

缩短生物工艺开发时间并提高效率，释放混合建模和工艺强化的力量

Thomas Zahel, Nawaz Mohammad

柯尔柏医药科技

摘要：许多制药公司正面临着缩短工艺开发时间和高效优化工艺的挑战，希望以更快的方式将产品推向市场。本文介绍了一种动态建模方法，通过利用工艺强化技术来简化实验，并将其与混合模型得出的直观结果相结合，可以减少实验次数，有助于开创生物制品工艺快速且高效开发的新时代。

关键词：动态建模；工艺强化技术；混合模型；生物工艺开发

生物工艺开发瞬息万变，在保证产品质量和工艺效率的同时，加速工艺开发是当务之急。这不仅有助于降低药品销售成本（Cost of goods sold, COGS），还可以使患者更快地获取药物，实现双赢。当今市场对挽救生命的药物的需求不断增加，因此越来越多的先进的工具和技术被应用于药品

生产工艺的开发之中，人们采用了各种创新的方法，力图加快开发的速度。本文展示了其中一种方法——动态建模，它可以减少实验次数，从而加速工艺开发。

1 充分利用所有数据

参见图 1A 中的上游批量发酵或

图 1B 中的制备色谱图：大多数执行的工艺生产步骤都具有“时间动态（time-dynamic）”的特性，如发酵的最终产品数量或色谱制备环节的杂质池浓度，这些最终特征值是时间演变的结果。可以利用诸如柯尔柏的维隆 PAS-X Savvy 这样的工具分析和处理一段时间内来自不同数据源的数据并识别趋势，

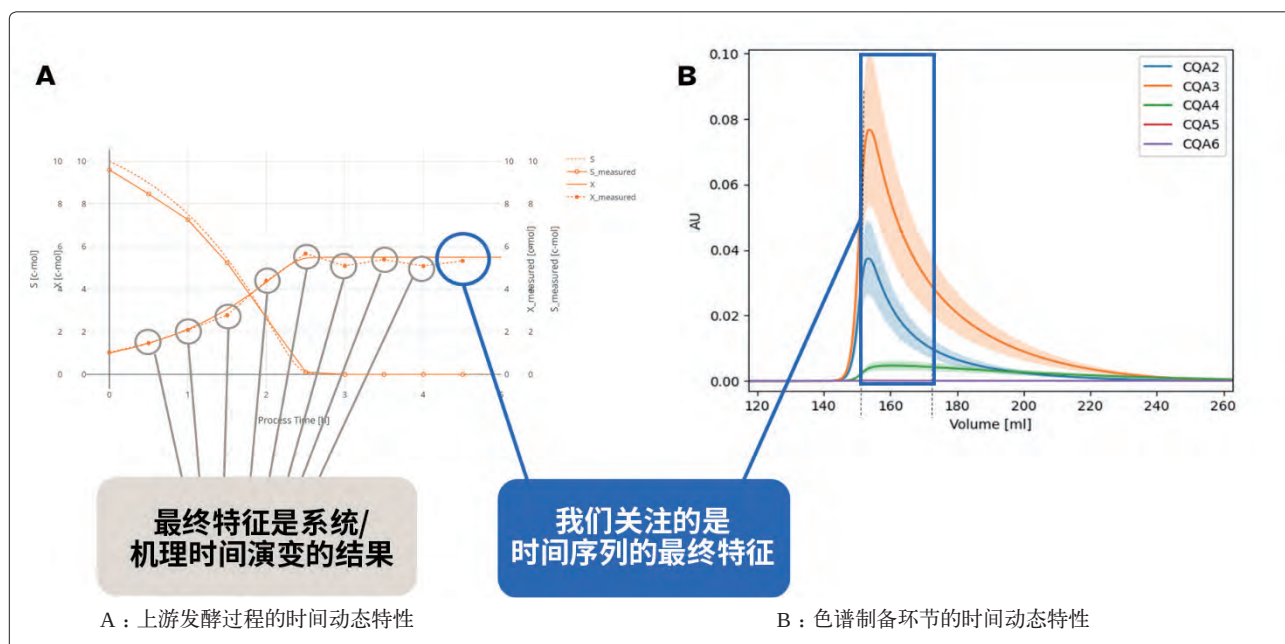


图 1 大多数执行的工艺生产步骤都具有“时间动态”的特性

作者简介：Thomas Zahel，柯尔柏医药科技 CMC 创新咨询部门的负责人；Nawaz Mohammad，柯尔柏医药科技的高级咨询顾问。

并在其基础上预测最终特征。通过使用滤波器（如卡尔曼滤波器）耦合的预测，可以增加最终特征值的精度，并因此增加识别临界效应的统计功效；或者相反，可减少识别这种临界效应所需要的实验次数。这便是动态建模可以减少实验数量并加速工艺开发的第一大原因。

