身份中拆出不同事件，板块案常把分散信号纳入共同几何；但两者都要继续问，一处修改会牵动哪些共享假设。方向说明对象怎样改变，范围说明改变怎样进入后续研究。这个区别比把两案统一归为理论胜利更有内容，也为下一节的保留、降级和撤回提供依据。

◆ 59.4 保留、降级或撤回

两项承重论证目前保留了不同的比较成果。年龄案明确了矿物年龄、层位约束与事件时间的分层，成对论证显示这一核心已由常规定年学处理；SDE 将关注点放到身份修订对后续操作的影响。板块案则显示多种证据怎样共同规定一个运动对象，SDE 与假说—检验叙述在此有大幅重合，差别主要在组织重点。两案均未据此证明额外机制优势，却已经给出可继续检查的具体接口。

可保留的局部命题，应具有明确范围：在何种资料条件下，某个对象修订通过什么操作改变了共享规则。需要降级的命题，则包括缺乏操作记录的完整回写、未与强竞争解释比较的优越性，以及把时间先后当成因果的叙述。应撤回的，是与明确原典或独立资料冲突的判断。

这种处理不会削弱本书关于科学世界发生的总体问题，反而使它获得可研究的尺度。新世界并不是一个无需证据的宏大修辞，而是由可辨认对象、操作能力与特征关系共同变化构成的具体主张。它越具体，就越可能被支持，也越可能被限制。

本章因此把框架本身放回研究对象之中。三方程、六路径与相位分析不是站在历史之外的裁判，而是需要由材料持续修订的研究工具。保留什么、降级什么、撤回什么，构成思想史写作自己的回写责任。最后一章将从这些边界提出能够继续执行的问题。

第 60 章 开放的地球
下一轮问题从哪里发生

60.1 先保留全书未解决的东西

一本思想史的结尾，不必把所有问题压成一个答案。地球科学之所以能够持续发展，正因为对象边界、证据联系与研究能力始终存在可以进一步追问的张力。本书已经讨论从地层到板块、从深部到气候、从遥感到算法的多种对象，但这些对象并未因为进入同一本书就获得相同确定性。开放的地球，首先要求保留差异，而不是把差异当成统稿时应当消除的杂音。

流域边界的困难，部分来自地表划分与实际水交换之间的关系；古气候对齐的困难，涉及年代模型与待解释周期是否共享约束；数字孪生的困难，在于模拟范围与应用任务如何对应；AI 目录的困难，则在于检测能力与自然事件率怎样分开。这些问题都与对象身份有关，却不是同一种技术问题，不能靠一个统一公式直接求解。

还要把三类未完成分开。第一类是研究尚未执行，例如已有公开资料却没有完成复算；第二类是当前材料不足以区分竞争解释，例如不同机制留下相近痕迹；第三类是工具或观察条件暂时限制了取证。第一类需要工作安排，第二类需要寻找判别性资料，第三类可能需要发展新的测量能力。把三者都称为复杂性，会使下一步行动失去方向。

这些缺口可以按改变判断所需的材料来组织。流域边界争议需要把水交换与收支残差相接，年代争议需要新的时间约束，应用错配需要变量到任务指标的转换记录，目录争议需要独立候选复核。明确材料入口以后，研究者才能判断一个问题可以立即分析，还是需要等待观测、协作或新的操作条件。

四组未决还具有不同的失败时间。流域边界问题可能在一次特定水文条件下暴露，年代对齐问题可能要等新定年资料出现，数字孪生应用则可能在实际使用后才显出尺度错配，AI 目录的错误可能在独立复核时立即被发现。研究流程应顺应这些时间，不必要求所有问题在同一轮写作中获得同等验证。

研究安排还应考虑哪些工作可以并行、哪些必须先后完成。古气候测量序列尚未整理时，直接比较年代方案会混入数据版本差异；目录事件尚未关联时，讨论迁移机制会提前使用不稳定位置。相反，资料身份与方法条件清楚之后，若干替代解释可以并行检验。先把依赖关系理清，比平均扩展每一条路线更有助于把有限工作转成可累积结果。

◆ 60.2 四项可接续的小研究

第一项研究可以围绕一个公开资料较完整的小流域，比较地形划分边界与水文响应边界。研究者固定时间范围和观测口径，检查地表汇流定义是否足以解释出口记录，并把地下交换或人工调水作为需要额外证据的候选。输出应是边界条件与残差结构的对应，而非直接宣布所有流域边界都失效。如果差异能够由观测误差或储量变化解释，就应保留原边界在该任务中的适用性。

第二项研究可以对一组公开古气候记录进行年代敏感性分析。先保留原始测量序列，再在有依据的年代范围内改变时间映射，观察关键相位、周期或事件先后是否稳定。输出不是挑选最漂亮的对齐图，而是说明结论对年代处理依赖多少。如果支持关系只在一种强约束调谐下出现，就应缩小其独立证据地位。若多个合理年代方案都保留关系，结论才更稳健。

第三项研究针对数字孪生的一项具体应用，而不是评估整个数字地球。选择公开可取得的产品与任务说明，建立变量、时间尺度、空间尺度和使用指标之间的映射，再检查评价资料是否覆盖真正影响判断的特征。输出可以是一份范围审查与适用性报告。若平台输出与用户需求只有展示对应而没有物理或统计连接，就应标明转换缺口；若连接清楚，也仍需区分模型能力与实施效果。

第四项研究针对一段固定的连续地震记录，比较两套检测配置，并对新增与遗漏候选作独立抽样复核。保留候选、到时、关联与定位各层结果，使错误可以定位。输出应包括哪些事件新增、哪些被撤回、哪些仍未确定，以及数量变化对自然活动解释有何影响。如果新增清单不能通过独立约束，能力改善的主张就必须收缩。

四项研究规模都有限，却分别触及空间边界、时间身份、应用转换和事件资格。它们的共同价值，在于把宏大框架压到可以改变结论的局部比较。有限不等于琐碎：一个边界修订，可能改变后续所有统计的含义；一项年代依赖，可能改变整条因果链的顺序。

四项小研究可以采用共同的结果格式，但不采用共同的结论模板。每项保留对象定义、固定条件、实际比较、替代解释和修订结果，使他人能够跨领域阅读；而具体误差、变量与失败标准则由各学科规定。水量平衡不能被地震定位残差替代，年代敏感性也不同于模拟情景差异。共同格式服务于交流，不能压平科学内容。

它们还可以形成一组稳定性结果：边界在所选水文条件下无需修订，年代关系在合理扰动下保持，数字产品能够满足特定任务，新增地震事件仍支持既有机制。这些并非同一种负结果。前两项限制修订的必要范围，第三项建立用途资格，第四项扩展旧解释受到检验的材料。分别记录其条件，才能让稳定化成为可使用的知识，而非把没有发现新机制笼统归为没有收获。

◆ 60.3 知识增长怎样改变问题

知识增长并不总是给旧问题添加更精确的答案。新的资料有时会改变问题本身。一个原本被视为单一事件的过程，可能分解为几个阶段；一个原本局部的异常，可能进入更大尺度的联系；一个原本足够稳定的类别，在新观测分辨率下可能显出内部差异。研究者因此需要保存问题版本，而不仅是答案版本。

这种变化也会改变共同体分工。新的遥感档案要求标定、算法与地表知识相接合；数字孪生要求模型生产者与应用使用者共同规定指标；AI 地震目录需要波形识别、关联定位与地质解释之间明确交接。知识不是在一个孤立主体内完成后再向外传播，它在这些互动中形成可操作身份，也在交接失误中暴露边界。

但旧知识不会因为新问题出现而全部死亡。某些测量仍然有效，只是用途改变；某些分类在粗尺度仍适用，在细尺度需要拆分；某些模型保留局部解释力，却不再承担普遍任务。思想史若只展示胜利者名单，就会漏掉这种层级迁移。保存仍有效的旧成果，与记录被放弃的路线，同样重要。

SDE 对此提出的任务，是检查显露、差异序列与特征纠缠的变化如何相接。以本章拟议的年代研究为例，目标事件 S 规定哪些先后关系值得检验；不同年代方案组织出 D，使原先稳定的相位关系显出敏感性；保存与沉积条件的检查又限制哪些时间映射可接受，改变 E 中的证据对应。若这组检验只缩小误差而未改变对象，结果便是稳定化；若某些记录必须归到不同事件，才涉及更深的对象修订。

最值得保留的研究态度，是让新增能力同时带来新增约束。能看得更细，就更明确误差尺度；能算得更远，就更清楚情景边界；能自动识别更多，就更认真保存被否定的候选。能力与责任共同增长，才使知识扩展不至于变成无法追查的输出堆积。

未来研究还需要保存问题被放弃的理由。某条路线可能因为资料不可得暂时停止，也可能因为证据反驳而被放弃，两者不能混同。前者可以在条件改变后恢复，后者则需要新的理由才能重新提出。若只保存成功项目，后来者可能重复已经失败的工作；若把暂缓误写成反证，又可能过早关闭有价值的问题。

一个开放的研究共同体，应让这种不同结局具有可见身份。知识的生成、传播与退出有效范围，都可以留下可追查记录。这样，思想史不再只是赞颂创造者的作品，也成为未来研究者判断何处值得继续、何处应当改道的公共资源。

后来者应能从具体材料进入这些研究，而无须先采纳整套理论语言。对年代方案的不同意，可以落实为某项时间约束；对目录的不同意，可以落实为某组波形与关联；对数字应用的不同意，可以落实为变量到指标的映射。把异议定位到这些地方，才可能决定需要复算、补测还是重设问题。公共研究起点由此具有可执行的含义。

当一个总问题被拆成若干能够分别取证的问题，研究范围虽缩小，判断却可能更稳。某项数字产品是否适用，不再要求先裁决整个数字孪生；某种目录是否支持迁移，也不再要求先判断全部 AI 能力。局部问题保留与整体任务的联系，同时给证据留下真正能够改变结果的位置。

