

时间验证的预测案例：后疫情时代的中国出境旅游预测



光华管理学院
Guanghua School of Management

李丰

feng.li@gsm.pku.edu.cn

<http://feng.li/>



- ① 研究背景
- ② 复苏预测
- ③ 我们的 RISE 框架
- ④ 回顾性分析

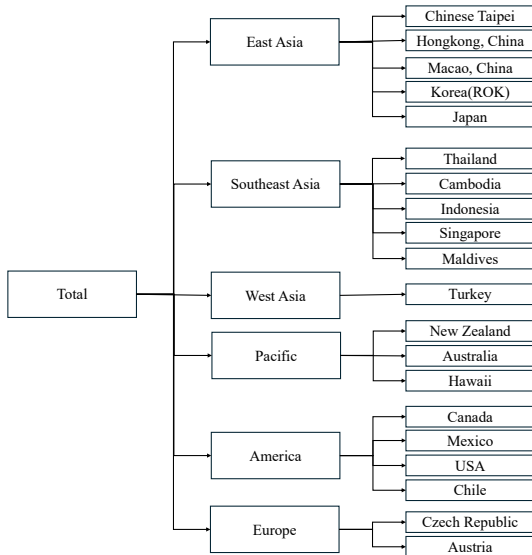




Figure: 我们的结果是全球预测团队中的点预测冠军 (Grand Prize)。

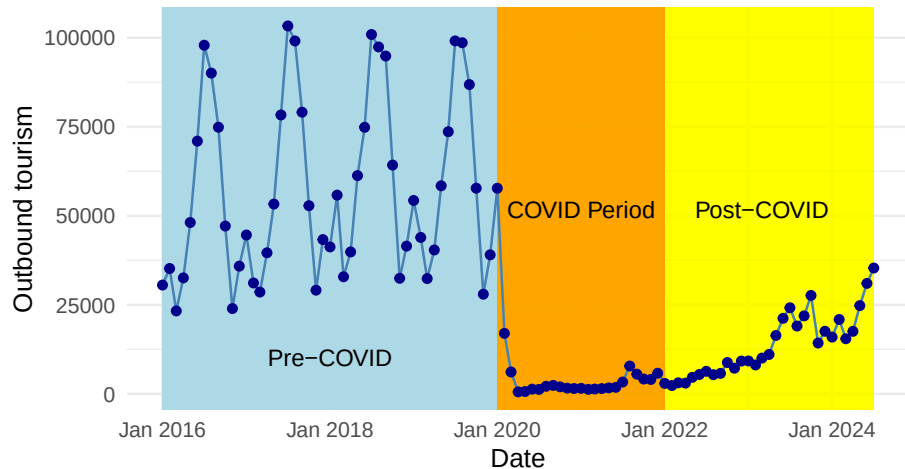


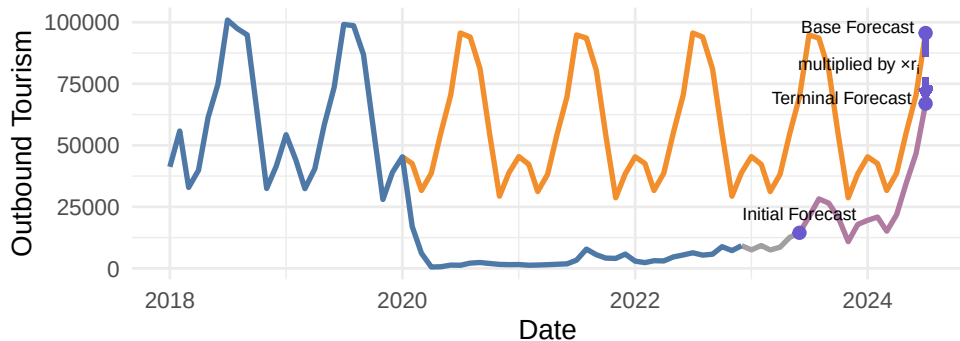
Figure: 我们的结果是区间预测赛道中荣获季军 (Second Runner-Up Prize)。

