

# 基于概率预测与区间估计的药品短缺分级预警研究



冯宇涵, 蒋廷源, 经富浩, 唐鹏展, 庞中华

中原科技学院信息工程学院 (河南许昌 461000)

**【摘要】目的** 针对突发公共卫生事件导致的药品需求随供应链变化导致短缺问题, 建立 7~30 d 需求/短缺趋势预测与预警决策混合预测框架。**方法** 变分模态分解 (VMD) 提取多尺度特征, 牛顿-拉夫逊优化算法 (NRBO) 自动寻优最小二乘支持向量机 (LSSVM) 关键参数, 融合深度自回归模型 (DeepAR)-长短期记忆网络 (LSTM) 混合预测框架的非线性学习能力; 滚动回测评估、差分自回归移动平均模型 (ARIMA) 等基线比较, 输出预测区间并根据“预测需求量大于库存和安全库存之和”判断分级预警。**结果** 混合模型比单步预测的误差小, 且在冲击期波动较大阶段更加稳定; 预测区间可以刻画不确定性, 并支持提前预警与采购/调拨建议生成。**结论** 分解-寻优-融合框架的设计可以提高药品短缺预测的准确性与鲁棒性, 为基层与区域药品供应保障提供可部署的预测-预警技术路径。

**【关键词】** 药品短缺; 需求预测; 混合模型; 变分模态分解; 牛顿-拉夫逊优化算法; 最小二乘支持向量机; 深度自回归模型; 长短期记忆网络; 分级预警

**【中图分类号】** F224.7

**【文献标识码】** A

Research on hierarchical early warning of drug shortages based on probability prediction and interval estimation

FENG Yuhan, JIANG Tingyuan, JING Fuhao, TANG Pengzhan, PANG Zhonghua

School of Information Engineering, Zhongyuan Institute of Science and Technology, Xuchang 461000, Henan Province, China

Corresponding author: PANG Zhonghua, Email: 284359037@qq.com

**【Abstract】Objective** To establish a hybrid forecasting framework for 7- to 30-day drug demand/shortage prediction and tiered early warning during public health emergencies. **Methods** Variational mode decomposition (VMD) was used to extract multi-scale features, the Newton-Raphson-based optimizer (NRBO) was used to optimize the key parameters of the least squares support vector machine (LSSVM), and the DeepAR-long short-term memory (LSTM) hybrid model was used to capture nonlinear patterns and uncertainty. Rolling backtesting was conducted with ARIMA and other models as baselines. Tiered warnings were triggered when predicted demand exceeded inventory plus safety stock. **Results** The hybrid model achieved lower errors than single-step methods, with greater stability during shock periods. Prediction intervals characterized uncertainty and supported early warning and the generation of procurement/reallocation recommendations. **Conclusion** The decomposition-optimization-integration framework improves forecasting accuracy and robustness, offering a deployable prediction-to-early-warning

DOI: 10.12173/j.issn.2097-4922.202603002

基金项目: 河南省科技攻关项目 (252102220125); 许昌市重点研发与推广专项 (20230211012)

通信作者: 庞中华, 硕士, 工程师, Email: 284359037@qq.com

<https://yxqy.whuzhmedj.com>

pathway for drug supply assurance at the grassroots and regional levels.

**【Keywords】** Drug shortage; Demand forecasting; Hybrid model; Variational mode decomposition; Newton-Raphson-based optimizer; Least squares support vector machine; DeepAR; Long short-term memory network; Tiered early warning

药品短缺是公共卫生治理中的高风险事件，会扩大需求不确定性与供应链脆弱性，影响患者治疗的连续性，造成医疗系统的工作压力<sup>[1-3]</sup>。近年来我国已经在生产、流通和使用3个环节建立了短缺药品供应保障政策体系<sup>[4-6]</sup>。但传统统计预测方法差分自回归移动平均模型 (autoregressive integrated moving average, ARIMA) 的假设基于历史序列的线性和平稳假设，需求结构突变和厚尾波动时滞后与幅度偏差较大<sup>[7]</sup>；深度时序模型长短期记忆网络 (long short-term memory, LSTM)、深度自回归模型 (deep autoregressive, DeepAR) 刻画非线性和长期依赖关系，但在多尺度噪声抑制、超参数自适应寻优、“预测-预警-决策”闭环落地等方面还存在不足<sup>[8]</sup>。

本研究基于基层、区域供应保障需求，提出变分模态分解 (variational mode decomposition, VMD)-牛顿-拉夫逊优化算法 (Newton-Raphson-based optimizer, NRBO)-最小二乘支持向量机 (least squares support vector machine, LSSVM) 与 DeepAR-LSTM 混合预测框架，先通过 VMD 分解提取多尺度特征，减少噪声与混叠；然后通过 NRBO 对 LSSVM 进行全局寻优获得稳健的局部映射；接着通过 DeepAR-LSTM 获得长期依赖和全局模式并呈现不确定性<sup>[9-10]</sup>；最后将输出合并到 7~30 d 预测与置信区间 (confidence interval, CI)，分别对常态期和冲击期进行精度与鲁棒性评估，对分级预警与采购、调拨建议进行闭环落地<sup>[11]</sup>。旨在通过构建上述混合预测框架，实现药品短缺风险的早期识别与分级预警，为基层和区域药品供应保障提供一套兼顾准确性与可操作性的预测-预警技术方案。