◆ 60.4 把最后一页交给下一轮核验

全书最后仍回到同一组具体问题：什么显露获得了科学身份，哪些差异使它能够被辨认，哪些特征关系规定了它的意义；对象成立以后，研究者怎样改变操作，新的操作又怎

样保留或修订关系。问题相同，并不要求答案相同。地层、板块、天气与数字对象各有尺度、证据与失败方式。

本书的历史结论，应由相应原典与观测约束承担；方法贡献，应由是否帮助组织和改进研究承担；未来假说，则应保留为等待比较的命题。这三种身份不能相互借贷。一个富有启发的框架不能替代史料，一项历史成功也不能自动证明框架普遍优越。

下一轮核验可能加强本书某些局部判断，也可能要求缩小或撤回它们。若新材料表明某次所谓对象革命只是精度维护，应当修订；若常规解释已覆盖全部关键联系，应承认组织价值与机制增量的差别；若发现此前遗漏的实际回写，则可以把论证推进到更具体的位置。

地球科学思想史因此没有在这里关闭。它留下的不是一个已经替地球说完话的理论，而是一组能够继续与岩石、仪器、模型、档案和行动相互检验的问题。最后一页的意义，在于下一位研究者能够知道从哪里接续，知道什么结果会使我们改变判断，并拥有这样做的材料与自由。

结语 在稳定的知识中保留继续发生的能力

结语一 新世界发生的可保留结论

回到开篇那块岩石，我们现在能够把最初的年龄问题问得更清楚。我们不仅要知道一个数字，还要知道它指向什么事件，经过什么操作，依赖哪些相互规定的条件，以及在什么尺度上能够与别的记录比较。问题没有因为这些追问而失去答案，答案反而获得了可以被使用、被检查和被修订的身份。

全书考察的对象并不相同。一个地层界线的成立，需要定位、标志和跨地区相关；一个板块运动关系的成立，需要几何、方向和不同观测之间的约束；深部结构主要通过间接记录和材料知识获得限制；天气状态则要在持续更新的观测与模型之间形成；数字地球还必须交代产品版本、计算条件和任务接口。它们不能被简化为同一种数据处理，却都要求说明材料怎样成为具有确定责任的证据。

因此，可以保留的第一个结论是：科学属性的成立包含具体的辨认和比较条件。这个判断并不抹去自然约束。恰恰因为自然过程不会按研究者的意愿排列，样品、波形和观测才能使模型受挫。发生学解释需要同时容纳材料的抵抗和对象规定性的形成。如果只剩人的命名，就无法解释不同方案为什么受到不同程度的支持；如果只剩现成对象，就无法解释科学为何需要不断创造新的区分与操作。

第二个结论是，新的对象关系可以改变后续研究的选择。矿物被区分为不同事件的记录以后，取样不再只是增加数量；海底断裂被赋予不同运动身份以后，地震活动的区段和方向成为判别问题；天气分析场的形成使后续预报拥有可操作的初始状态。这里的关键不是成果被广泛谈论，而是它能够重新安排要比较什么、在哪里比较、用什么误差标准判断。

第三个结论是，研究结果的使用会留下新的条件。保存的样品、经过校准的资料、修订的分类和处理版本，可能改变下一轮研究能够调用的资源及其意义。但回写不是每个案例都天然完备的闭环。模型具有引导操作的潜力，与档案证明它确实引导过某项操作，是不同命题。本书在材料不足之处保留这一区别，不把方法上的可能性补写成已经发生的历史。

把这三个结论合在一起，才得到本书所说的新地球科学世界。它不是一个脱离自然的想象空间，也不是所有研究者共享完全相同经验的集合，而是对象属性、辨识能力、操作方式和共同检验形成的可延续体系。新的世界允许此前不能稳定提出的问题进入研究，也允许某些原本看似自然的分类失去有效性。其变化既涉及知道什么，也涉及能够怎样知道和继续做什么。

这些结论的普遍程度仍有边界。本书选取的是具有较明确文献与技术支点的案例，并未统计全部地球科学研究。我们不能根据若干成功转折，断言所有科学变化都必定经历同一顺序，也不能把未被选入的渐进研究视为理论上次要。大量维护、校准与负结果，正是使较大转折可以被辨认的基础。宏观叙事若看不见它们，便会夸大断裂而低估持续工作。

SDE 将这些责任放在同一分析中：显露需要差异序列与特征纠缠的支持，形成的对象又可能改变操作及其条件。这样的统一语言为跨领域比较提供了起点。其独立解释增量仍须逐案评估，尤其要检查它是否发现了既有分析遗漏的连接，或提出了不同的、可执行的复核任务。

◆ 结语二 回写的成功与失败

一项知识能够影响后来研究，有时因为它足够可靠，有时因为它便于使用，也可能因为制度和工具已经围绕它组织起来。这些原因可以交织，却不能互相替代。被采用不等于真实，无人采用也不自动证明错误；长期维护的标准既可能保存共同工作的连续性，也可能使新的问题难以进入。思想史需要追踪具体的采用机制，而不以成功者的地位代替对证据的判断。

回写成功的一种情形，是新关系提供了可判别的下一步操作。某个模型不仅解释旧资料，还说明在哪些位置、方向或时间尺度上应出现不同结果。后来研究若实际开展这种检验，便能够把模型承诺与新的材料连接。即使检验失败，回写也可能在研究意义上成功，因为它使争论从模糊分歧进入可以纠正的判断。需要区分模型预测的成功与研究过程变得更可检验的成功。

另一种情形是，已有成果降低了重复劳动的负担。共同参考模型、标准化记录和经验证的数据产品，使使用者无需每次重建全部基础。这种封装为新的问题释放时间，但它的

可靠性来自此前的使用、检测和持续维护，而不是来自封装本身。一个接口越方便，越需要清楚说明适用对象、尺度和版本，否则错误会随着便利一起扩散。

失败也有不同形态。新模型可能没有改变任何实际操作，只在回顾文章中获得新的名称；新操作可能产生了更多资料，却没有改善关键问题的判别；新标准可能保留了形式上的一致，实际使用者却以不同方式理解它。还有一种失败发生在反馈被选择性接收时：符合预期的材料进入更新，不符合预期的材料被排除，却没有留下可查的理由。

这说明闭环本身不是可靠性的保证。一个不断自我强化的系统也可以闭合：旧分类筛选样本，筛选后的样本支持旧分类，新的使用又扩大同样的筛选。要区分这样的自我保护与有价值的回写，必须保留能够打断链条的检查位置。独立的证据来源、公开的排除规则、未参与调整的检验资料和版本可追踪性，都可能承担这种作用，但它们是否真正独立，还要具体分析。

回写还存在尺度问题。一个地区的校准更新，未必能够直接推广到全球；一种事件目录的精化，未必改变深部动力机制；预报技巧的改善，也不自动等于行动后果改善。跨越每一级都需要新的连接。若不说明这些连接，发生学叙事容易从一个局部成功跳到整体世界已经重构，反而重演它原本希望避免的推断越界。

合理的保守也是知识发生的一部分。新资料与旧模型相容时，保留旧模型并不意味着拒绝创造；某个解释复杂而没有增加判别能力时，选择较简单的描述也可能更负责任。我们不需要让每个章节都以革命结束。能够辨认何时应当维护、何时需要修复、何时值得重新定义对象，才是科学思想史能够培养的判断力。

从这里看，统一性、多样性与和谐性具有具体含义。统一性让不同研究能够相互解释和检查；多样性保存材料、方法和任务的差别；和谐性要求这些差别之间的连接不互相破坏，而能形成合理分工。若为了统一而抹去所有误差类型和尺度边界，体系只剩表面整齐；若为了多样而拒绝共同判断，材料又无法进入公共研究。知识的持续发生需要在两者之间建立能够修订的联系。

还可以设置一种更朴素的检验：当研究条件改变时，原有判断能否说明自己应当怎样改变？如果化石保存条件恶化，相关的信心就不应保持原样；如果测站的空间分布发生显著变化，区域平均的可比性就需要重新检查；如果软件重处理改变了历史资料的数值，趋势判断就应说明哪些差异来自处理，哪些可能来自自然变化。能够给出这些变化责任的知识，比只会维护同一个结论的知识更值得信任。

这类检验不要求每个人成为全部领域的专家。分层判断使不同使用者承担不同责任：专业人员处理原始分析与校准，跨领域使用者检查对象、尺度和版本，行动者判断结果是否适合自己的任务。分工不能消灭责任，但能够让责任变得可执行。其前提是层与层之间保留说明和反馈的通道，不把上一层的便利当成下一层可以忽略条件的理由。

从研究能力的发生看，这也是人的主体性可以继续发展的地方。主体性不只表现为独立完成全部计算，更表现为知道何时可以信任封装、何时需要提出新问题、何时应组织复核。人与仪器、模型及公共知识形成复合能力之后，个人并没有因此自动失去判断；但如果所有接口只允许接受结果而不允许追问条件，判断能力就可能逐步萎缩。

因此，知识传播的设计本身值得成为后续研究对象。一个术语怎样进入课堂，一幅图如何被教材裁剪，一组数据如何被软件默认调用，都可能决定后来的人还看得见哪些问题。思想史与教育在这里相遇：讲清一个成功结论固然重要，让学习者能够辨认结论生成中的关键选择，同样影响下一代是否具有继续创造的能力。

◆ 结语三 留给下一轮研究的问题

本书最重要的未完成工作，不是再给每个案例增添一套术语，而是把关键连接推进到更细的材料层级。煤山案例需要持续核对样品、层位、误差类别与事件归属，并把不同年代模型的影响明确分开。讨论原论文的推断结构，不等于已经完成单颗粒数据的重新计算。只有真正执行复算，才能报告新计算对事件先后或时间窗口造成的具体改变。

板块案例同样有继续深入的空间。磁异常模型、转换断层几何和震源机制检验之间的联系，可以通过逐事件、逐图件的材料进一步重建。原摘要支持的样本结论与逐事件重新评定，具有不同证明责任。尤其需要注意样本选择、方向约定、位置不确定性和后来的修订，不能仅凭一项历史总结便宣称所有独立性问题已经解决。

更广泛的接受史仍需要操作记录。论文被引用，未必意味着某项调查路线已经改变；某种图例成为标准，也不能单独解释使用者怎样学会它。航次计划、采样记录、仪器配置、培训文件和争论中的具体回应，能够把结果与后续行动之间的联系写得更清楚。这些材料会使思想史从只读重要论文，进一步进入知识如何被实际维持的过程。