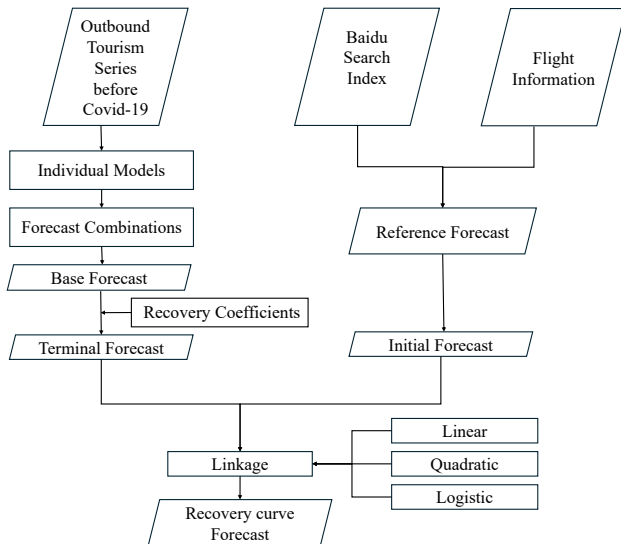
- 复苏预测 (**Recovery forecasting**) 是指在系统经历冲击后, 估计其回归至干扰前水平或达到新均衡状态的轨迹与时间过程的实践。
- 主要目标:
 - 预测恢复将于何时发生;
 - 估计恢复的路径与速度;
 - 识别影响恢复模式的关键因素;
 - 为规划、资源分配与政策制定提供支持。
- 典型应用领域:
 - 经济学: 预测经济衰退后的 GDP 或就业反弹 (Dominguez & Shapiro, 2013, AER);
 - 供应链: 估计中断后物流能力的恢复 (Ivanov & Dolgui, 2020);
 - 医疗健康: 预测危机后医疗服务水平回归基线的进程 (Kruk et al., 2017, BMJ);
 - 自然灾害: 预测基础设施或人口的恢复状况 (Gourio, 2008, AER);
 - 旅游业: 预测危机 (如 COVID-19) 后的游客到访数量 (本报告主题)。



- 基于模型的方法
 - 干预分析与中断时间序列分析 (Box & Tiao, 1975)
 - 状态转换模型 (Regime-Switching Models) (Hamilton, 1989; Kim & Nelson, 1999)
 - 基于离散状态马尔可夫过程的自回归模型 (Hamilton, 1989)
 - 动态因子模型 (Stock & Watson, 2002)
 - 向量自回归模型 (Karlsson, 2013)
 - 贝叶斯预测方法 (Koop & Korobilis, 2010)
- 局限性
 - 仅针对单个或少量时间序列的碎片化或特定领域方法。
 - 未利用外部信息作为恢复启发式因素。
 - 未充分利用预测组合的优势 (Wang, Hyndman, Li & Kang, 2023)







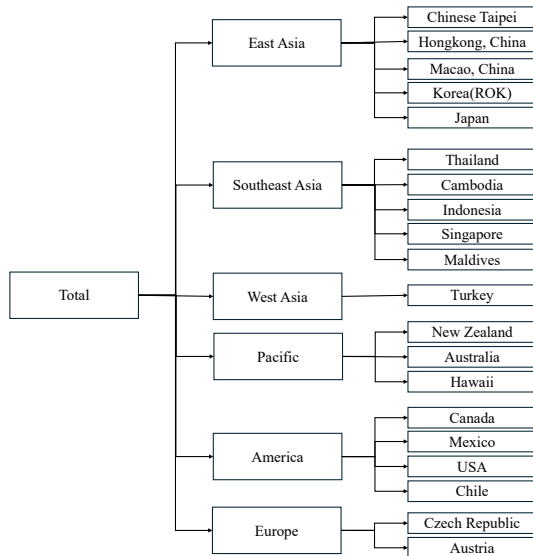


- 提出了一个综合性、模块化的三阶段框架，将结构突变、实时指标与恢复轨迹有机结合。
- 通过整合百度搜索指数和航班数据作为外生变量，利用恢复启发式信息实现前瞻性预测，并对预测能力进行比较评估。
- 基于地理距离、政策和恢复数据形式化定义恢复系数，并应用灵活的恢复曲线形式进行建模。