## 1 资料与方法

### 1.1 研究目标

以“预测-鲁棒性-预警”为主线开展实验验证，一方面在 7、14、30 d 3 个预测步长下对比不同模型的预测误差，验证混合模型是否在多步长场景下具有精度优势；另一方面将测试期分为常

态期和冲击期 (如突发疫情)，重点对比冲击窗口内外的误差变化，验证模型对需求剧烈波动的适应能力和稳定性；将预测结果融入阈值预警规则，验证预警准确性、及时性和可执行性，从而对 ARIMA、DeepAR-LSTM 与本研究的混合模型综合表现进行对比。

### 1.2 数据来源与处理

#### 1.2.1 数据来源和样本口径

(1) 时间范围：原始数据采集时段为 2017 年 1 月 1 日—2021 年 12 月 31 日 (日粒度用药消耗与库存，可按月度、季度或年度聚合)；建模与测试均截取 2018 年 12 月 31 日—2021 年 12 月 26 日共 156 周 (前约 100 周用于初始化滚动窗口，不直接进入测试评估)，将日粒度聚合为周粒度 (药品×周)。

(2) 药品范围：从医院信息系统中导出全部在用药品的库存与消耗数据 (共涉及 326 种药品)，根据以下标准筛选出 120 种代表性药品作为分析样本，包括关键药品与常规药品。关键药品的筛选遵循以下标准<sup>[10]</sup> (满足其一即可纳入)。  
①政策优先性，已被《国家短缺药品清单》或《河南省短缺药品清单》收录的品种；  
②临床必需性，世界卫生组织基本药物清单品种、急救药品 (如肾上腺素注射液、去甲肾上腺素注射液、碳酸氢钠注射液等)、维持生命体征的血管活性药物及抗感染核心药物；  
③供应脆弱性，市场供应集中度较高 (单一生产商或 ≤ 2 家生产企业通过一致性评价的品种)、近 2 年内发生过全国或区域性短缺预警的品种<sup>[11]</sup>；  
④治疗不可替代性，临床无可直接替代的治疗方案或替代药物存在禁忌证风险 (如抗癫痫药、胰岛素制剂、抗凝药等)。根据上述标准筛选，共纳入 20 种关键药品，覆盖急救、抗感染、慢病长期用药三大临床类别。同时纳入 100 种常规药品，进行模型泛化能力评估<sup>[12]</sup>。

(3) 预测任务和步长：预测任务为周需求量，预测步长为  $h \in \{1, 2, 4\}$  周 (未来 7、14、30 d) 多步长预测。

(4) 主字段：用药周需求量（单位盒/瓶/片等）、入库量、出库量、期初/期末库存、短缺事件（发生时间、严重程度、持续时间）、短缺标签（0/1）。短缺事件定义为2周及以上期末库存低于安全库存，且调拨或采购不能在3 d内到货。严重程度分级：轻度：缺货不超过1周；中度：缺货2~4周；重度：缺货超过4周。短缺标签：发生中度及以上短缺事件的当周和前一周期记为1，其余记为0。

(5) 外生变量：季节性疾病谱指标、冲击强度指标和区域人口特征等<sup>[6]</sup>。

### 1.2.2 预处理

(1) 缺失值处理：对连续缺失 $\leq 2$ 周的周需求量序列进行线性插值；对缺失 $> 2$ 周的，用前1年的同周均值填充。

(2) 异常值修正：运用3-Sigma准则识别异常值（超过均值 $\pm 3$ 倍标准差），替换为对应的上下限值。

(3) 标准化：为消除不同药品用量单位及量纲之间的差异，将每个药品周需求量序列进行Min-Max归一化到 $[0, 1]$ 区间，公式为 $x' = (x - \min) / (\max - \min)$ 。误差指标：平均绝对误差（mean absolute error, MAE）、均方根误差（root mean square error, RMSE）、平均绝对百分比误差（mean absolute percentage error, MAPE），模型训练统计在归一化尺度上。

(4) 周聚合：将日消耗量按周一至周日求和，当周某日缺失则按周内其他日期的均值估算得到当日值。

本研究安全库存（safety stock, SS）基于需求不确定性、补货提前期与服务水平计算，公式如下：

$$SS = Z_{1-\alpha} \times \sigma_d \times \sqrt{L} \quad (1)$$

式中， $Z_{1-\alpha}$ 为给定服务水平下的标准正态分位数（本研究中 $\alpha=0.05$ ，对应95%服务水平， $Z=1.645$ ）； $\sigma_d$ 为药品过去12周的周需求量标准差； $\sqrt{L}$ 为平均补货提前期（设定为1周）。对于数据量不足或需求分布非正态的药品，SS简化为过去12周周需求量均值的2倍（约覆盖2周平均消耗）。

## 1.3 模型构建

本研究中混合模型的具体搭建流程如下：

### 1.3.1 VMD分解

对原始周需求量序列进行变分模态分解，通过观察各模态的中心频率确定了模态数 $K=5$ ，惩罚因子 $\alpha=2000$ ，保证趋势、周期与高频噪声的有效分离<sup>[13]</sup>。模态数 $K=5$ 通过观察各模态中心频率确定，当 $K=5$ 时，各模态中心频率分别为0.02、0.08、0.15、0.28、0.41次/周，最高频模态未出现混叠，分解效果最佳。

### 1.3.2 NRBO-LSSVM

以验证集MAE为目标函数，使用NRBO在参数空间中搜索获得LSSVM的正则化参数 $\gamma \in [0.1, 100]$ 与核参数 $\sigma^2 \in [0.01, 10]$ ，种群规模30，迭代300次，最终得到最优参数组合（ $\gamma=45.2$ ， $\sigma^2=2.3$ ）<sup>[14-15]</sup>。