2 作为主力工具的动态和混合建模

顾名思义，混合建模（hybrid modelling）是机理模型和数据驱动模型的结合，并且模型基于数据学习机理，更加灵活。基于第一性原理的机理模型，也称为“白箱”，能够解释所涉及的潜在生物和化学现象。通常，生物制药生产中使用的模型都不是纯粹基于第一性原理的，因为反应机制和结合机制不是根据第一性原理获知的。所有这些所谓的“机理模型”都是基于经验方程，如莫诺动力学（Monod kinetics），来评估反应机制的。数据驱动模型则是基于工艺过程中所产生的实验数据建立的。

典型的混合模型 workflow，由两个部分组成：

- 使用经验性信息（即数据驱动方法）估算特定反应速率；
- 将上述这些估计值输入到机理模型中进行进一步分析 / 处理，如图 2 所示。

因此，混合模型可以利用机理模型和黑箱模型（即数据模型）的组合效应，更好地洞察整个工艺过程，从而产

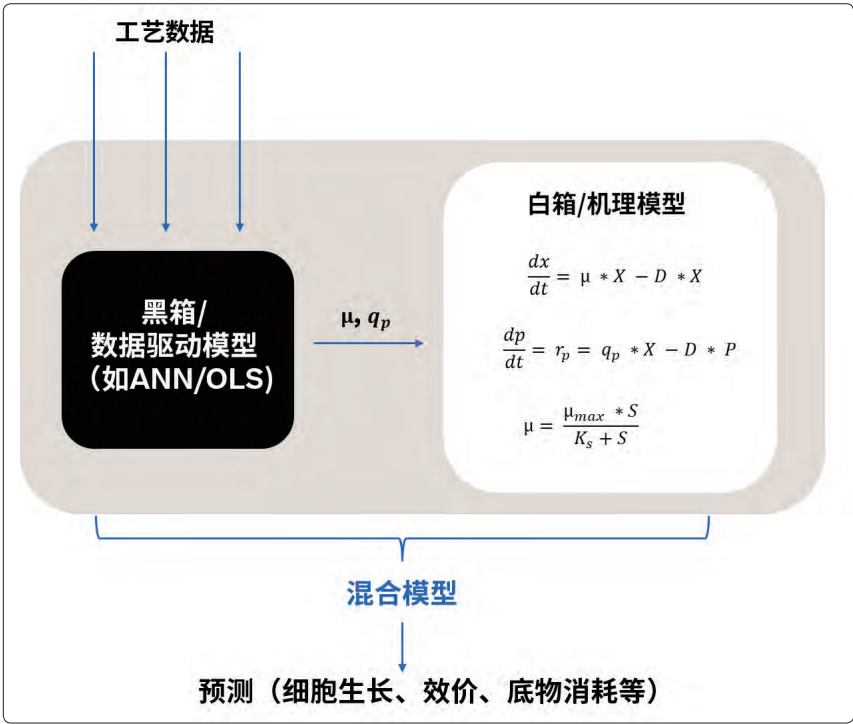


图 2 典型的混合模型结构

生稳健且高度可预测的模型。

3 工艺强化技术（Process intensification）

所有系统均由两种类型的变量组成：

- 工艺参数：可随时时间控制的工艺参数（如温度、进料量、洗脱梯度等）；
- 状态值：任何受工艺参数影响的结果（活细胞密度 VCD、细胞活率、效价、代谢产物、杂质洗脱谱图等）。

于是，基于给定初始条件（如初始活细胞密度 VCD、代谢物等）和随时间变化的可控工艺参数，就可以预测状态值。

如果想随着时间的推移改变工艺

参数，如进料曲线，则需在经典线性回归中引入多个周期 [如实验设计 (Design of Experiments, DoE)]。这也被称为时间序列的展开，如图 3 所示。在经典的 DoE 设定中，需要比模型中的效应多一次运行；因此，如果想将所有周期的影响都纳入模型中，至少需要 n+1 次运行，其中 n 是周期的数量。如果假设每个周期是独立的，即从一个周期到另一个周期过程中，没有影响下一周期的机理关系的记忆效应，则只需要一次运行来校准动态模型。