模型	描述	参数设定
SN	季节性朴素法 (seasonal naïve), 即预测值等于相同季节中最后一个观测值。	<code>snaive()</code>
RWF	带漂移项的随机游走模型, 相当于一个可选漂移系数的 ARIMA(0,1,0) 模型。	<code>rwf(drift=False)</code>
ARIMA	自回归积分滑动平均模型族 (Hyndman & Khandakar, 2008)。	<code>auto.arima()</code>
ETS	指数平滑状态空间模型族 (Hyndman et al., 2002)。	<code>ets(alpha=NULL, beta=NULL)</code>
HOLT	简单指数平滑扩展, 用于具有趋势的数据预测 (Holt, 1957)。	<code>holt(alpha=NULL, beta=NULL)</code>
HOLT-WINTERS	Holt 方法的扩展, 用以捕捉季节性 (Holt, 1957; Winters, 1960)。	<code>hw(alpha=NULL, beta=NULL)</code>
STL-A	使用 Loess 的季节与趋势分解方法, 其中趋势部分通过 ARIMA 模型估计, 季节部分采用:	<code>mstl(iterate=2, s.window=NULL)</code>
STL-B	(A) 过去三年的平均季节成分。	
STL-C	(B) 上一年的季节成分。	
TBATS	(C) 通过 ARIMA 估计季节成分。	
NNAR	具有 Box-Cox 变换、ARMA 误差、趋势与季节成分的指数平滑状态空间模型 (De Livera et al., 2011)。	<code>tbats(use.trend=NULL, seasonal.periods=NULL)</code>
	单隐层前馈神经网络, 以时间序列的滞后值作为输入。	<code>nnetar(p=12, P=1, size=7)</code>



- 前述的独立预测结果未考虑空间相关性。
- 来自 20 个目的地的旅游数据根据地理区域自然形成一个三层级的地理层次结构，这一结构有助于提升预测精度。
- 我们将这 20 个目的地划分为六个区域组：东亚、东南亚、西亚、太平洋、美洲和欧洲。
- 我们采用了在预测领域广泛认可的层次化预测方法；综述可参见 [Athanasopoulos et al. \(2024\)](#)。我们的做法主要包含两种层次化预测策略：自上而下（Top-Down）与最优一致化（Optimal Reconciliation）。



- 设 $\mathbf{y}_t$ 表示时刻 t 的所有时间序列向量（包括聚合层与底层）， $\hat{\mathbf{y}}_t$ 表示独立预测（可能不一致）， $\hat{\mathbf{y}}_t^{\text{rec}}$ 表示经一致化处理后的预测（即一致预测）。一致化预测可表示为：

$$\hat{\mathbf{y}}_t^{\text{rec}} = \mathbf{S}\hat{\mathbf{y}}_t.$$

- 其中， $\mathbf{S}$ 为表示层次结构的求和矩阵：

$$\mathbf{S} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ & & & & & & & & & \mathbf{I}_m & & & & & & & & & \end{bmatrix}.$$



- 自上而下 (**Top-Down**) 方法：首先在层次结构的最高层进行预测，然后根据历史信息或单个预测值估计的比例，将预测结果分配到低层。此时， $\mathbf{G}$ 为一个 $m \times n$ 的矩阵：

$$\mathbf{G} = \begin{bmatrix} p_1 & 0 & \cdots & 0 \\ p_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ p_m & 0 & \cdots & 0 \end{bmatrix}$$