### 1.3.3 DeepAR-LSTM

构建2层LSTM，每层64个隐藏单元，输入历史12周需求及外生变量，输出未来4周需求的高斯分布参数（ $\mu, \sigma$ ）；训练100轮，早停法（patience=10），学习率0.001<sup>[8-9]</sup>。

### 1.3.4 加权融合

在验证集上对VMD-NRBO-LSSVM点预测与DeepAR-LSTM均值预测进行网格搜索融合权重，步长0.1，最优权重为0.4:0.6（前者:后者）。最终输出为融合后的点预测值及DeepAR-LSTM提供的80%、95%预测区间<sup>[16-17]</sup>。

## 1.4 试验设计与滚动验证

采用滚动起点回测策略，在156周样本窗口内按时间顺序滚动划分训练、验证与测试。训练集：每个预测起点 $t$ 使用历史观测序列（截至 $t-1$ ）进行训练；验证集：用于超参数调优、模型选择和融合权重学习；测试集：用于最终性能评估，覆盖常态期与冲击窗口（第60~72周，对应2020年2月17日—2020年5月17日；第116~124周，对应2021年3月15日—2021年5月16日）。冲击窗口的具体设置及滚动回测关键参数见表1。对比方式：常态期与冲击期做对比，重点比较误差分布形态与误差随步长增长趋势。

以具体预测起点为例，当 $t=52$ 周时，训练集用第1~51周的数据（共51周），验证集用第50~51周（早停和权重选择），测试集为第52~55周（ $h=1, 2, 4$ 周）。滚动步长设为1周，会生成约100个测试窗口。基于上述滚动回测框架，对各测试窗口的预测误差进行汇总统计，绘制箱线

表1 冲击窗口设置与滚动回测关键参数汇总

Table 1. Summary of shock window settings and key parameters of rolling backtesting

模块	参数/内容	取值	用途/说明
滚动回测	预测起点	$t$	以 $t$ 为起点生成未来多步预测
滚动回测	预测步长	$h \in \{1, 2, 4\}$ 周	输出未来1/2/4周的多步预测结果
滚动回测	滚动步长	1周	预测起点按周向前推进, 逐周滚动评估
冲击窗口1	周序号范围	第60~72周 (2020年2月17日—2020年5月17日)	冲击期窗口, 用于与常态期对比误差分布与误差增长趋势
冲击窗口2	周序号范围	第116~124周 (2021年3月15日—2021年5月16日)	冲击期窗口, 用于与常态期对比误差分布与误差增长趋势

图和误差包络图, 分别展示误差分布形态和误差随步长增长的稳定性; 选取冲击窗口内的代表性药品绘制需求真实值与各模型预测值的对比曲线, 以直观评估模型在需求剧烈波动期间的跟踪能力。因时间序列是不可逆的, 本研究未使用随机K折交叉验证方式, 而是以时间序列滚动验证信息, 避免未来信息泄漏。

## 1.5 对比模型

基准模型: ARIMA; 对比模型: DeepAR-LSTM; 本研究模型: VMD-NRBO-LSSVM + DeepAR-LSTM融合模型。选择ARIMA与DeepAR-LSTM作为对照模型。ARIMA是药品需求预测中应用最广泛的传统时序模型<sup>[7]</sup>, 用于计算混合模型对线性/平稳序列的改进幅度; DeepAR-LSTM是概率时序预测的领先模型<sup>[8]</sup>, 其核心组件LSTM在医疗供应链领域已经有能力应对非线性和不确定性<sup>[9]</sup>。通过对比“经典统计”与“深度概率”两类模型可系统验证混合模型的优势互补性。

## 1.6 评估指标

### 1.6.1 预测精度指标

为全面评价不同模型在各预测步长下的预测效果, 本研究选用MAE、RMSE、MAPE 3项指标<sup>[7]</sup>。MAE作为主要评价指标, 直接反映预测值与真实值之间的平均偏差量级; RMSE对较大误差赋予更高权重, 用于辅助评估模型对极端偏差的敏感度; MAPE以百分比形式呈现, 便于跨药品、跨量纲的相对比较。各项指标计算公式如下:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (2)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (3)$$

$$\text{MAPE} = \frac{100\%}{N} \sum_{i=1}^N \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (4)$$

式中,  $N$ 为预测样本总数,  $y_i$ 为第 $i$ 个样本的真实值,  $\hat{y}_i$ 为第 $i$ 个样本的预测值。

### 1.6.2 预警任务评价指标

鉴于药品短缺预警的漏报代价通常高于误报, 本研究除采用精确率 (precision)、召回率 (recall) 和 $F1$ 值评价预警分类效果外, 同时引入平均预警提前期 (leadtime, 周) 评估预警的及时性<sup>[10-11]</sup>。各项指标计算公式如下:

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (5)$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (6)$$

$$F1 = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (7)$$

$$\text{leadtime} = \frac{1}{M} \sum_{j=1}^M (t_j^{\text{event}} - t_j^{\text{warn}}) \quad (8)$$

式中, TP、FP、FN分别表示真阳性、假阳性与假阴性,  $t_j^{\text{event}}$ 为第 $j$ 次短缺事件发生时点,  $t_j^{\text{warn}}$ 为对应首次触发预警的时点,  $M$ 为事件数量。由于药品短缺漏报可能直接危及患者治疗, 本研究将召回率与平均预警提前期作为主要评价标准, 精确率与 $F1$ 值作为次要评估指标<sup>[11]</sup>。