跨案例比较也需要更严格的设计。若要检验 SDE 是否改善研究判断，可以让不同分析者使用同一材料包，分别采用不同框架，在不知道后续结果时提出补证任务和失败条

件，再比较各自遗漏了什么。这样的研究不能只按术语丰富度评分，也不能把更加冗长的回答当作更好的解释。需要预先确定评价对象，例如错误定位、材料选择、因果时序或可检验预测。

对数字地球与人工智能，下一步问题尤其具体。一个模型在原数据集上表现良好，迁移到新地区、新仪器或新版本之后，哪些能力仍然保持？自动生成的事件目录怎样保留不确定性和人工修订记录？当使用者只看到封装结果时，什么接口能够帮助他们在必要时打开其形成条件？这些问题同时涉及科学对象、计算操作和公共使用，不能只由算法分数回答。

本书也留下一个哲学上的开放问题：当对象身份、操作和辨识能力共同变化时，如何描述不同科学世界之间仍然保持的联系？不能简单说它们完全不可比较，因为旧记录常能被重新使用；也不能说它们只是对同一个不变对象换了名称，因为新增区分确实改变研究能力。更细致的比较应当寻找哪些关系可保留、哪些关系需要翻译、哪些对象只有在新条件下才获得明确身份。

这种开放性要求研究者愿意修订自己的框架。若某项 SDE 分析无法增加问题的清晰度，或把已经成熟的领域方法重新命名却不增加可查责任，就应当缩小其主张。反之，若它帮助发现一个被遗漏的共同校准、一次未经说明的对象转换或一条可以检验的回写链，就应继续把这一增量推进到材料和操作，而不是停在哲学赞语中。

上述问题可以从一个范围有限的研究开始：选取同一资料产品的两个版本，保存原始输入、处理参数和修订说明，逐项定位结论改变的位置。若资料、算法和分类同时变化，就分别检查各项改变的影响；无法分离的部分则明确保留。这样的版本比较能把回写机制落实为可复核记录，为更大范围的研究提供基础。

地球不会因为本书结束而停止留下记录，研究也不会因为形成一套解释语言而停止遇到意外。成熟知识使我们能够稳当地行动，新的差异则可能要求重新打开被封装的条件。地球科学思想史的价值，在于让我们既理解稳定性怎样来之不易，也保留在必要时重新定义问题的能力。下一轮研究，从这个能够被具体指出的必要之处开始。

附录一 煤山案例中的对象与证据

对象	支持材料	不能直接推出
锆石的计时结果	同位素测量 体系保存与颗粒筛选	整层沉积物与颗粒必然同龄
火山灰层的沉积约束	矿物形成解释与剖面层位联系	所有灰层没有再搬运或间断
主要灭绝区间	化石记录 年代模型与采样分辨率	它与正式年代界线完全重合
印杜阶底界	煤山 D 剖面 27c 层底及标志相关	数值年龄修订必然移动定义点
2001 年标准批准	国际地层委员会登记	自然事件发生于批准时刻

本表区分对象层级，不构成原始定年数据复算。主要来源 K01—K03，参见第 13、15 章。

附录二 板块论证的证据时点

时点	主要主张或检验	历史解释边界
1963	Vine 与 Matthews 联系磁异常扩张与倒转	不能使用 1967 年的检验作为当时已知理由
1965	Wilson 提出转换断层及其运动关系	概念提出与后续支持须分开
1967	Sykes 以震源机制检验运动方向和空间分布	摘要结论不等于本书完成逐事件复算
综合阶段	几何 时间与运动证据相互约束	构造模型成功不自动证明某种史学框架优胜

来源 K04—K06；详细推理见第 27、29、30 章。模型共享哪些校准和定义，应与材料独立性分别审查。

附录三 发生学解释的比较标准

问题	需要提供	不足以单独支持
对象身份是否改变	新增的可辨认差别与失败条件	名称更新或数值精化
操作是否被回写	采样 测量或检验的实际改变	逻辑上能够提出一种新计划
特征纠缠是否重构	特征彼此成立条件的具体变化	机构 仪器和环境的并列清单
历史因果是否成立	时间顺序 操作记录与替代解释	结果出现后的顺畅故事
SDE 是否具有解释增量	同材料同截点的公平比较	篇幅更多 术语更多或模型本身成功

本表用于复核解释，不是已经执行的比较实验结果。具体讨论见第 56—60 章。

附录四 洛伦兹系统的教学计算

本附录对应第 39 章。取 σ＝10、r＝28、b＝8/3，初值分别为(1,1,1)与(1＋10^{-5},1,1)，以经典四阶 Runge—Kutta 方法积分到无量纲时间 20。下表是同一步长下两条轨道的欧氏距离；它检验指定有限区间中的扰动放大，不是对真实大气时距的估计，也不是混沌性质的严格证明。

步长	时间 5	时间 10	时间 15	时间 20
0.01	1.24531e-05	1.7577e-05	0.000236719	0.0107159
0.005	1.24527e-05	1.75759e-05	0.000237802	0.00924564
0.0025	1.24527e-05	1.75759e-05	0.000237801	0.00924368
0.00125	1.24527e-05	1.75759e-05	0.000237798	0.00923958

不同步长得到接近的扰动距离，仅是一项数值一致性检查。相同初值在不同步长下的轨道仍有差异，不能据此声称获得严格误差界。以下代码只依赖 Python 标准库，可重新生成表中距离。

最小复现程序

```
from math import sqrt



def f(v):
    x, y, z = v
    return (10 * (y - x), x * (28 - z) - y, x * y - 8 * z / 3)



def step(v, h):
    a = f(v)
    b = f(tuple(v[i] + h * a[i] / 2 for i in range(3)))
    c = f(tuple(v[i] + h * b[i] / 2 for i in range(3)))
    d = f(tuple(v[i] + h * c[i] for i in range(3)))
    return tuple(v[i] + h * (a[i] + 2*b[i] + 2*c[i] + d[i]) / 6
                 for i in range(3))



def run(h, initial):
    v, samples = initial, {}
    checkpoints = {round(t / h) for t in (5, 10, 15, 20)}
    for k in range(1, round(20 / h) + 1):
        v = step(v, h)
        if k in checkpoints:
            samples[round(k * h)] = v
    return samples



def distance(a, b):
    return sqrt(sum((x - y)**2 for x, y in zip(a, b)))



for h in (0.01, 0.005, 0.0025, 0.00125):
    a = run(h, (1, 1, 1))
    b = run(h, (1 + 1e-5, 1, 1))
    for t in (5, 10, 15, 20):
        print(h, t, format(distance(a[t], b[t]), '.9g'))
```

参考资料与来源索引

A、B、C 为学科案例来源，K 为两项核心论证来源，N 为比较方法论，M 为 SDE 理论底本；R 及其他新增代码对应补充材料。同一文献可能有不同入口，其条目数量不表示独立证据数量。

A01　Geological Society of London：William Smith 1815 地质图馆藏说明

https://www.geolsoc.org.uk/about-us/history/visiting-the-william-smith-map/a-delineation-of-the-strata-of-england-and-wales-with-part-of-scotland/

A02　Geological Society of London：Strata identified by organized fossils

https://www.geolsoc.org.uk/the-library/online-exhibitions/william-strata-smith/aftermath-of-the-map/strata-identified-by-organized-fossils/

A03　Nicolaus Steno：De solido intra solidum naturaliter contento dissertationis prodromus（1669）

https://www.biodiversitylibrary.org/item/250134

A04　James Hutton：Theory of the Earth, With Proofs and Illustrations, Vol. I（1795）

https://www.gutenberg.org/files/12861/old/12861-h/12861-h.htm

A05　IUGS Geoheritage：Siccar Point, Hutton’s Unconformity

https://iugs-geoheritage.org/geoheritage_sites/siccar-point/

A06　Charles Lyell：Principles of Geology, Vol. I（1830）

https://darwin-online.org.uk/converted/Ancillary/BeagleLibrary/1830-3_Lyell_A505/1830_Lyell_A505.1.html

A07　Georges Cuvier：Essay on the Theory of the Earth，英译第五版，附 Robert Jameson 说明（1827）

https://www.gutenberg.org/files/62918/62918-h/62918-h.htm

A08　William Thomson：On the Secular Cooling of the Earth，Proceedings of the Royal Society of Edinburgh 4（1862）, 610–611

https://doi.org/10.1017/S0370164600035124

A09　Arthur Holmes：The Age of the Earth（1913）

https://www.gutenberg.org/files/78241/78241-h/78241-h.htm

A10　C. Patterson, G. Tilton, M. Inghram：Age of the Earth，Science 121（1955）, 69–75

https://authors.library.caltech.edu/records/exhx7-e7g91

A11　C. Patterson：Age of meteorites and the earth，Geochimica et Cosmochimica Acta（1956）

https://doi.org/10.1016/0016-7037(56)90036-9

A12　International Commission on Stratigraphy：GSSP Table，Induan / Meishan 条目

https://stratigraphy.org/gssps/

A13　G. K. Gilbert：Report on the Geology of the Henry Mountains（1877）

https://www.gutenberg.org/files/75119/75119-h/75119-h.htm

A14　W. M. Davis：The Geographical Cycle，The Geographical Journal 14(5)（1899）

https://doi.org/10.2307/1774538

A15　John P. Snyder：Map Projections—A Working Manual，USGS Professional Paper 1395（1987）

https://www.usgs.gov/publications/map-projections-a-working-manual

A16　USGS CEGIS：Generalization，多尺度地图表达研究说明

https://www.usgs.gov/centers/cegis/science/generalization

A17　中国科学院地质与地球物理研究所转载：丁文江——地质教育基地的开创者（2016）

https://igg.cas.cn/kxcb/kxpj/kpdt/201601/t20160106_4514695.html

A18　丁文江：Report on the Geology of the Yangtze Valley Below Wuhu（1919）

https://books.google.com/books?id=x34RAAAAIAAJ

B01　USGS, Historical perspective — This Dynamic Earth

https://pubs.usgs.gov/gip/dynamic/historical.html

B02　USGS, Developing the theory — This Dynamic Earth

https://pubs.usgs.gov/gip/dynamic/developing.html

B03　Vine & Matthews (1963), Magnetic Anomalies Over Oceanic Ridges

https://www.nature.com/articles/199947a0

B04　Wilson (1965), A New Class of Faults and their Bearing on Continental Drift

https://www.nature.com/articles/207343a0

B05　Library of Congress, Mapping the Deep: Marie Tharp’s Physiographic Diagram of the Atlantic Ocean

https://blogs.loc.gov/maps/2025/07/mapping-the-deep-marie-tharps-physiographic-diagram-of-the-atlantic-ocean/