- 最优一致化 (**Optimal Reconciliation**) 方法 (Wickramasuriya et al., 2019, JASA)：通过统计方法（如加权最小二乘法 WLS 或最小迹估计 MinT）对各个独立预测进行最优调整。MinT 方法的目标是设定 $\mathbf{G}$ ，使一致化误差的协方差矩阵的迹最小。其解为：

$$\mathbf{G} = (\mathbf{S}^\top \mathbf{W}^{-1} \mathbf{S})^{-1} \mathbf{S}^\top \mathbf{W}^{-1}$$

其中， $\mathbf{W}$ 为独立预测误差的协方差矩阵。

模型	描述	参数设定
TOPDOWN-A-ARIMA	层次化预测方法，通过层级结构分配顶层预测结果。在模型名称中，字母 A/B 表示比例分配方式，而 ARIMA/ETS 表示 TOPDOWN 方法的基础模型。分配比例由 (A) 各个独立预测结果确定。	<code>forecast(method="tdfp", fmethod="arima")</code>
TOPDOWN-A-ETS		<code>forecast(method="tdfp", fmethod="ets")</code>
TOPDOWN-B-ARIMA	(B) 历史出境旅游数据确定分配比例。	<code>forecast(method="tdgsa", fmethod="arima")</code>
TOPDOWN-B-ETS		<code>forecast(method="tdgsa", fmethod="ets")</code>
MIN-T	最小迹一致化方法，在保证一致性的同时最小化总体预测误差。	<code>forecast(method = "comb", weights = "mint")</code>
WLS	加权最小二乘法 (Weighted Least Squares)，通过为层次结构的不同层级分配权重以最小化一致化误差。	<code>forecast(method = "comb", weights = "wls")</code>

基准预测评估：验证集（2018~2019）中的 MAPE 与 MASE



模型	MAPE	MASE
SN	0.2679	1.4648
RWF	0.4158	1.8741
ARIMA	0.2430	1.2175
ETS	0.1494	0.9523
HOLT	0.4350	3.3307
HOLT-WINTERS	0.1851	1.1028
STL-A	0.1769	1.1512
STL-B	0.2249	1.2802
STL-C	0.1859	1.1830
TBATS	0.2796	1.4949
NNAR	0.1994	1.3337
TOPDOWN-A-ARIMA	0.2220	1.2520
TOPDOWN-A-ETS	0.1550	0.9860
TOPDOWN-B-ARIMA	0.4132	2.1673
TOPDOWN-B-ETS	0.4663	2.1810
MinT	0.1527	0.9861
WLS	0.1465	0.9961
简单平均 (Simple Average)	0.1609	0.9468
误差加权平均 (Error-weighted Average)	0.1816	1.1029
Stacking-Lasso	0.0961	0.5421
Stacking-Ridge	0.1051	0.5922

- 预测策略的第二阶段重点在于通过引入外部因素（如搜索指数数据与航班信息）来估计参考预测值。
- 该阶段对于提升初始预测的准确性至关重要。
- 我们利用相关外部指标，对参考期（2023 年 2 月至 7 月）的中国出境旅游人数进行预测，这一阶段的预测对整体准确率提升具有显著贡献。
- 使用的两种时间序列模型包括：**ARIMAX** 和 Facebook 的 **Prophet** 模型 (Taylor & Letham, 2018)。
- 我们使用 VariFlight APP（非常准）¹，提供 2019–2023 年的航班恢复参考信息。



参考预测：MASE 评估结果

目的地	最终值 (Final)	直航航班 (Direct flight)	综合搜索指数 (Composite search index)
加拿大	0.4708	0.8802	0.8085
智利	1.0963	—	1.0963
墨西哥	0.6944	—	0.6944
中国台湾	1.9542	1.3703	2.5381
中国香港	3.0323	3.8043	3.0323
日本	0.9430	0.9670	0.9190
韩国	0.2937	1.1461	1.3760
中国澳门	0.9888	1.0328	1.1545
马尔代夫	6.3710	11.2578	1.4841
柬埔寨	1.4194	1.3805	2.5733
印度尼西亚	3.0946	3.1808	4.4610
新加坡	0.9567	2.1413	1.1924
新西兰	1.5182	1.7451	1.2914
美国	1.0160	0.7334	1.3443
泰国	2.3467	5.5711	1.5206
土耳其	1.4537	3.3289	0.8521
澳大利亚	2.1087	4.8322	1.2256
夏威夷	1.6194	—	1.6194
奥地利	0.9694	1.2176	0.8884
捷克共和国	2.3953	—	2.3953
平均值 (Average)	1.7371	2.7868	1.6233
加权平均 (Weighted Average)	2.0248	2.6610	2.0779