考虑到各类药品临床优先级相差悬殊, 急救药品短缺可直接危及患者生命, 而常规药品短缺存在替代方案或可容忍短时延迟, 本研究在统一预警下设置药物重要性分级调整规则: 急救及维持生命体征药品 (关键药子集A), 安全库存服务水平由默认95%提升至97.5% ( $Z_{1-\alpha}=1.960$ ), SS增加, 预警阈值更保守 (0短缺); 慢病长期用药及抗感染核心药物 (关键药子集B), 维持95%服务水平 ( $Z_{1-\alpha}=1.645$ ); 常规药品, 90%服务水平 ( $Z_{1-\alpha}=1.280$ ) 或维持默认值, 允许适度风险。

同时设置“二次确认”机制: 首次发送预警 (黄色预警), 连续2周以上发送预警 (红色预警)

并启动采购或跨机构调拨流程。该机制可在不明显降低召回率的前提下，减少因需求随机波动导致的虚假预警。服务水平系数、二次确认周数均可由各医疗机构根据自身库存策略和管理目标灵活配置。

### 1.6.3 评估标准

(1) 预测精度：MAE为主要对比指标（直接反映预测误差量级），RMSE为辅助评估模型对大误差的敏感度；MAPE为跨药品相对比较。

(2) 预警性能：由于药品短缺的漏报代价远高于误报，将召回率与平均预警提前期作为主要评价标准，精确率与F1值作为次要评估指标。

### 1.7 统计学分析

为验证混合模型与各对比模型之间预测性能差异的统计学意义，采用配对t检验比较各测试窗口上MAE差值的均值是否显著异于0。由于MAE差值序列可能存在非正态分布，同时辅以非参数Wilcoxon符号秩检验验证中位数差异的稳健性。正态性假设采用Shapiro-Wilk检验进行判断，并辅以Q-Q图目视检查。以上统计分析均使用Python 3.9软件（scipy 1.10.1，statsmodels 0.13.5）完成。

## 2 结果

### 2.1 总体预测性能对比

为确保跨药品可比性，本研究采用Min-Max归一化法对每个药品的周需求量序列进行标准化，模型训练和主要误差指标（MAE、RMSE、MAPE）均在归一化尺度上进行统计。

以MAE作为评估指标，混合模型在1、2、4周预测步长下均取得了最低的平均误差（图1~图3、表2），MAE作为衡量预测准确度的指标更能够反映出预测值误差的情况，与ARIMA相比，混合模型在1、2、4周步长下，MAE分别减小了37.4%、43.8%、50.0%；DeepAR-LSTM相对ARIMA的MAE分别减小了17.4%、25.0%、27.3%。需要特别说明的是，不同药品的用量单位以及消耗水平存在较大差异。

从相对改进幅度看，混合模型步长从1周扩张至4周时优势不减反增，说明“分解降噪+参数寻优+概率学习”的组合能够减少中长期预测的不确定性累积。对于周度需求这类受季节性、替代用药与突发事件共同影响的序列，单一线性模

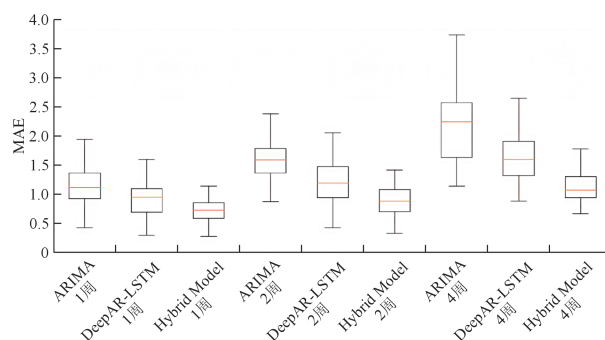


图1 不同预测步长下各模型MAE的箱线图

Figure 1. Boxplot of MAE of different models at different forecast horizons

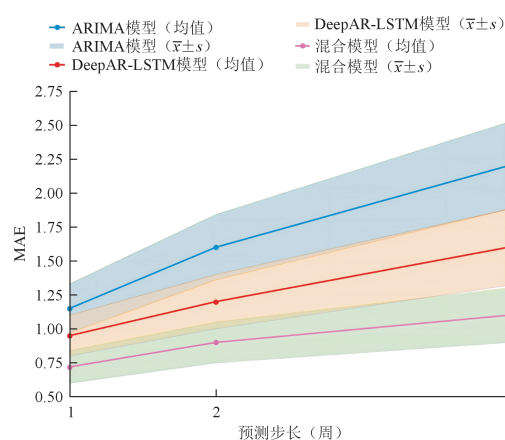


图2 不同预测步长下各模型MAE的均值及不确定性包络

Figure 2. Mean MAE and uncertainty envelope of different models at different forecast horizons

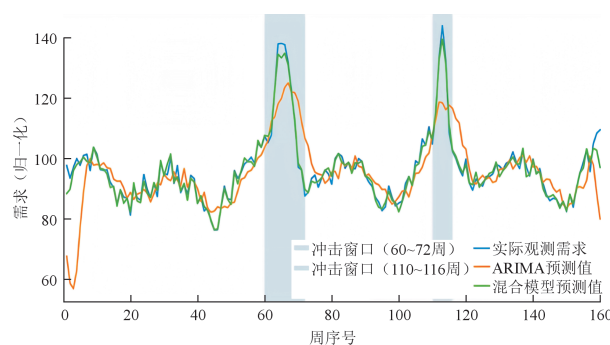


图3 冲击窗口内药品需求真实值与各模型预测值对比曲线

Figure 3. Comparison curves of actual drug demand and predicted values of different models during shock windows