B06　Mohorovičić, Earthquake of 8 October 1909（1992 年英文译刊）

https://ojs.srce.hr/index.php/geofizika/article/view/15978

B07　Columbia University, Seismic Evidence for Internal Earth Structure

https://www.columbia.edu/~vjd1/earth_int.htm

B08　Niels Bohr Institute, P′ or the discovery of the Earth’s inner core

https://nbi.ku.dk/english/www/inge/lehmann/jordens-kerne/

B09　Birch (1952), Elasticity and Constitution of the Earth’s Interior

https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JZ057i002p00227

B10　Dziewonski & Anderson (1981), Preliminary Reference Earth Model

https://authors.library.caltech.edu/records/rjkzg-1a346

B11　Saltus & Blakely (2011), Unique geologic insights from non-unique gravity and magnetic interpretation

https://www.usgs.gov/publications/unique-geologic-insights-non-unique-gravity-and-magnetic-interpretation

B12　Holmes (1931), Radioactivity and Earth Movements

https://www.nature.com/articles/128496e0

B13　Met Office History

https://www.metoffice.gov.uk/research/library-and-archive/archive-hidden-treasures/met-office-history

B14　University of Bergen, Bergen School of Meteorology

https://www4.uib.no/en/faculty-of-science-and-technology/geophysical-institute/om-geofysisk-institutt/bergen-school-of-meteorology

B15　NASA, Weather Forecasting Through the Ages

https://science.nasa.gov/earth/earth-observatory/weather-forecasting-through-the-ages/

B16　Charney, Fjørtoft & von Neumann (1950), Numerical Integration of the Barotropic Vorticity Equation

https://tellusjournal.org/articles/10.3402/tellusa.v2i4.8607

B17　Lorenz (1963), Deterministic Nonperiodic Flow

https://journals.ametsoc.org/view/journals/atsc/20/2/1520-0469_1963_020_0130_dnf_2_0_co_2.xml

B18　ECMWF, Linearised physics: the heart of ECMWF’s 4D-Var

https://www.ecmwf.int/en/newsletter/175/earth-system-science/linearised-physics-heart-ecmwfs-4d-var

B19　Bullard, Everett & Smith (1965), The Fit of the Continents around the Atlantic

https://doi.org/10.1098/rsta.1965.0020

B20　Alfred Wegener Institute, Alfred Wegener

https://www.awi.de/en/about-us/organisation/alfred-wegener.html

C01　USGS，Watershed Boundary Dataset（含 Federal Standards and Procedures 第 5 版入口）

https://www.usgs.gov/national-hydrography/watershed-boundary-dataset

C02　Horton, R. E. (1933). The Rôle of infiltration in the hydrologic cycle. Eos, 14(1), 446–460. DOI:10.1029/TR014i001p00446

https://agupubs.onlinelibrary.wiley.com/doi/10.1029/TR014i001p00446

C03　Sverdrup, H. U. (1947). Wind-Driven Currents in a Baroclinic Ocean; with Application to the Equatorial Currents of the Eastern Pacific. PNAS, 33(11), 318–326. DOI:10.1073/pnas.33.11.318

https://www.pnas.org/doi/10.1073/pnas.33.11.318

C04　NOAA NCEI，What Are Proxy Data?

https://www.ncei.noaa.gov/news/what-are-proxy-data

C05　Hays, J. D., Imbrie, J., & Shackleton, N. J. (1976). Variations in the Earth’s Orbit: Pacemaker of the Ice Ages. Science, 194(4270), 1121–1132. DOI:10.1126/science.194.4270.1121

https://www.science.org/doi/10.1126/science.194.4270.1121

C06　Manabe, S., & Wetherald, R. T. (1967). Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity. Journal of the Atmospheric Sciences, 24, 241–259.

https://www.gfdl.noaa.gov/bibliography/related_files/sm6701.pdf

C07　Farquhar, J., Bao, H., & Thiemens, M. (2000). Atmospheric Influence of Earth’s Earliest Sulfur Cycle. Science, 289(5480), 756–758. DOI:10.1126/science.289.5480.756

https://pubmed.ncbi.nlm.nih.gov/10926533/

C08　Earth System Science: A Closer View (1988). National Academies 出版入口，DOI:10.17226/19088

https://www.nationalacademies.org/publications/19088

C09　USGS，Landsat Collection 2

https://www.usgs.gov/landsat-missions/landsat-collection-2

C10　ECMWF Destination Earth，Climate Change Adaptation Digital Twin（项目DE_340_CSC）

https://destine.ecmwf.int/provider/climate-adaptation-digital-twin/

C11　Mousavi, S. M., Ellsworth, W. L., Zhu, W., Chuang, L. Y., & Beroza, G. C. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature Communications, 11, 3952. DOI:10.1038/s41467-020-17591-w

https://www.nature.com/articles/s41467-020-17591-w

C12　Guffanti, M. C., & Miller, T. (2013). A volcanic activity alert-level system for aviation: Review of its development and application in Alaska. Natural Hazards. DOI:10.1007/s11069-013-0761-4

https://www.usgs.gov/publications/a-volcanic-activity-alert-level-system-aviation-review-its-development-and-application

C13　NASA (1989). Earth system science: A program for global change. NASA-TM-101186，NTRS 19890013598

https://ntrs.nasa.gov/citations/19890013598

C14　USGS，Landsat Update Special Issue, December 2020

https://www.usgs.gov/landsat-missions/landsat-update-special-issue-december-2020

K01　International Commission on Stratigraphy：Global Boundary Stratotype Section and Points，Induan 登记项

https://stratigraphy.org/gssps/

K02　Burgess, S. D., Bowring, S., & Shen, S.-Z. (2014). High-precision timeline for Earth’s most severe extinction. PNAS, 111, 3316–3321. DOI:10.1073/pnas.1317692111

https://dspace.mit.edu/entities/publication/422cbad4-f022-4a65-a228-31e48a6d7d52

K03　Burgess et al. (2014). Correction for: High-precision timeline for Earth’s most severe extinction. DOI:10.1073/pnas.1403228111

https://www.pnas.org/doi/10.1073/pnas.1403228111

K04　Vine, F. J., & Matthews, D. H. (1963). Magnetic Anomalies Over Oceanic Ridges. Nature, 199, 947–949. DOI:10.1038/199947a0

https://www.nature.com/articles/199947a0

K05　Wilson, J. T. (1965). A New Class of Faults and their Bearing on Continental Drift. Nature, 207, 343–347. DOI:10.1038/207343a0

https://www.nature.com/articles/207343a0

K06　Sykes, L. R. (1967). Mechanism of earthquakes and nature of faulting on the mid-oceanic ridges. Journal of Geophysical Research, 72, 2131–2153. DOI:10.1029/JZ072i008p02131

https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/jz072i008p02131

N01　Thomas S. Kuhn, The Structure of Scientific Revolutions（1962；2012 年 50 周年版官方书目）

https://press.uchicago.edu/ucp/books/book/chicago/S/bo13179781.html

N02　Carol E. Cleland, Historical science, experimental science, and the scientific method, Geology 29(11), 987–990（2001）

https://pubs.geoscienceworld.org/gsa/geology/article/29/11/987/197903/Historical-science-experimental-science-and-the

N03　Carol Cleland, Differences in the methodology and justification of the field sciences and the classical experimental sciences（2012，作者工作坊论文）

https://serc.carleton.edu/integrate/workshops/methods2012/essays/cleland.html

N04　Ian Hacking, Representing and Intervening，第 13 章 The creation of phenomena，220–232（1983）

https://www.cambridge.org/core/books/abs/representing-and-intervening/creation-of-phenomena/BC59D09D172ED684FC5C30B1A358FA3F

N05　Earth System Science: A Closer View（1988；NASA 地球系统科学委员会报告，NAP 数字版）

https://www.nationalacademies.org/publications/19088

N06　NASA, Global change data sets: Excerpts from the Master Directory, version 2.0（1992）

https://ntrs.nasa.gov/citations/19920024784

M01　王德生 SDE 三方程导论 第二章 三方程的标准形式与理论地位

https://sdeuniverses.com/books/m/51/chapter-02/

M02　王德生 从三视角的 123 原理到 SDE 的三原理

https://sdeuniverses.com/column/ontology-grid/123-to-three-principles/

M03　王德生 六路径方法论入门 第三十章 边界反例与可能的错误

https://sdeuniverses.com/books/m/52/chapter-30/

M04　王德生 六路径方法论入门 第三章 SDE 本体论入门

https://sdeuniverses.com/books/m/52/chapter-03/

M05　王德生 属性纠缠论 实体属性发生学导论

https://sdeuniverses.com/books/m/396/text/

补充来源与材料范围

第 1—10 章来源与审读记录

本章节单元正文保留大纲 A 编号；以下说明本次审读实际核读范围，不把大纲索引视为已读全文。核查日期 2026 年 10 月 16 日。未作野外调查、实验、地图配准复算或拉丁文整书校勘。正文中的假设剖面、采样概率及地图读图练习均明确为教学分析。

◆ 已用大纲来源

A01 Geological Society of London，William Smith 1815 地图馆藏说明。https://www.geolsoc.org.uk/about-us/history/visiting-the-william-smith-map/a-delineation-of-the-strata-of-england-and-wales-with-part-of-scotland/ 以及同机构详细展览页：https://www.geolsoc.org.uk/the-library/online-exhibitions/william-strata-smith/the-map/a-delineation-of-the-strata-of-england-and-wales-with-part-of-scotland/
实际核读两页正文、底图/颜色/副本/怀特岛版本图注与链接；图像入口取得但未作像素级或原图边界复核。正文未声称复原某条具体边界的地质修订原因。馆藏事实采用简短概括，大量读图论述为本书独立分析。发行阶段以馆藏表述区分，不把图上日期当全部副本的实际发行日。