表中同时列出了 MASE 的平均值与加权平均值，权重为 2023 年 2 月至 6 月期间的平均出境旅游人数。



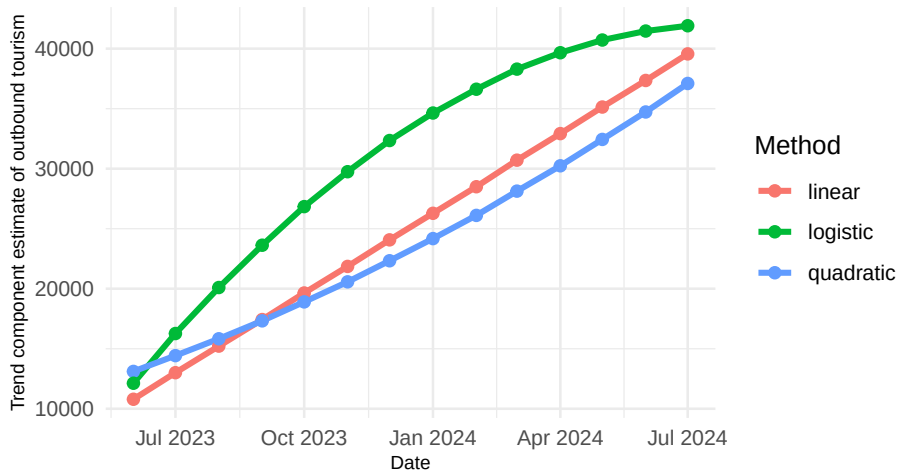
- 我们通过 $\tilde{F}_i^{13} = \widehat{BF}_i^{202407} \times r_i$ 来调整基准预测，其中 $0 < r_i \leq 1$ ，该系数用于向下修正预测值，以反映该时点预期的恢复程度。
- 由于中国出境旅游在短期内不太可能完全从新冠疫情的影响中恢复 (Škare et al., 2021)，因此 2024 年 8 月的基准预测必须加以修正，以考虑这一影响。
- 理想情况下，恢复系数应通过数据学习得到。然而在预测竞赛中，真实值 $\tilde{F}$ 是未知的，因此我们将恢复系数视为一个超参数，其取值来自先验知识。
- 我们根据地理距离与政策信息作为主要的先验依据，为各目的地设定参数。每个目的地只需为预测期终点指定一个恢复系数。

区域	目的地	恢复系数
美洲	加拿大	0.7
	智利	0.7
	墨西哥	1.0
	美国	0.65
东亚	中国台湾	0.6
	中国香港	0.85
	中国澳门	0.85
	韩国 (ROK)	0.8
	日本	0.8
东南亚	泰国	0.8
	柬埔寨	0.8
	印度尼西亚	0.8
	新加坡	0.8
	马尔代夫	0.8
太平洋地区	新西兰	0.65
	澳大利亚	0.8
	夏威夷	0.8
西亚	土耳其	0.75
欧洲	奥地利	0.65
	捷克共和国	0.65