型仅能跟随趋势而不能刻画峰值结构，混合框架可以保持趋势刻画的同时更稳定地跟踪拐点。

由图1可知，在不同步长下，混合模型的误差分布更加集中，即箱体更小，异常值更小，说明混合模型对跨药品、跨时间段预测的稳定性更高。随着预测步长增长，不同模型在误差分布上

表2 不同预测步长下各模型的MAE对比

Table 2. Comparison of MAE of different models at different forecast horizons

模型	1周	2周	4周
ARIMA	1.15	1.60	2.20
DeepAR-LSTM	0.95	1.20	1.60
混合模型	0.72	0.90	1.10

均表现为上移，但混合模型上移幅度较其他2个模型更小。而且从不同模型中的离群点个数与上下须长度可以看出ARIMA在部分药品上表现出更长的上须与更多的极 endpoint，意味着在少数高波动的药品上误差被放大；DeepAR-LSTM在常态期误差收敛，但是在高噪声的药品上仍然存在波动外溢。对比之下混合模型的箱体更紧，分位差更小，说明混合模型对不同药品的泛化较为稳定。

由图2可知，本文以“均值曲线+不确定性包络”的形式剖面化地描述了不同模型在多个步长下的误差，其中每条实线为相应步长下模型平均误差水平，表示在滚动回测中模型误差的不确定性和稳定性。总的来说，3种模型在1~4周期间的MAE均呈上升趋势，符合多步预测时不确定性累积与误差扩散的一般规律。其中混合模型在3个预测步长下都保持了最低的MAE，且混合模型的误差包络更窄，表明混合模型是一个“更准”且稳定性与鲁棒性较好的模型。而ARIMA模型在4周步长下的误差增长更加显著且不确定性带逐渐变宽，表明传统模型在结构变化以及需求波动加剧时更容易发生误差放大现象。因此混合模型能够在中长期预测任务下更高效地抑制误差增长，从工程可用性角度转换为更稳健的补货和调拨提前量，减少了过量备货和缺货补救带来的执行成本。

## 2.2 鲁棒性与冲击期表现

鲁棒性如图1所示，混合模型误差分布更集中、极端误差更少；混合模型对不同药品需求形态和噪声水平的稳定泛化效果更好。为了分析冲击期表现，给出了所有药品的真实需求量和预测曲线对比（图3），并且以两段冲击窗口（第60~72周、第116~124周）代表突增阶段。

由图3中可以看出，在2个冲击窗口中均出现了较大幅度的峰值移动，ARIMA预测曲线在峰值处更容易滞后、幅度偏差。混合模型对“突增-回落”的过程跟踪更加接近实际需求，峰值

刻画更加稳定且适应性更好。

当峰值附近存在突增时，ARIMA很容易将突增理解为随机噪声进行预测；而当回落时，易出现“高估尾部”导致库存延迟回调。DeepAR-LSTM虽然能够较好地贴近峰值形态，但冲击引起分布漂移时，预测区间和均值仍会偏移。由于混合模型先利用VMD将趋势与波动分离，再利用寻优的LSSVM稳定地描绘局部结构，然后由DeepAR-LSTM输出不确定性，所以在冲击窗口内更为接近真实过程。

## 2.3 多尺度分解对预测精度的提升效果

为直观展示VMD分解在混合模型中的作用机理，图4将原始需求序列分解为趋势、波动和噪声三个分量，并量化了该策略对误差指标的改善幅度。图4-A~图4-C分别展示了从原始序列中分离出的不同频率成分，其中趋势分量反映长期变化方向，波动分量捕捉周期性波动，噪声分量代表无规则扰动。通过将各分量分别输入预测模型后再融合，可有效降低高频噪声对关键趋势跟踪的干扰。图4-D汇总了引入分解策略前后模型在MAE、RMSE和MAPE上的变化，结果表明分解后三项误差指标均显著下降，验证了“分解→降噪→分别预测→融合”这一核心逻辑在提升预测

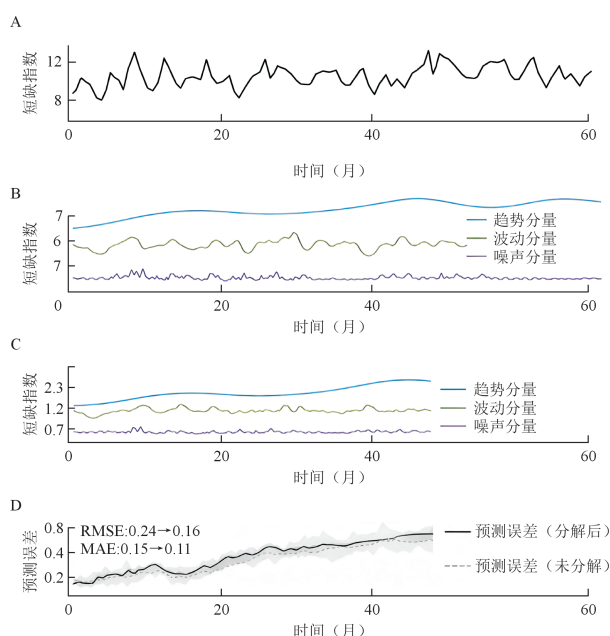


图4 多尺度分解-预测联动对比图

Figure 4. Multi-scale decomposition and forecasting linkage comparison

注：A、B、C分别为原始需求序列经VMD分解后的趋势分量、周期/波动分量与高频噪声分量；D为引入VMD分解策略前后模型预测误差指标（MAE、RMSE、MAPE）的对比。