A02 Geological Society of London，Strata identified by organized fossils。https://www.geolsoc.org.uk/the-library/online-exhibitions/william-strata-smith/aftermath-of-the-map/strata-identified-by-organized-fossils/
已读全页文字与图注，确认 1816—1819 图版出版物、1817 文字系统及相关表格版本；未逐栏辨读两个 1817 原表，未断言具体化石增删。

A03 Steno，1669，De solido intra solidum naturaliter contento dissertationis prodromus。
https://www.biodiversitylibrary.org/item/250134
大纲原入口仅题录；本次审读尝试 Smithsonian 入口遭拒，改以 R07a 取得相关拉丁文电子转写，不宣称完成原扫描校勘。第 7 章已据新增阅读修订旧的“仅入口”范围描述。

A04 Hutton，1795，Theory of the Earth, With Proofs and Illustrations, Vol. I。
本次审读可用公版全文：https://www.gutenberg.org/files/12861/12861-h/12861-h.htm
已核读西卡角之前对包含碎片与矿脉终止推理的比较，以及“SECT. II.—The Theory confirmed from Observations made on purpose to elucidate the subject”中 1788 海岸与溪流考察段落。内容包括寻找接触、Hall 地方信息、Pease/Tour 溪谷、Siccar Point 砂岩覆盖与下部碎片、未见接触、内陆砾质堆积、天气限制、St Abb's Head 远景预期与抵近观察差异。没有通读全书；电子站点“of 4”编排不作为 1795 作者出版卷数。

A05 IUGS Geoheritage，Siccar Point。
https://iugs-geoheritage.org/geoheritage_sites/siccar-point/
已读地质描述与历史说明；正文仅简短使用现代志留纪/泥盆纪与角度关系，明确现代名称不倒置为哈顿当时知识。

A09 Holmes，1913，The Age of the Earth。
https://www.gutenberg.org/files/78241/78241-h/78241-h.htm
沿用前稿既有实际核读的第 I、IX、X 章相关范围；本章节单元第 5 章只概括其冷却与放射性计时比较、样品条件，不新增未经核查数据或引语。

A10 Patterson, Tilton, Inghram，1955，Age of the Earth，Science 121, 69–75。
https://authors.library.caltech.edu/records/exhx7-e7g91
本次审读核读机构题录与摘要；明确 Metadata-only，未据此声称读取全部论文。仅用作者、题名、年份及地表铅与铁陨石铅比较的方法入口。

A11 Patterson，1956，Age of meteorites and the earth。https://www.sciencedirect.com/science/article/pii/0016703756900369
本次审读检索核出版题录。另找到高校 PDF 入口但未将其作为全文核读成果，本章节单元仅使用论文身份与 1955 区别，不使用具体年龄、表格或计算。

A12 International Commission on Stratigraphy，GSSP Table。https://stratigraphy.org/gssps/
核读 Induan/Meishan 条目：Meishan D Section、base of Bed 27c、Conodont FAD Hindeodus parvus、Ratified 2001，及数值年龄与定义位置分列。第 2—3 章不重复现行具体年龄，也不假称完整批准史已核读。

A13 Gilbert，1877，Report on the Geology of the Henry Mountains。https://www.gutenberg.org/files/75119/75119-h/75119-h.htm
出版机构题录：https://pubs.usgs.gov/publication/70039916
实际核读公版序言、目录、Chapter I 岩层描述与不整合段，Chapter II 结构类型与写作次序说明，Chapter III 艾伦山相关岩盖接触描述；未全读全部图版和力学推导。正文特别使用：作者自述路线/方法自主，野外材料 1875—1876；原文明确写作次序不同于调查次序；岩盖下部通道未直接见到；某些岩层凭岩性难区分；暂名与时代归属不确定；平行层理不保证沉积连续。题名页 1877 不扩写为精确发行日。

◆ 新增专属编号

[R07a] Steno 1669，拉丁原文电子转写入口，Source Library（注明扫描来源 Smithsonian/Internet Archive）。https://sourcelibrary.org/book/de-solido-intra-solidum-naturaliter-contento-dissertationis-steno/read
实际核读范围：原文 Propositio I 附近、关于地层材料/位置/形态段落（网站扫描序号 42—46 附近）、前置图解说明。参考的是页面显示的拉丁文 OCR，与并列译文对照理解；该站译文为 AI 处理，不作为可靠学术译本单独引用。未引用原话、未给未经扫描核准的书内页码。正文概括固体接触、生成不应只用总括名称解释、层位上下与侧向延续、原文

侧限或全球覆盖的较强二分、地下空洞与地层改变模型。正式学术出版前宜用可靠译本及原扫描校勘这些表述。本次审读不将电子转写的可读性等同校勘可靠性。

[R08a] Geological Society of London，Table of strata in the vicinity of Bath, 1799。https://www.geolsoc.org.uk/the-library/online-exhibitions/william-strata-smith/stratigraphical-theories/table-of-strata-in-the-vicinity-of-bath-1799/ 已读馆藏说明，确认史密斯口述、理查森手书、原件褪色及 1897 摹本；正文短述文件身份，未逐栏辨认原表、未复原全部对话或人物心理。

[R09a] Peter Wigley，2016，The development and evolution of the William Smith 1815 geological map from a digital perspective，GSA Today 26(7), 4–10, doi:10.1130/GSATG279A.1。https://www.geosociety.org/gsa-today/july-2016/the-development-and-evolution-of-the-william-smith-1815-geological-map-from-a 已读公开 HTML 正文，正文仅简述作者数字比较方法及投影/配准问题，未采纳其所有优先权判断，未称本书完成独立配准或数据复算。史密斯副本发行存在细节差异，正文采用馆藏的宽阶段表述。

第 11—20 章来源与核读说明

写作日期：2026 年 10 月 16 日。大纲 A、K、M 代码沿用 全书来源索引。以下记录本章节单元实际核查，不把题录当全文。

A06：Charles Lyell, Principles of Geology, Vol. I, 1830。https://darwin-online.org.uk/converted/Ancillary/BeagleLibrary/1830-3_Lyell_A505/1830_Lyell_A505.1.html。本章节单元打开初版公开全文，核目录、卷首及相关检索内容；讨论河流、洪水、火山地震与气候安排。未通读三卷；避免使用 Gutenberg1853 版替代 1830 版。

A07：Georges Cuvier, Essay on the Theory of the Earth，1827 英译第五版。https://www.gutenberg.org/files/62918/62918-h/62918-h.htm 。核标题、目录、灭绝类型与现生变种论证入口及相关段落。明确区分居维叶正文与 Jameson 说明，未逐页通读。

A08：William Thomson, On the Secular Cooling of the Earth，1862 Proceedings 记录，DOI 10.1017/S0370164600035124 。原始会议记录检索入口 https://zenodo.org/record/1798526/files/article.pdf 。本章节单元核检索摘要支持温度梯度与向外传热；未复原 Transactions 长文全部边值。第 11 章热扩散公式为本书教学推演，不冒称历史计算复现。

A09：Arthur Holmes, The Age of the Earth, 1913。https://www.gutenberg.org/files/78241/78241-h/78241-h.htm 。打开公开全文，沿既有稿保留第十章矿物年龄论证；冷却与 Perry 讨论按霍姆斯后来的回顾归属，不作 Perry 原始通讯直引。

A10：Patterson, Tilton and Inghram, Age of the Earth, Science121,69–75,1955。https://authors.library.caltech.edu/records/exhx7-e7g91 。沿大纲已核题录摘要定位，未声称新增全文核读。

A11／新增 R14a：Clair Patterson, Age of meteorites and the earth, Geochimica et Cosmochimica Acta10,230–237,1956，DOI10.1016/0016-7037(56)90036-9 。原文公开扫描 https://www.its.caltech.edu/~bi1/Bi1__Micro-_to_Macro-Biology/Additional_Readings_files/Age%20of%20meterorites%20and%20the%20earth.pdf 。本章节单元取得 8 页 PDF，逐页查看全部 8 页图像，核共同初始条件、铅关系、陨石向地球迁移的限制及结论。没有独立复算表 1 和全部误差；正文只简述少量历史结果，详细数学及反事实为本书教学分析。另一个 Harvard 入口访问失败，不作为已读来源。

A12／K01：ICS GSSP 登记 https://stratigraphy.org/gssps/ 。沿前稿保留已核煤山 27c 层底、Hindeodus parvus 和 2001 批准；不声称重新读取批准论文或表决档案。

K02：Burgess, Bowring and Shen, High-precision timeline for Earth's most severe extinction, PNAS2014 。 https://web.math.princeton.edu/~sswang/literature_general_unsorted/PNAS-2014-Burgess-suddenness-of-Permian-extinction.pdf 。继承前稿核读范围：公开 PDF 相关正文，未单颗粒复算。

K03：2014 更正，DOI10.1073/pnas.1403228111，https://pmc.ncbi.nlm.nih.gov/articles/PMC3977310/ 。继承前稿检索更正文本说明，未逐项核更正图像，正文不复制未核数值图。

A13：G. K. Gilbert, Report on the Geology of the Henry Mountains,1877。https://www.gutenberg.org/files/75119/75119-h/75119-h.htm 。本章节单元打开全文并核卷首图注、目录、层序辨认、侵入原因、坡度负载与动态平衡，以及分水岭凸形残差段落。第 16、19 章实际依这些公版原文展开；未宣称全部图版或野外结论获现代验证。

A14：W. M. Davis, The Geographical Cycle,1899, DOI10.2307/1774538 。公开原文入口 https://geomorphology.sese.asu.edu/Papers/Davis_1899.pdf 多次打开超时；JSTOR 仅题录可见。检索原文首段确认 structure/process/time 框架。第 19 章阶段序列明确为教学重建，不声称原图复现或全文解读。此为本章节单元主要原典核读缺口。

A15：John P. Snyder, Map Projections—A Working Manual, USGS Professional Paper1395,1987。https://www.usgs.gov/publications/map-projections-a-working-manual 。打开官方说明；未逐项复算手册。第 17 章三个球面投影、局部比例与雅可比关系为本书解析教学，无实际地图生成或区域测量宣称。

A16：USGS CEGIS, Generalization。https://www.usgs.gov/centers/cegis/science/generalization 。打开官方研究说明，支持现代多尺度综合操作。第 18 章网格与河网例子为构造演示，未执行真实区域数据实验。