4 协同作用：将动态建模和工艺强化相结合，加速工艺开发

通过上述介绍，我们了解了动态



图 3 时间序列数据的展开

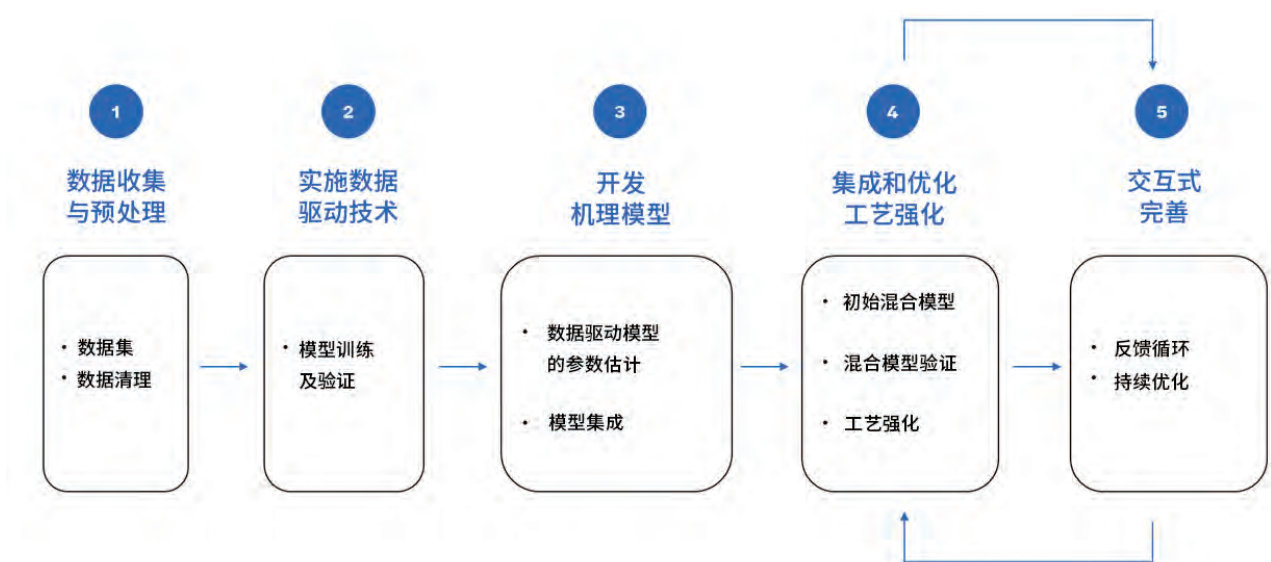


图 4 将混合模型和工艺强化结合在一起，加快早期工艺开发的通用工作流程

建模和工艺强化技术的优势，接下来有必要将这两种方法结合起来，并利用二者的潜力来取得更大的效果：制定一个能大幅缩短工艺开发时间的体系，从而更快地向患者提供药物。

为了建立这个体系结构，我们在图 4 中描述了通用工作流程所需要遵循的步骤。

第一步，数据收集与预处理，主要包括以下内容：

- 总体数据集：收集建模所需要的工艺参数和状态值的数据；
- 数据清理：确保数据准确、相关，并检查潜在的异常值；
- 借助 PAS-X Savvy 这样的计算机化系统平台，能够大幅减少人工收集数据的工作；同时，借助系统的数据标记与过滤功能，可以对大量外部数据实现统一的自动清理和标记，将原始数据转变为可用信息。

第二步，实施数据驱动技术，具体内容如下：

- 模型训练和验证：利用 PAS-X Savvy 等计算机系统平台的建模功能以及普通最小二乘 (Ordinary Least Square, OLS)、人工神经网络 (Artificial Neural Networks, ANN)

等多样的算法建立一个良好的预测模型，将反应速率描述为状态值和工艺参数的函数。

第三步，开发机理模型，主要包括以下内容：

- 参数估计：使用先前步骤中经过标记和清理的数据驱动计算机系统模型来估计机理模型的参数，如特定增长率、产物生成速率等；
- 模型集成和验证：将上述估计参数集成到机理模型中，并通过数据验证模型，以获得更好的预测效果和准确性。

第四步，通过工艺强化技术进行集成和优化，具体内容如下：

- 初始混合模型：使用实验数据运行该混合模型，并检查模型的预测效果和准确性；
- 工艺强化：优化混合模型，特别关注可能提高模型整体可预测性的参数组合。

第五步，迭代改进（交互式完善），主要包括以下内容：

- 反馈循环 (Feedback loop)：执行迭代程序，用其中混合模型的预测结果指导后续实验；
- 持续优化：以结果为基础，逐渐

提高混合模型的精确度和有效性。

5 结论：加速生物工艺开发的未来

当下，无论是工艺开发还是商业化生产，都需要争分夺秒。为了缩短早期工艺的开发时间，通过机理建模或者混合建模进行动态建模是一个很有前景的解决方案。通过利用工艺强化技术来简化实验，并将其与混合模型得出的直观结果相结合，研究人员可以在生物工艺开发中获得令人惊叹的效率。这种结合不仅显著缩短了开发时间，还能够确保开发出稳健、可扩展和可优化的工艺，从而尽可能地提高产量。通过使用 PAS-X Savvy 这样的数据管理与分析平台技术，能够对工艺开发和模型训练阶段的全部数据有效地进行管理，为后续的风险分析以及追溯提供坚实的基础。随着生物工艺行业的不断发展，采用这些创新技术将是保持竞争优势的关键，也是开创快速高效生物工艺产品开发新时代的关键。