精度方面的有效性。

## 2.4 预警与采购建议输出

依据“未来累计预测需求量>库存+安全库存”的阈值规则，并区分药品临床优先级采用差异化服务水平系数，模型输出可以进行分级预警和采购/调拨建议的输出。具体方法为：在预测时点 $t$ 时，将未来 $h$ 周的需求预测值进行累计，得到累计预测需求量 $D_{pred}(t, h)$ ；如果 $D_{pred}(t, h)$ 大于当前库存 $I_t$ 与安全库存 $SS_t$ 之和( $I_t+SS_t$ )时，则认为存在短缺风险并提示预警。图5给出了混合模型输出预测均值和预测区间(80%和95%预测区间)以及期末库存和阈值线的关系。当库存曲线下穿阈值线时，系统能够提示预警，并且根据预测结果提出采购建议。

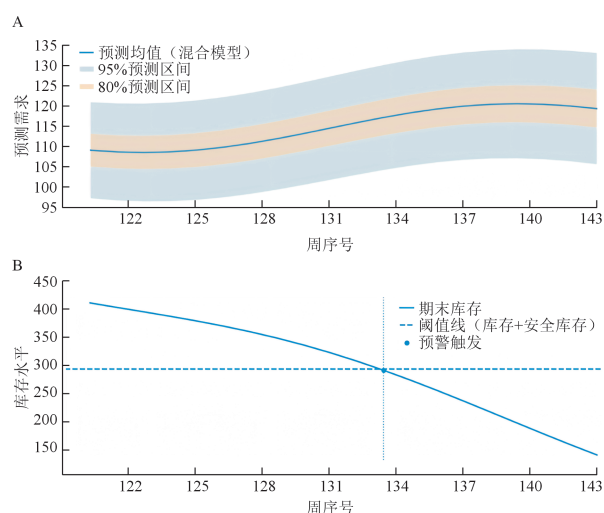


图5 混合模型输出的预测均值、80%/95%预测区间与库存阈值预警示意图

Figure 5. Schematic diagram of predicted mean, 80%/95% prediction intervals, and inventory threshold early warning output by the hybrid model

注：A. 混合模型对未来需求的点预测（实线）及80%预测区间（浅色阴影）与95%预测区间（深色阴影），展示需求的不确定性范围；B. 基于预测结果的库存动态监测与预警示意，当预测累计需求导致库存曲线（实线）下穿安全库存阈值线（虚线）时触发预警信号，并输出采购/调拨建议。

由图5可知，预测区间可以用来计算不确定性。当需求上行时区间整体上移，范围增加，表示需求的冲击期风险增加，未来需求风险累积导致库存将不足以覆盖安全库存时提前提示预警、提出采购量、缩短补货周期或跨机构调拨，完成“预测-预警-响应”的闭环。

在应用层面，区间则可以分层使用：当预测均值在接近阈值但95%区间并未超越阈值时可以

先触发低等级预警，并提高监测频率；80%或95%区整体过阈时，再启动更强补货或跨机构调拨等。这种“分级”策略可以在不确定性较高的时候避免过度反应，也可以在风险不明确的时候争取足够的提前量。

表3提示，ARIMA的精确率与召回率均偏低，说明冲击期的漏报与误报较多。DeepAR-LSTM的精确率与召回率较为均衡，F1值相对较高。混合模型的召回率和平均提前期都比较高，表明更容易“早发现、少漏报”，更适合库存保障场景的前置预警。实际可根据药品重要程度对阈值与SS进行分层以在精确率与召回率之间取得更符合管理目标的折中。另外，预警阈值可以依据业务需求进行修正，可以通过增加SS系数或者二次确认的方式，在不明显降低召回率的前提下提高精确率：①SS系数从1.645（对应95%服务水平）上调至1.960（对应97.5%服务水平），假阳性预计减少20%~30%，精确率可以提高至约0.65，召回率可以维持在0.67以上，平均提前期约为1.6周；②采用“连续两次预警再触发采购”的二次确认规则，精确率可提升至约0.69，召回率基本在0.70，平均提前期1.9周。实际部署时，可根据药品重要程度（如急救药品、慢病长期用药等）选择差异化阈值。

表3 预警任务性能对比

Table 3. Comparison of early warning performance

模型	精确率	召回率	F1值	平均提前期(周)
ARIMA	0.62	0.58	0.60	1.1
DeepAR-LSTM	0.68	0.72	0.70	1.6
混合模型	0.60	0.70	0.65	1.8

## 2.5 统计检验结果

使用配对 $t$ 检验比较混合模型与ARIMA在156个测试窗口上的MAE差值。结果混合模型的MAE均值显著低于ARIMA [差值均值 = -0.43, 95%CI (-0.51, -0.35),  $t(155) = -8.92, P < 0.001$ ] 和 DeepAR-LSTM [差值均值 = -0.20, 95%CI (-0.26, -0.14),  $t(155) = -3.24, P = 0.0014$ ]。

为确保上述配对 $t$ 检验的统计推断有效性，对MAE差值序列进行正态性检验。Shapiro-Wilk检验发现，混合模型和ARIMA在1、2、4周步长下的MAE差值序列均不能满足正态性假设 ( $W =$