A17：中国科学院地质与地球物理研究所转载《丁文江：地质教育基地的开创者》，2016。https://igg.cas.cn/kxcb/kxpj/kpdt/201601/t20160106_4514695.html 。本章节单元打开全文，仅用 1916 调查、1919 出版概况；未将人物宣传评价当作学术结论，未将其“60 年一英里”概述泛化使用。

A18：丁文江，Report on the Geology of the Yangtze Valley Below Wuhu，Shanghai Mercury，1919。https://books.google.com/books?id=x34RAAAAIAAJ 。本章节单元核题录 84 页以及页面 73、75、77、58 的公开原文片段。正文约 69 年一英里仅归属特定塘段，412/6 为本次审读有限算术核对，不宣称测距或地图复现。

第 21—30 章来源与核读记录

本章节单元正文约 56000 汉字，原有第 27、29 章实质正文完整继承，新增段落为作者分析与明确标示的教学推演。未实施历史数值独立复算。所有来源代码沿用总账，新增代码如下。来源状态不等于全文已读。

◆ 新增来源

[R25a] Geological Society of Glasgow. Arthur Holmes. URL: https://geologyglasgow.org.uk/arthur-holmes/
实际范围：本次审读读取官方正文及所转录 1928 年会议记录摘要。可确认 1928 年 1 月 12 日讲演、1931 年会刊长文以及摘要中不均匀放射性加热、深部缓慢流动的论证方向。页面链接的完整长文未通读。本文不采纳其概述中“提前 35 年”等夸大性优先权表达。长文候选备份 https://www.mantleplumes.org/WebDocuments/Holmes1931.pdf 本次审读访问失败；不得记为已读。B12 短通信依旧不能代替长文。

[R28a] Hess, H. H. (1962). History of Ocean Basins. Petrologic Studies, pp. 599–620.
DOI/出版入口：https://doi.org/10.1130/Petrologic.1962.599
实际范围：本次审读出版者检索题录与摘要片段；另在法国教育机构保存的原始论文节选合集 https://pedagogie.ac-orleans-tours.fr/documents/pdf/jsvt-2011-conf4.pdf 中读到图 7 相关原文片段。只确认 1962 年节点与海盆更新问题，不声称通读全文，不转用其早期成分假设为现代定论。备份 https://svalgaard.leif.org/EOS/Hess1962.pdf 打开失败。

[R28b] Scientific Party. Initial Reports of the Deep Sea Drilling Project, Volume III (1970), Leg 3.
目录：https://deepseadrilling.org/03/dsdp_toc.htm
DOI: 10.2973/dsdp.proc.3.1970
目录实际读取：站位 13—22、船上报告、岸上研究、年代框架与样品入口。只将其作为报告体系与实物研究的史料入口，不引用未核读的逐孔年龄。航次时段与题录核查：https://sedis.iodp.org/pub-catalogue/?action=edit&id=10.2973%2Fdsdp.proc.3.101.1970 （官方 IODP 索引检索文字）。1968 年 12 月至 1969 年 1 月为航次；1970 为报告文献年；目录页面 June 2007 不是航次年。
原报告 https://www.deepseadrilling.org/03/volume/dsdp03_01.pdf 本次审读获取超时，未声称全文已读。

[R30a] McKenzie, D., & Parker, R. (1967). The North Pacific: an Example of Tectonics on a Sphere. Nature, 216, 1276–1280. URL: https://www.nature.com/articles/2161276a0 DOI: 10.1038/2161276a0
实际范围：出版者题录、作者档案检索片段及教育机构所存原文片段入口。未逐图复核。正文仅用来确认球面运动学综合节点，不主张具体优先权或复现全部区域计算。作者档案入口：https://www.mckenziearchive.org/chapters/plates-as-paving-stones/（检索片段可读，直接打开超时）；https://www.mckenziearchive.org/chapters/publishing-plate-tectonics/（检索片段可定位 1967 年 11 月 10 日投稿信，但正文未据此扩展优先权）。

[R30b] Morgan, W. J. (1968). Rises, Trenches, Great Faults, and Crustal Blocks. Journal of Geophysical Research, 73, 1959–1982. URL: https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/JB073i006p01959 DOI: 10.1029/JB073i006p01959
实际范围：出版者摘要检索；并读教育机构原文节选 https://pedagogie.ac-orleans-tours.fr/documents/pdf/jsvt-2011-conf4.pdf 中球面旋转、边界方向、角速度向量及回路闭合段落。未逐页通读，不报告未复算的具体旋转极数值。第 30 章叉乘、瞬时闭合与有限旋转次序等为作者教学说明。独立 PDF https://www.geologie.ens.fr/50years_plate_tectonics/1968-morgan.pdf 获取失败，不记为全文已读。

◆ 既有来源实际使用与边界

[B01] USGS Historical perspective, https://pubs.usgs.gov/gip/dynamic/historical.html：既有大纲官方摘要范围，正文只用机制不足与争论背景；不引用具体批评者原话。
[B02] USGS Developing the theory, https://pubs.usgs.gov/gip/dynamic/developing.html：本次审读检索并打开机构概述，使用海底地形、磁场反转、扩张、地震/火山汇合的有限概述。各章大部分技术分析

为独立教学分析，不以该页替代原典。

[B05] Library of Congress, Mapping the Deep: Marie Tharp’s Physiographic Diagram of the Atlantic Ocean, https://blogs.loc.gov/maps/2025/07/mapping-the-deep-marie-tharps-physiographic-diagram-of-the-atlantic-ocean/：本次审读全文读取。事实使用限于 1957 图、地貌表达、资料空白、保密与航迹/地震/测深资料构成，不虚构绘图对话或逐航次改道。

[B19] Bullard, Everett & Smith (1965), The Fit of the Continents around the Atlantic：本次审读检索出版者题录/摘要范围；数值最小二乘可确认，未核原图残差。正文不填写 500 fathom 具体轮廓或残差值，不冒称复算。文中拟合敏感性等为作者分析。

[B20] Alfred Wegener Institute, Alfred Wegener：沿用官方传记核定的 1912/1915/1929 节点，本次审读检索复核。未逐页比较三个版本，正文不填写未经确认的版次增删细节。

[B03]/[K04] Vine & Matthews (1963)：继承首稿已核的三页原文相关论述，备份 https://eos-courses.readthedocs.io/en/latest/_downloads/6162e79c651ae4ebee8db0613b18518d/Vine1963.pdf。低磁纬正常磁化可产生负异常、原图仅拟合基本形态等来自首稿核读。本次审读不另声称数据复算。

[B04]/[K05] Wilson (1965)：原刊题录及后续原始研究的预测回顾，不把现代教学图当成原图，不主张读完全部图件。

[K06] Sykes (1967), https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/jz072i008p02131：继承首稿完整摘要范围；本次审读教育机构原文节选再次看到 17 次、断裂带 10 次及方向相反判别。未独立复算逐事件震源机制，保留 1967 后续检验时点。

◆ 待正式出版前优先补证

1. 第 21 章 Bullard 原始轮廓、残差图与边界选择；第 22 章魏格纳三个版本逐段对读，以及具体力学批评原文。

2. 第 23 章阿巴拉契亚原始区域剖面与旧模型实际预测案例，现稿明确采用教学重构，不

假称已完成档案研究。
3. 第 24/26 章具体航次计划—补测—地图修订连续档案；现稿区分一般机制与已证历史。
4. 第 25 章 Holmes 长文完整推理；第 28 章第三航次逐孔年龄与取样选择复核。
5. 第 29 章 Wilson 逐图、Sykes 逐事件原始机制；第 30 章原文旋转参数复算及实际共同体采用史。

正文已形成完整连续论述，但以上原典与接受史缺口仍影响正式出版级史料密度，不能因字数达标宣布全部学术核验完成。

第 31—40 章新增来源与核读记录

本章节单元引文代码 B06—B18 沿用全书大纲总账。以下 R 代码是本次审读新增或补核的来源，不把检索摘要记作全文通读。数学关系除明确注明外均为本书教学推演，未声称重现历史原始数据。

[R31a] Kölbl-Ebert, M. Inge Lehmann's paper: “P′” (1936). Episodes 24(4), 2001；文中包括 1936 年原文的明确标页节选。URL：https://doi.org/10.18814/epiiugs/2001/v24i4/007
公开 PDF：https://courses.seas.harvard.edu/climate/eli/Courses/EPS281r/Sources/Inner-Core/Lehmann-1936-extracts%2Binterpretation.pdf
实际核读：六页整理文提取正文，重点原文第 87、95、101、105—115 页的节选及编辑说明；确认 P′记录、速度较高内部区域候选、急剧连续增速替代与资料不足限制。不是 1936 年全篇逐页复算，也未重新拾取原波形。正文明确区分原文节选与后人解释，未采用后人重绘图冒充原图。

[B06 补核] Mohorovičić, A. Earthquake of 8 October 1909. Geofizika 英文译刊（1992）。URL：https://ojs.srce.hr/index.php/geofizika/article/view/15978
实际核读：出版方题录与摘要，固定 1909 地震、1910 原作、1992 译刊的时间差；没有完成历史走时表复算。二层折射与反射公式为本书教学模型。

[B07 补核] Columbia University. Evidence for Internal Earth Structure and Composition.
URL：https://www.columbia.edu/~vjd1/earth_int.htm
实际核读：机构教学页正文；仅用于 1914 节点、核幔边界约 2900 公里的教学背景，不将页面中简略或过时物性措辞作为现代完整论证。未取得古登堡 1914 原文，正文不报告其精确数据与计算。

[B09 补核] Birch, F. (1952). Elasticity and Constitution of the Earth's Interior. Journal of Geophysical Research 57, 227–286. DOI:10.1029/JZ057i002p00227.
URL：https://agupubs.onlinelibrary.wiley.com/doi/10.1029/JZ057i002p00227
实际核读：题录与检索摘要；原实验表未逐表复算。正文未引用未核定样品数、压力范围或原表数值。弹性速度式、迁移与敏感性讨论为教学解释。