- 我们采用乘法形式的 STL 分解来推导最终的恢复曲线，其形式为：

$$\hat{F}_i^t = \widehat{Trend}_i^t \times \widehat{Seasonal}_i^t$$

- 季节成分 $\widehat{Seasonal}_i^t$ 通过对 2020 年之前的历史出境旅游序列的对数形式进行乘法 STL 分解得到。
- 虽然存在其他替代方法可用于估计季节成分——例如使用 2023 年 8 月之后的基准预测——但我们认为季节效应在历史数据中更为稳定，而依赖基准预测可能会引入额外噪声。
- 在本研究中，我们提出恢复曲线的趋势部分可采用三种函数形式之一：线性、二次或逻辑斯蒂（Logistic）形式，并据此对趋势成分进行拟合。

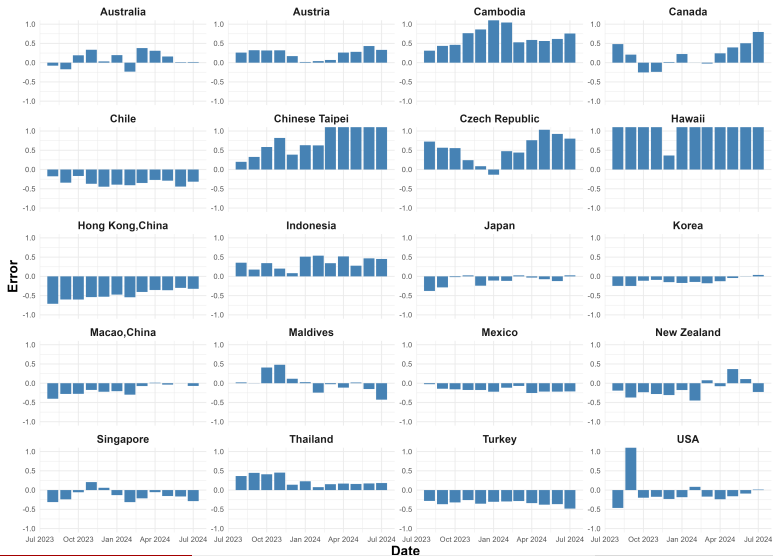


趋势成分：MASE 评估

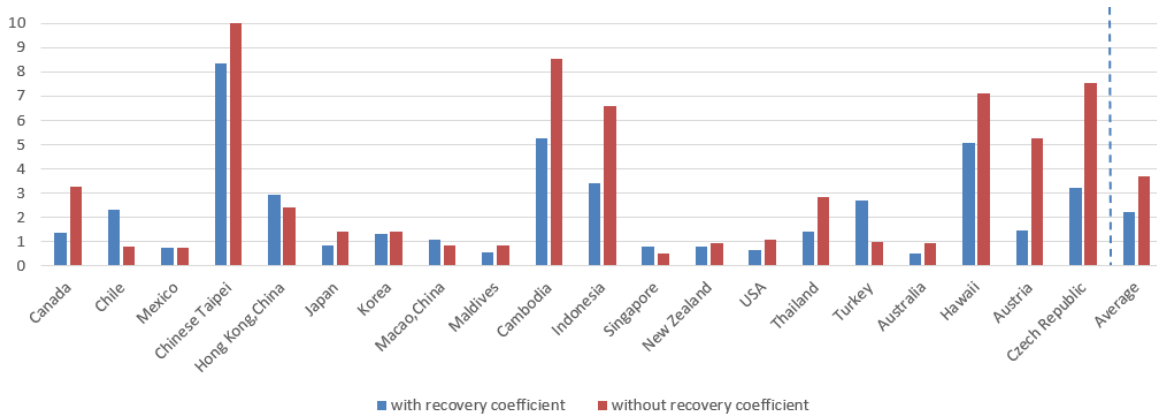
目的地	最终值 (Final)	线性 (Linear)	二次 (Quadratic)	逻辑斯蒂 (Logistic)
加拿大	1.3871	1.5731	1.3975	1.4430
智利	2.3311	2.0116	2.1836	2.7980
墨西哥	0.7668	0.4868	0.5054	1.4691
中国台湾	8.3389	6.6705	5.2014	13.1448
中国香港	2.9475	3.7195	4.0265	1.1836
日本	0.8486	0.9979	0.9479	1.6394
韩国	1.3411	0.8470	1.6343	2.5268
中国澳门	1.0657	0.9919	1.4723	1.0298
马尔代夫	0.5389	1.2788	0.5194	0.9948
柬埔寨	5.2619	4.7207	3.9109	7.1542
印度尼西亚	3.4175	6.4888	2.7449	1.5641
新加坡	0.7896	0.5813	0.9811	0.8821
新西兰	0.8169	0.6338	0.9715	1.2857
美国	0.6401	0.7420	0.8192	0.7203
泰国	1.4301	2.2996	1.0091	0.9816
土耳其	2.7209	2.1860	2.4690	3.5079
澳大利亚	0.4904	0.6411	0.4525	0.6395
夏威夷	5.0964	3.8336	3.3592	8.1253
奥地利	1.4804	2.4500	2.0315	1.7284
捷克共和国	3.2130	3.1708	3.1744	3.4261
平均值 (Average)	2.2461	2.3162	1.9906	2.8122
加权平均 (Weighted Average)	1.8479	2.1719	2.1848	1.4904