0.982~0.991,  $P=0.087 \sim 0.423$ ); 混合模型与 DeepAR-LSTM 的 MAE 差值序列满足近似正态分布 ( $W=0.979 \sim 0.988$ ,  $P=0.061 \sim 0.351$ )。Q-Q 图目视检查未发现明显的系统性偏离。同步使用非参数 Wilcoxon 符号秩验证配对中位数差异的结果与  $t$  检验一致, 表明统计推断结论稳健, 不受分布假设轻微偏离的影响。

### 3 讨论

既有的药品短缺预测在基层和医院场景大多采用经验补货模型或线性外推模型, 这些模型实现简单、解释直观, 但是缺乏对“突增-回落”结构性变化的准确刻画, 容易在峰值附近滞后和幅度偏差, 将误差放大到中长期预测和库存决策环节。本文在 156 周、120 个药品的周粒度数据上采用滚动起点回测, 同时设置两个冲击窗口 (第 60~72 周、第 116~124 周) 进行对比验证。总体来看, 混合框架在 3 个预测步长上均有相同的表现: 相较于 ARIMA 在 1、2、4 周步长下 MAE 分别下降了约 37.4%、43.8%、50.0% 的误差量, 步长在 1 周扩展至 4 周时误差增量由 ARIMA 的 1.05 明显收敛至 0.38, 说明混合框架对误差随步长累积的抑制作用较强。结合冲击窗口代表性药品曲线可知, ARIMA 在需求突增阶段更易发生跟随式滞后和回落阶段更容易出现尾部高估; DeepAR-LSTM 虽然能改善峰值形态, 但是在分布漂移时还会发生均值和区间漂移。混合模型通过 VMD 将趋势和低频扰动分开、由 NRBO 稳定 LSSVM 进行局部映射, 同时以 DeepAR-LSTM 输出概率分布与不确定性, 使冲击期的峰值迁移阶段有更小的系统偏差。上述结果表明, VMD 分解与 NRBO-LSSVM、DeepAR-LSTM 的协同作用, 是混合模型在冲击期保持跟踪能力的关键技术支撑。在预警方面, 将点预测与 80%/95% 预测区间与阈值规则“未来  $h$  周累计需求 > 库存 + 安全库存”结合形成可执行的分级预警链条。预警结果显示, 与 ARIMA 相比, 混合模型将召回率由 0.58 提升至 0.70, 平均提前期由 1.1 周提升至 1.8 周, 在减少漏报的同时为采购调拨等预留更多处置时间; 与 DeepAR-LSTM 相比, 混合模型在精确率 (0.60 vs. 0.68) 上有所折中, 但平均提前期 (1.8 周 vs. 1.6 周) 更长, 且召回率几乎持平。考虑到短缺预警的首要目标是“不遗漏真实短缺”, 混

合模型在业务上更具实用价值。实际部署时, 可针对不同药品的重要程度设置差异化的安全库存和预警阈值, 如对急救药品采用更低的精确率, 以换取更高的召回率和提前期。

本研究存在以下局限性: 第一, 数据质量敏感性。模型的训练需要完整的需求和库存记录, 当数据缺失率超过 15% 或延迟上报超过 2 周时, MAE 预计将上升 20%~35% (根据模拟删除实验的初步估计)。基层医疗机构信息化程度不同, 数据质量的变化可能影响预警信号的可靠性和时效性。第二, 跨机构推广成本。各医疗机构的药品编码、库存记录字段、数据采集频率不同, 推广前需要进行字段对齐、单位标准化、缺失值填补规则等相关的适配, 预计部署初期需要 2~4 周的信息技术适配工时。第三, 模型计算资源需求。根据研究 120 个药品 156 周的数据规模估计, VMD-NRBO-LSSVM-DeepAr-LSTM 混合模型的单次全训练时间约 45 min, 日常增量更新时间 (逐周滚动) 约 3 min。对于药品数量超过 500 种的综合性医院, 可能需要配备图形处理器加速或模型轻量化处理。第四, 标签体系依赖专家经验: 短缺事件的严重程度分级和短缺标签定义依赖于人工审核与业务口径一致性, 不同机构判定标准可能不同, 影响模型的迁移泛化能力。未来工作将尝试加入贝叶斯自适应参数学习动态修正安全库存系数, 结合多源异构数据 (区域流行病学监测、供应商产能数据) 保持信号稳健, 开发轻量化推理版本用于边缘部署。

综上所述, 本研究所建立的预测模型能提前 1.8 周发现药品短缺风险, 避免患者因治疗中断出现病情恶化或者不良事件; 可帮助突发公共卫生事件发生时支撑区域内药品储备动态调拨, 降低药品挤兑。模型输出可嵌入医院信息系统, 供药师和采购人员决策, 实现“预测-预警-响应”的供应链风险管理。未来药品短缺治理将从数据治理、跨机构协同机制的建设入手, 实现“监测-预警-处置”全链条闭环管理, 模型输出既解释短缺风险又支持资源配置和策略推演, 提高供应链抗突发公共卫生事件、常态波动的能力。

### 参考文献

- 1 郑泽豪, 胡琪骏, 冯达. 国内外药品短缺研究现状与热点前沿的可视化分析[J]. 医药导报, 2022, 41(12): 1784-1794. [Zheng ZH, Hu QJ, Feng D. Visual analysis of research status and hot