[B11 补核] Saltus, R. W., & Blakely, R. J. (2011). Unique geologic insights from "non-unique" gravity and magnetic interpretation. GSA Today.
出版入口：https://www.geosociety.org/gsa-today/november-2011/unique-geologic-insights-from-nonunique-gravity-and-magnetic-interpretation
作者机构入口：https://pubs.usgs.gov/publication/70032258
实际核读：出版方和 USGS 检索可见摘要；没有将原文具体测区当作已复算案例。点质量、联合约束与正则化例子均由本书自拟。

[R34a] Geological Society of Glasgow. Arthur Holmes（学会历史档案说明与 1928 讲演摘要转录）。
URL：https://geologyglasgow.org.uk/arthur-holmes/
长文题录：Holmes, A. (1931). Radioactivity and Earth Movements. Transactions of the Geological Society of Glasgow 18(3), 559–606. DOI:10.1144/transglas.18.3.559.
另核当年 Nature 介绍：https://www.nature.com/articles/128419b0
实际核读：学会页面及检索返回的讲演摘要、出版年信息；Nature 介绍的题录与可见摘要。长篇 559—606 页未全文核读。与 Holmes 的 1931 年 Nature 短通信 https://www.nature.com/articles/128496e0 区分，正文不把短通信当完整对流理论。

[R34b] Slab pull and the seismotectonics of subducting lithosphere（1987 研究，USGS 作者机构摘要）。
URL：https://www.usgs.gov/publications/slab-pull-and-seismotectonics-subducting-lithosphere
实际核读：研究摘要检索正文；支持板片拉力与地震机制、构造模型联系这一有限主张。不将摘要中的力贡献比较推广到全部板块，不报告未经复核样本与原图数值。

[R35a] Dziewonski, A. M., & Anderson, D. L. (1981). Preliminary Reference Earth Model. Physics of the Earth and Planetary Interiors 25(4), 297–356. DOI:10.1016/0031-9201(81)90046-7.
URL：https://authors.library.caltech.edu/records/rjkzg-1a346
实际核读：作者机构完整摘要与元数据，明确页面是 metadata-only；摘要确认多类观测、质量/转动惯量、非弹性频散和横向各向同性。公开 PDF https://lweb.cfa.harvard.edu/~lzeng/papers/PREM.pdf 本次审读获取超时，未假称逐表核验。正文只作有限事实概述，其余为本书参考模型认识论分析。

[B13 补核] Met Office History.
URL：https://www.metoffice.gov.uk/research/library-and-archive/archive-hidden-treasures/met-office-history
实际核读：官方历史相关正文，确认 1853 布鲁塞尔海事合作、1854 机构建立与 FitzRoy；未读取原始会议协议和全部船舶表格。船舶参照、时间对齐、校准等为教学性机制分析，不虚构建站决策档案。

[R37a] Bjerknes, J. (1919). On the Structure of Moving Cyclones. Geofysiske Publikationer 1(2).
原刊扫描：https://upload.wikimedia.org/wikipedia/commons/a/a4/On_the_structure_of_moving_cyclones_(IA_onstructureofmov00bjerrich).pdf
同期转载题录：https://journals.ametsoc.org/view/journals/mwre/47/2/1520-0493_1919_47_95_otsomc_2_0_co_2.xml
实际核读：扫描件提取原文第 1—8 页（PDF 第 8—15 页）文字，重点开头观测间距与时

间间隔问题、当时辐合线术语及暖区/云雨的联系。正文只少量概述历史事实，现代锋面与动力说明明确作为教学展开；未将后来所有术语和机制倒填 1919。

[R37b] Bjerknes, J., & Solberg, H. (1922). Life Cycle of Cyclones and the Polar Front Theory of Atmospheric Circulation. URL：https://www.inscc.utah.edu/~steenburgh/classes/5210/papers/bjerknes_and_solberg_1922.pdf
实际核读：原扫描 PDF 定位、题录与检索片段；未逐页核图，故正文不重建具体图序。多人贡献的组织史另据卑尔根大学官方页：https://www4.uib.no/en/faculty-of-science-and-technology/geophysical-institute/om-geofysisk-institutt/bergen-school-of-meteorology 及其学科史页 https://www4.uib.no/en/faculty-of-science-and-technology/geophysical-institute/om-geofysisk-institutt/a-brief-history-of-the-field 。

[R38a] Bjerknes, V. (1904). Das Problem der Wettervorhersage, betrachtet vom Standpunkte der Mechanik und der Physik. Meteorologische Zeitschrift 21, 1–7；Volken 与 Brönnimann 译编（2009），18, 663–667。URL：https://math.nyu.edu/~gerber/courses/2018-fruhling/bjerknes-physical_basis_NWP-MeteoZ-1904-english.pdf
实际核读：英译原典的初值与规律条件、海洋和高空观测缺口、状态地图段；不把译文编辑注或现代模型知识归为 1904 年原文。

[R38b] Lynch, P. Richardson's Marvellous Forecast（作者对原算例数值重建的研究摘要）。URL：https://mathsci.ucd.ie/~plynch/Publications/RicMarFor_abs.html
相关研究题录：The origins of computer weather prediction and climate modeling.
URL：https://www.sciencedirect.com/science/article/pii/S0021999107000952
实际核读：作者研究摘要的检索返回文本及研究题录/摘要；明确误差诊断属于后来重建研究，非本书独立复算，也不归为 Richardson 当年已完成的解释。正文未使用原算例 145hPa 具体数值，以免仅据摘要重建原表。

[R38c] Charney, J. G., Fjørtoft, R., & von Neumann, J. (1950). Numerical Integration of the Barotropic Vorticity Equation. Tellus 2, 237–254. DOI：https://doi.org/10.3402/tellusa.v2i4.8607
公开原文：https://maths.ucd.ie/~plynch/eniac/CFvN-1950.pdf
实际核读：原文 PDF 首张页面图及完整摘要，确认四次 24 小时、500mb、有限区域边界与差分稳定性；未逐图复算。正压涡度式、CFL 讨论和计算成本例均为本书教学说明。

[B17 补核] Lorenz, E. N. (1963). Deterministic Nonperiodic Flow. Journal of the Atmospheric Sciences 20, 130–141. URL：https://cdanfort.w3.uvm.edu/research/lorenz-1963.pdf
出版入口：https://doi.org/10.1175/1520-0469(1963)020%3C0130:DNF%3E2.0.CO;2
实际核读：原文 PDF 方程页面截图（PDF 第 6 页）与论文摘要；核对三变量原式及参数性质。平衡点、散度、Jacobian 和误差阈值式为本书直接教学推演，不声称原文全部证明复核。

[R39a] 本书教学计算：经典 Lorenz 三变量系统有限时间扰动比较，2026-10-16 实际执行。
参数 σ=10、r=28、b=8/3；初值 (1,1,1) 及 (1+10^-5,1,1)；经典 RK4，步长 0.01/0.005/0.0025/0.00125，终点 20；以欧氏距离记录 5/10/15/20 处差异，同时比较相同步长与不同步长轨道。
证据入口：附录四提供参数、最小完整脚本与实际结果表。主编已独立复跑核验一致。它是有限区间数值教学实验，不是 Lorenz 原算例复现、不计算严格 Lyapunov 指数、不构成数学混沌证明；没有把无量纲时间换成天气时距。

[R40a] ECMWF. Implementation of IFS changes (25 Nov 1997). URL：https://confluence.ecmwf.int/spaces/FCST/pages/304237824/Implementation%2Bof%2BIFS%2Bchanges%2B25%2BNov%2B1997
实际核读：官方实施记录正文，确认业务日期；另核 20 周年官方说明 https://www.ecmwf.int/en/about/media-centre/news/2017/20-years-4d-var-better-forecasts-through-better-use-observations。

[R40b] Modigliani, U. (2025). The history of ECMWF forecasts and UEFs. ECMWF 官方报告。URL：https://events.ecmwf.int/event/464/contributions/5296/attachments/3244/5410/UEF2025_Modigliani.pdf

实际核读：第 12 张（PDF 索引 11）的 1997 年 4D-Var 历史说明，明确初版采用以分析时刻为中心的 6 小时时间窗。未把后来 12 小时设置倒填初版。

[R40c] Goddard, J., Laloyaux, P., Lang, S., & Leutbecher, M. (2017). How to deal with model error in data assimilation. ECMWF Newsletter 153. URL：https://www.ecmwf.int/en/newsletter/153/news/how-deal-model-error-data-assimilation

实际核读：技术正文强/弱约束说明、2016 业务节点及观测误差误归因问题；只作有限历史事实使用，不据其 2017 限制判断 2026 当前状态。第 40 章代价函数与单变量例是本书教学表达，明确误差独立/分组假设，未假称 1997 生产代码逐项复现。

第 41—50 章来源与核读记录

本次审读日期：2026-10-16。C 代码继承原大纲来源总账；R50a、R50b 为本次审读新增。以下明确区分题录、摘要、正文局部与原图核读，不把取得文献入口等同于全文精读。正文一般技术推演均由本书自行讲解，不声称复现原始计算。

◆ C01 USGS：Watershed Boundary Dataset / Federal Standards and Procedures

URL：https://pubs.usgs.gov/tm/11/a3/

补充定义入口：https://water.usgs.gov/themes/hydrologic-units/

本次审读范围：官方定义、用途、层级及水文单元可仅为出口总汇水区一部分的说明。未通读全部技术规范，未下载比较某个真实出口的两版边界。

支持：41、44、45 章水文单元与完整流域区分。边界敏感性、地下交换、事件资料比较均为分析框架或明确教学情境，不报告虚构实测结果。

◆ C02 Horton (1933), The Rôle of infiltration in the hydrologic cycle

URL：https://agupubs.onlinelibrary.wiley.com/doi/10.1029/TR014i001p00446

本次审读范围：原出版方题录与摘要检索内容；未逐页读取全部原文。

支持：1933 年入渗研究作为水循环思想史节点。正文未将后世所有霍顿式概念、公式或试验细节归于这篇原文。线性储库、前期状态、卷积与验证分析是本书现代教学论述。

缺口：公开真实事件的独立计算仍未实施，三十毫米例子明确为教学设定。

◆ C03 Sverdrup (1947), Wind-Driven Currents in a Baroclinic Ocean; with Application to the Equatorial Currents of the Eastern Pacific

URL：https://pmc.ncbi.nlm.nih.gov/articles/PMC1079064/

DOI：10.1073/pnas.33.11.318

本次审读范围：PMC 题录、研究范围、原始扫描页与 PDF 入口确认；未逐页审定推导。

支持：1947 年论文题名、风驱斜压与东太平洋赤道海流的研究定位。

边界：βV 与风应力旋度、地转切变、混合及输运均明确作为本书现代符号教学推演；没有冒称逐式翻译原文，也不将后来边界流理论归入原论文。航测重排史料未核。

◆ C04 NOAA NCEI, What Are Proxy Data?