注：误差计算的时间范围为 2023 年 8 月至 2024 年 7 月。

2023 年 8 月至 2024 年 7 月的预测误差



含与不含恢复系数的 MASE 对比



基准方法与 RISE 框架的 MASE 比较



目的地	SARIMA	ETS	TBATS	干预模型	状态转换模型	RISE
加拿大	2.9667	2.2082	1.9100	2.4581	3.2378	1.3871
智利	5.1277	5.2081	5.1712	4.1005	3.2805	2.3311
墨西哥	1.9099	1.2430	1.5048	1.1106	2.4990	0.7668
中国台湾	9.3979	5.6816	5.6808	9.7510	6.2613	8.3389
中国香港	5.7767	6.0076	5.9116	6.2879	6.2879	2.9475
日本	7.3596	7.1312	7.1193	6.7520	5.8375	0.8486
韩国 (ROK)	12.8643	10.1430	10.3485	13.7329	10.7890	1.3411
中国澳门	3.3220	3.6869	3.7170	3.4781	4.0182	1.0657
马尔代夫	2.9486	2.7917	2.4113	2.7061	0.9282	0.5389
柬埔寨	3.6381	4.7759	4.5931	5.8800	6.8498	5.2619
印度尼西亚	7.7420	6.5528	6.5633	7.3903	4.2689	3.1475
新加坡	3.6390	3.3745	3.3896	3.3778	1.0681	0.7896
新西兰	3.2980	2.3175	2.7306	3.2857	1.3701	0.8169
美国	1.0768	1.8760	1.6801	1.0049	2.2823	0.6401
泰国	6.7429	6.6517	6.0061	6.6644	5.1343	1.4301
土耳其	5.3094	4.9897	5.1738	5.1808	2.7298	2.7209
澳大利亚	2.1565	2.4560	2.4519	1.3970	1.3821	0.4904
夏威夷	2.1750	0.4405	1.0382	0.9277	3.5185	5.0964
奥地利	5.2992	3.6351	3.0496	4.0602	3.1671	1.4804
捷克	2.2702	1.2652	1.4420	3.1012	2.4587	3.2130
平均值	4.7510	4.1218	4.0946	4.6324	3.8684	2.2326



- 我们提出的 RISE 框架具有简洁、通用且易于实现的特点。
- 可将启发式的恢复系数替换为数据驱动的动态估计方法，以进一步提升模型的自适应性。
- 该框架还可与优化与决策任务相结合，例如旅游推广规划或航线调度，从而构建以决策为中心的预测工具（待开发）。
- 代码已开源于：<https://github.com/feng-li/RISE-Forecasting>



感谢!

<https://feng.li>
feng.li@gsm.pku.edu.cn



Athanasopoulos, G., Hyndman, R. J., Kourentzes, N. & Panagiotelis, A. (2024). "Forecast Reconciliation: A Review". *International Journal of Forecasting* 40.(2), pp. 430–456.