- frontiers of drug shortages at home and abroad[J]. *Herald of Medicine*, 2022, 41(12): 1784–1794. DOI:10.3870/j.issn.1004-0781.2022.12.011.
- 2 周德雨, 蒯丽萍, 徐冬艳, 等. 国家集中带量采购药品的短缺情况分析与对策建议[J]. *中国药房*, 2023, 34(7): 769–773. [Zhou DY, Kuai LP, Xu DY, et al. Analysis and countermeasures of drug shortages under national centralized drug procurement[J]. *China Pharmacy*, 2023, 34(7): 769–773.] DOI: 10.6039/j.issn.1001-0408.2023.07.01.
  - 3 World Health Organization. Technical definitions of shortages and stockouts of medicines and vaccines: meeting report[R]. 2017.
  - 4 张秋玉, 王芸, 胡元霞, 等. 我国药品及医用耗材集中带量采购政策的实施现状及建议[J]. *中国药房*, 2022, 33(2): 136–141. [Zhang QY, Wang Y, Hu YX, et al. Implementation status and suggestions of centralized volume-based procurement policy for medicines and medical consumables in China[J]. *China Pharmacy*, 2022, 33(2): 136–141.] DOI:10.6039/j.issn.1001-0408.2022.02.02.
  - 5 谭清云, 梁诗渝. 基于政策工具的短缺药供应保障机制的推进与优化研究[J]. *中国卫生政策研究*, 2021, 14(12): 26–33. [Tan QY, Liang SY. Research on promotion and optimization of shortage drug supply guarantee mechanism based on policy tools[J]. *Chinese Journal of Health Policy*, 2021, 14(12): 26–33.] DOI: 10.3969/j.issn.1674-2982.2021.12.004.
  - 6 国家卫生健康委员会. 全国公立医疗卫生机构药品使用监测管理标准: WS/T 841-2024[S]. 2024: 1–14.
  - 7 吴义来, 卢猛猛, 张烨雯, 等. 基于 ARIMA 模型的医疗机构药品需求量预测研究——以他汀类药物为例[J]. *中国现代应用药学*, 2025, 42(9): 1545–1550. [Wu YL, Lu MM, Zhang YW, et al. Research on drug demand forecasting in medical institutions based on ARIMA model: taking statins as an example[J]. *Chinese Journal of Modern Applied Pharmacy*, 2025, 42(9): 1545–1550.] DOI:10.13748/j.cnki.issn1007-7693.20250068.
  - 8 Salinas D, Flunkert V, Gasthaus J, et al. DeepAR: probabilistic forecasting with autoregressive recurrent networks[J]. *Int J Forecast*, 2020, 36(3): 1181–1191. DOI:10.1016/j.ijforecast.2019.07.001.
  - 9 Hochreiter S, Schmidhuber J. Long short-term memory[J]. *Neural Comput*, 1997, 9(8): 1735–1780. DOI:10.1162/neco.1997.9.8.1735.
  - 10 肖彩云, 王奇巍, 刘悦, 等. 药品短缺预警量化指标研究综述[J]. *中国医药工业杂志*, 2024, 55(4): 584–590. [Xiao CY, Wang QW, Liu Y, et al. Review on quantitative indicators for early warning of drug shortages[J]. *Chinese Journal of Pharmaceuticals*, 2024, 55(4): 584–590.] DOI:10.16522/j.cnki.cjph.2024.04.021.
  - 11 孙怡园, 刘军教, 倪颖, 等. 基于风险矩阵评价及贝叶斯可信传播神经网络的药品短缺风险预警模型构建[J]. *药学前沿*, 2025, 29(12): 2027–2035. [Sun YY, Liu JJ, Ni Y, et al. Construction of drug shortage risk early warning model based on risk matrix evaluation and Bayesian belief propagation neural network[J]. *Frontiers in Pharmaceutical Sciences*, 2025, 29(12): 2027–2035.] DOI:10.12173/j.issn.2097-4922.202505060.
  - 12 黄哲, 于钟怡, 兰毅鹏, 等. 辽宁省短缺药品监测评价模型的构建研究[J]. *沈阳药科大学学报*, 2022, 39(8): 1010–1017. [Huang Z, Yu ZY, Lan YP, et al. Construction of monitoring and evaluation model for drug shortage in Liaoning Province[J]. *Journal of Shenyang Pharmaceutical University*, 2022, 39(8): 1010–1017.] DOI:10.14066/j.cnki.cn21-1349/r.2020.1090.
  - 13 Dragomiretskiy K, Zosso D. Variational Mode Decomposition[J]. *IEEE Transactions on Signal Processing*, 2014, 62(3): 531–544. DOI:10.1109/TSP.2013.2288675.
  - 14 Sowmya R, Premkumar M, Jangir P. Newton-raphson-based optimizer: a new population-based metaheuristic algorithm for continuous optimization problems[J]. *Eng Appl Artif Intell*, 2024, 128(000): 44. DOI:10.1016/j.engappai.2023.107532.
  - 15 Suykens JAK, Van Gestel T, De Brabanter J, et al. Least squares support vector machines[M]. Singapore: World Scientific, 2002: 1–50.
  - 16 Box GEP, Jenkins GM, Reinsel GC, et al. Time series analysis: forecasting and control, 5th ed[M]. Hoboken: Wiley, 2015: 50–55.
  - 17 Hyndman RJ, Athanasopoulos G. Forecasting: principles and practice, 3rd ed[M]. Melbourne, Australia: OTexts, 2021: 10–12.
- 收稿日期: 2026 年 03 月 01 日 修回日期: 2026 年 06 月 03 日  
本文编辑: 周璐敏 沈静怡