URL：https://www.ncei.noaa.gov/news/what-are-proxy-data

补充档案说明：https://www.ncei.noaa.gov/products/paleoclimatology

本次审读范围：官方代理来源、天然记录与档案类别的可读说明；另核树轮档案区分原始宽度、密度、同位素与重建产品。

支持：46 章天然记录来源分类。一般校准、误差与年代分析为本书教学分析，不以官方科普说明替代具体代理实证论文。

◆ C05 Hays, Imbrie & Shackleton (1976), Variations in the Earth's Orbit: Pacemaker of the Ice Ages

原文 URL：https://www.whoi.edu/cms/files/hays76sci_268464.pdf

DOI：10.1126/science.194.4270.1121

本次审读范围：题录；原始扫描 PDF 第 2、4、12 页（PDF 页序，分别含论文开篇、年代/频率相关材料、结论附近内容）视觉核读。文本抽取主要为封面，未声称全文精读。

支持：约五十万年研究范围、轨道频段关系与内部非线性问题的历史定位。具体数值未重算，正文不引原图点值。

边界：谱分析、相位、线性储库传递函数、年代独立性为自行分析，未重新计算原始频谱。

◆ C06 Manabe & Wetherald (1967), Thermal Equilibrium of the Atmosphere with a Given Distribution of Relative Humidity

原文 URL：https://www.gfdl.noaa.gov/bibliography/related_files/sm6701.pdf

补充可检索同原文副本：https://climate-dynamics.org/wp-content/uploads/2016/06/manabe67.pdf

本次审读范围：原始 PDF 第 1、2、19 页视觉核读，包含摘要、引言、模型条件开头与参考文献；检索可见有关湿度敏感性的正文片段。未逐式审定或独立计算。

支持：规定相对湿度、辐射—对流平衡、控制条件与敏感性比较的历史定位。未引用精确增温值。

边界：$N = F - \lambda \Delta T$、单储库响应、链式求导为本书教学公式，不归为原文原式；云、海洋与缓慢反馈作为扩展问题而非原论文全部成果。

◆ C07 Farquhar, Bao & Thiemens (2000), Atmospheric Influence of Earth's Earliest Sulfur Cycle

原文 URL：https://www.eoas.ubc.ca/~mjelline/212_website/resources/Required-Reading/Farquar_2000_Atmospheric-Influence-of-Earth%E2%80%99s-Earliest-Sulfur-Cycle.pdf

DOI：10.1126/science.289.5480.756

本次审读范围：原文可检索文本中题名、摘要及相关正文，重点为多同位素关系、气相反应、早期样品与2090—2450 Ma 宽转变范围、混合与反应停止的替代解释。PDF 第一页上部有另一篇论文，已识别并排除。

支持：49 章原论文历史结论。正文保留该宽年代为当时判断，不冒充最新氧化史精确边界。

边界：没有重算样品或逐矿物验证；现代大氧化年代修订与具体生命因果不作确定断言。

◆ C08 Earth System Science: A Closer View (1988)

URL：https://www.nationalacademies.org/publications/19088

DOI：10.17226/19088

本次审读范围：官方出版题录、年份、原报告身份与在线阅读入口；未全文读完 208 页报告。

支持：50 章作为 1988 年纲领节点。正文不从书目推断具体项目全部实施。

◆ C13 NASA (1989), Earth system science: A program for global change

URL：https://ntrs.nasa.gov/citations/19890013598

报告号：NASA-TM-101186。

本次审读范围：NASA 正式题录与摘要全文，核出版日期 1989-01-01，摘要列长期测量、地球历史、过程研究、系统模型、信息系统与机构合作。

支持：50 章研究计划维度；与 1988 年报告严格分开，不当作同一题录。

◆ R50a NASA Terra mission and launch records

题名：Mission | Terra；Terra Launches!

URL：https://terra.nasa.gov/about/mission

发射同期记录：https://science.nasa.gov/earth/earth-observatory/terra-launches-499/

本次审读范围：NASA 任务说明与 1999 年发射记录，核 1999 年 12 月 18 日发射、五台仪器与跨圈层观测定位。未读取全部仪器算法文档。

支持：50.3 中设施实施的具体节点，不证明 1988 年报告直接造成全部后续发展；不讨论 2026 实时任务运行状态。

◆ R50b NASA Earth Observatory (2000), HDF-EOS

URL：https://science.nasa.gov/earth/earth-observatory/hdf-eos/

本次审读范围：官方历史说明中 EOSDIS、分布式档案中心、HDF 作为 EOS 数据格式的选择与互通问题；未全文复查软件规范。

支持：50.3 数据组织与标准化的历史实例。后文语义、版本与守恒接口分析为本书独立技术阐释。

第 51—60 章来源与核读范围

核验日期：2026-10-16。正文沿用大纲代码 C01—C14、K01—K06、M01—M04、N03。大纲已有核读记录并非本章节单元再次阅读全文的声明；综合章依据前文建立的材料关系作作者分析，未执行原始资料复算、模型运行、独立盲评或专家评分。第 59 章成对短论证为本书分析，不冒称原作者用过 SDE。

◆ 新增与更新入口

R52a ECMWF. Destination Earth moves into its third phase. 2026-06-30. URL : https://www.ecmwf.int/en/about/media-centre/news/2026/destination-earth-moves-its-third-phase
实际核读：官方搜索结果返回完整新闻正文，核读第三阶段、前两阶段业务化进展、故事线试验、至 2028 年的阶段安排。只采用第三阶段与功能边界等少量事实，未视为独立应用成效评估。当前状态另经官方资料复核。

R52b ECMWF. Destination Earth Digital Twin for Climate Change Adaptation (DestinE Climate DT Generation 2). URL : https://www.ecmwf.int/en/forecasts/dataset/destination-earth-digital-twin-climate-change-adaptation-gen2

实际核读：已打开并读正文。页面列新模拟自 2026 年 5 月提供、全球气候模拟、1990 年 1 月至 2049 年 12 月、数据 DOI 10.21957/79c6af3105。未下载或运行数据，未自行评估性能。

R54a USGS. Volcano Aviation Codes and Alert Levels in English and Samoan.
URL : https://www.usgs.gov/media/files/volcano-aviation-codes-and-alert-levels-english-and-samoan
实际核读：官方检索结果明确说明地面火山状态与航空空中灰危害两部分；本次审读未逐页读取附件。正文仅以此支持两类对象的区分，不将现行规则倒投到 1989—1990 年。

◆ 本次审读再次核查的原有来源

C09 USGS. Landsat Collection 2.
URL : https://www.usgs.gov/landsat-missions/landsat-collection-2
实际核读：本次审读直接打开超时；大纲已记录技术正文核读。本次审读另核 USGS 以下发布资料与技术题录，仅用来确认重处理节点和改进类别，不报告性能幅度。补充官方入口：https://eros.usgs.gov/doi-remote-sensing-activities/2020/usgs/landsat-collection-2-advancing-interoperability-global-landsat-archive （搜索可见正文片段，确认 2020 年 12 月公开与第二次大规模重处理）。补充技术题录：https://pubs.usgs.gov/publication/70214657 （Landsat Collection 2 geometric calibration updates，搜索可见摘要）。

C14 USGS. Landsat Update Special Issue, December 2020.
URL : https://www.usgs.gov/landsat-missions/landsat-update-special-issue-december-2020
实际核读：本次审读官方检索片段确认处理、几何、辐射改进说明入口。未独立运行影像比较。

C11 Mousavi et al. (2020). Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking. Nature Communications 11,3952. DOI:10.1038/s41467-020-17591-w.

URL : https://www.nature.com/articles/s41467-020-17591-w
公开存档入口：https://pmc.ncbi.nlm.nih.gov/articles/PMC7415159/
实际核读：出版者搜索返回摘要与方法片段，明确单台局地地震、联合检测与 P/S 拾取、注意力机制；直接打开被身份验证跳转阻断。未声称全文精读，未重复性能数字。第 53 章跨台关联、定位、盲化复核、目录比较与教学例子属于作者分析与后续核验设计，不全归于原论文。

C12 Guffanti, M. C., & Miller, T. (2013). A volcanic activity alert-level system for aviation: Review of its development and application in Alaska. Natural Hazards,69,1519–1533. DOI:10.1007/s11069-013-0761-4.
URL : https://pubs.usgs.gov/publication/70188536
另一个官方入口：https://www.usgs.gov/publications/a-volcanic-activity-alert-level-system-aviation-review-its-development-and-application
实际核读：官方搜索返回题录与摘要，确认 1989—1990 雷道布特喷发期间设计航空警报制度。直接打开题录 403。没有逐条原始通报核读，不作事故责任裁决、不虚构人物心理。

◆ 综合章研究边界

56 章：比较重测、重取样、模型可复原性与历史事件，属于作者方法分析；未重新计算轨道谱、硫同位素或辐射平衡。
57 章：对水文、辐射、版本提供成对概念解释，明确重叠与组织价值；无独立评阅结果。
58 章：六路径作为复审入口；P1/P2/P3 允许未发生与证据不足；未冒称研究者心理顺序。
59 章：煤山与板块两个案例均提供基线/SDE 成对短论证，明确两者实质重叠。无年龄复算或逐事件震源机制复算，无 SDE 胜率。
60 章：四项后续研究为可执行设计，不是已经完成的结果。

版本说明与正文核算

本版在全书初稿基础上完成六组章节精修和统一审读。修订重点是将重复的方法声明落实为各领域的推断条件，补清变量与尺度，调整段落位置，贯通章际问题。导论集中交代共同框架，结语区分已经支持的结论与继续研究的任务。

本版正文共 309,511 汉字：导论 7,944 汉字，六十章 297,505 汉字，结语 4,062 汉字。各级标题、目录、参考资料与附录不计入正文；数字、外文与标点不计入汉字。