Wang, X., Hyndman, R. J., Li, F. & Kang, Y. (2023). "Forecast Combinations: An over 50-Year Review". *International Journal of Forecasting* 39.(4), pp. 1518–1547.



Škare, M., Soriano, D. R. & Porada-Rochoń, M. (2021). "Impact of COVID-19 on the Travel and Tourism Industry". *Technological Forecasting and Social Change* 163, p. 120469.



Ivanov, D. & Dolgui, A. (2020). "Viability of Intertwined Supply Networks: Extending the Supply Chain Resilience Angles towards Survivability. A Position Paper Motivated by COVID-19 Outbreak". *International Journal of Production Research* 58.(10), pp. 2904–2915.



Wickramasuriya, S. L., Athanasopoulos, G. & Hyndman, R. J. (2019). "Optimal Forecast Reconciliation for Hierarchical and Grouped Time Series through Trace Minimization". *Journal of the American Statistical Association* 114.(526), pp. 804–819.



Taylor, S. J. & Letham, B. (2018). "Forecasting at Scale". *The American Statistician* 72.(1), pp. 37–45.



Kruk, M. E., Ling, E. J., Bitton, A., Cammett, M., Cavanaugh, K., Chopra, M., el-Jardali, F., Macauley, R. J., Muraguri, M. K., Konuma, S., Marten, R., Martineau, F., Myers, M., Rasanathan, K., Ruelas, E., Soucat, A., Sugihantono, A. & Warnken, H. (2017). "Building Resilient Health Systems: A Proposal for a Resilience Index". *BMJ* 357, j2323.



Dominguez, K. M. E. & Shapiro, M. D. (2013). "Forecasting the Recovery from the Great Recession: Is This Time Different?" *American Economic Review* 103.(3), pp. 147–152.



Karlsson, S. (2013). "Forecasting with Bayesian Vector Autoregression". In: *Handbook of Economic Forecasting*. Vol. 2. Elsevier, pp. 791–897. ISBN: 978-0-444-62731-5.



De Livera, A. M., Hyndman, R. J. & Snyder, R. D. (2011). "Forecasting Time Series with Complex Seasonal Patterns Using Exponential Smoothing". *Journal of the American statistical association* 106.(496), pp. 1513–1527.



Koop, G. & Korobilis, D. (2010). "Bayesian Multivariate Time Series Methods for Empirical Macroeconomics". *Foundations and Trends® in Econometrics* 3.(4), pp. 267–358.



Gourio, F. (2008). "Disasters and Recoveries". *American Economic Review* 98.(2), pp. 68–73.



Hyndman, R. J. & Khandakar, Y. (2008). "Automatic Time Series Forecasting: The Forecast Package for R". *Journal of Statistical Software* 27.(3), pp. 1–22.



Hyndman, R. J., Koehler, A. B., Snyder, R. D. & Grose, S. (2002). "A State Space Framework for Automatic Forecasting Using Exponential Smoothing Methods". *International Journal of Forecasting* 18.(3), pp. 439–454.



Stock, J. H. & Watson, M. W. (2002). "Forecasting Using Principal Components From a Large Number of Predictors". *Journal of the American Statistical Association* 97.(460), pp. 1167–1179.



Kim, C.-J. & Nelson, C. R. (1999). *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. Cambridge, Mass: Mit Pr. 250 pp. ISBN: 978-0-262-11238-3.



Hamilton, J. D. (1989). "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle". *Econometrica* 57.(2), pp. 357–384.



Box, G. E. P. & Tiao, G. C. (1975). "Intervention Analysis with Applications to Economic and Environmental Problems". *Journal of the American Statistical Association* 70.(349), pp. 70–79.



Winters, P. R. (1960). "Forecasting Sales by Exponentially Weighted Moving Averages". *Management Science* 6, pp. 324–342.



Holt, C. C. (1957). "Forecasting Trends and Seasonals by Exponentially Weighted Moving Averages". *ONR Research Memorandum, Carnegie Institute of Technology* 52.