

基于预测与评价模型的疫情防控研究

摘 要

自新冠疫情爆发以来,新型冠状病毒(COVID-19)对世界各国的人口和经济方面产生了严重的影响。因此,本文以中国、美国等八个具有代表性的国家为例,通过构建新型冠状病毒感染防控评价模型,对世界各国防疫水平做出评价。为了更深入地研究不同疫情防控水平对人口增长以及经济恢复情况的影响,本文选取防疫水平显著的中国和防疫水平低下的美国,利用灰色预测模型、ARIMA 模型对疫情爆发后的人口数据进行预测分析,再利用多元线性回归模型对疫情爆发后的经济数据进行预测分析,得出结论为新冠疫情防控对人口和经济方面都带来了显著的负面影响,但是中国的疫情防控对人口的负增长有一定的抑制作用并且有效缓解了疫情对国民经济造成的冲击。

关键词: 疫情防控 灰色预测 ARIMA 多元线性回归 预测评价

目录

摘要.....	I
表格与插图清单.....	V
一、 引言	1
(一) 研究背景	1
(二) 相关研究	3
二、 模型建立与检验	4
(一) 数据来源	4
(二) 指标选取	4
(三) 模型建立	6
1. 疫情防控水平综合评价模型	6
2. 新冠疫情下的人口预测模型	10
3. 新冠疫情下的经济预测模型	13
三、 模型求解与分析	13
(一) 疫情防控水平综合评价模型的求解	13
1. 最优权重求解	13
2. TOPSIS 综合评价法	14
3. 结果分析	15
(二) 新冠疫情下的人口预测模型的求解	15
1. 灰色预测模型	15
2. ARIMA 预测模型	16
(三) 新冠疫情下的经济预测模型的求解	17
1. 模型求解	17
2. 模型修正	18
3. 模型检验	20
4. 模型预测	21
四、 模型评价	22
(一) 模型的优点	22
(二) 模型的不足	22
五、 结论与建议	23
(一) 主要研究结论	23
(二) 相关建议	23
参考文献	25
附录	26

表格与插图清单

- 表 1. 反映八个国家防疫水平的评价指标表
- 表 2. 八个国家五个指标的最优权重
- 表 3. TOPSIS 评价法计算结果
- 表 4. 八个国家五个指标的综合水平得分
- 表 5. 中国灰色预测模型预测结果表
- 表 6. 美国灰色预测模型预测结果表
- 表 7. 中国 ARIMA 预测模型预测结果表
- 表 8. 美国 ARIMA 预测模型预测结果表
- 表 9. 中美 2020-2021 年年末人口总数预测值与真实值对比表
- 表 10. OLS 回归结果
- 表 11. 相关系数表
- 表 12. 对数变换回归结果
- 表 13. 剔除 X_4 后的回归结果
- 表 14. 中美 2020-2021 年国民生产总值预测值与真实值对比表
-
- 图 1. 八个国家五个指标相关性
- 图 2. 1992-2019 年中国人口数量变化图
- 图 3. 1992-2019 年美国人口数量变化图
- 图 4. 中国灰色预测模型拟合预测图
- 图 5. 美国灰色预测模型拟合预测图
- 图 6. 最优权重图
- 图 7. 八个国家得分图

基于预测与评价模型的疫情防控研究

一、引言

（一）研究背景

自 2019 年新型冠状病毒疫情爆发以来，全球在很长一段时间内都笼罩在病毒的阴霾下。面对病毒的肆虐，各个国家都采取了一定的疫情防控措施。相关防控措施不仅对疫情防控发挥了一定的作用，对各国人口数量和经济发展也有一定的影响。截至目前，中国的新型冠状病毒疫情得到了有效控制，但仍然有许多国家和地区的疫情持续反弹，不少人多次感染新冠，全球的人们都渴望疫情不断反复的情况彻底结束，各国政府也致力于恢复由于疫情带来的不可避免的人口和经济方面的冲击。因此，对各个国家的疫情防控水平及其带来的人口增长和经济发展方面的影响展开研究，具有十分重要的现实意义[1]。

面对全球突发公共卫生事件，每个国家会根据本国实际国情采取一系列预防、控制、救治、恢复等方面的措施。为遏制新冠病毒疫情的传播，各国普遍加强了对除医疗物资外的货物的检验检疫工作，部分国家还采取了暂停国际航班或国内公共交通等限制性措施，控制人员流动以阻止新冠病毒进一步蔓延。但同时，也有部分国家疫情防控措施较为松懈，甚至刻意淡化疫情警告、消极应对疫情防控工作，导致疫情大面积覆盖全国。

疫情防控期间，亚洲国家较早地对疫情进行了有效管控，以中国、越南、新加坡为代表的几个国家迅速采取了相关措施。其中中国实现了疫苗全覆盖，疫情期间采取了严格执行减少人员流动、物理隔离、免费治疗患者等防控防治措施，有效地控制了新冠疫情的扩散，及时克服了疫情引发的经济下滑的势头。横跨亚欧大陆的发展中国家土耳其后期解除了普通居民的出行限制，并取消户外戴口罩的要求，对疫情防控也较为松懈。由于欧洲一体化存在结构性缺陷，部分国家的体制性弊端以及民众对自由生活和隐私保护极为重视，欧洲的大多数国家在疫情初期较难迅速进行有效控制。其中发展中国家俄罗斯全面禁止人员流动，对疫情防控起到了一定的作用；发达国家如英国、法国、德国采取了一定程度上的积极政策进行防控。而地处北美洲的美国在疫情期间刻意淡化疫情警告，甚至发布虚假信息误导民众，美国联邦和地方政府在应对

疫情的措施上缺乏协调统一。正是由于各国对疫情防控的不同重视程度和不同防控防治措施，各国的疫情控制水平和经济复苏发展状态也不尽相同。

随着新冠感染病例的不断增加，全球的死亡人口数急剧增加，部分国家的出生人口也大幅下降。中国采取的控制病例流动和免费救治、接种疫苗等相关控制与救治措施，在极大程度上缓冲了死亡人数激增的严峻势态。反观美国初期对疫情的认识和应对存在一些问题，导致疫情蔓延威胁人民健康，出现了大批的死亡病例，人口总数骤降。新冠疫情导致很长一段时间内，世界经济体系发展形势极不稳定，经济增长速度减缓，甚至出现负增长。在疫情最严重的时期，中国是全球唯一实现经济正增长的国家，中国向其他国家和国际组织提供疫苗援助，帮助全球疫情得到医药物质保障，在一定程度上推动其他各国尽快恢复生产、实现经济复苏，中国经济的稳定为全球经济的发展做出了巨大贡献。

对各国疫情防控水平的研究以及疫情防控手段对人口与经济影响的研究为制定有效的防控措施提供参考，有助于预防和应对类似公共卫生事件，促进全球卫生安全，也保障再次遭遇类似事件打击后的各个国家和地区能够快速恢复人口、经济发展的正常状态。

为研究各国防疫水平，本文结合相关背景，根据地理位置、国家性质以及疫情期间采取的防控手段挑选了中国、土耳其、俄罗斯、美国、英国、法国、德国、西班牙八个有代表性的国家。参考世界卫生组织官网发布的 COVID-19 相关数据，选择累计确诊人数、累计死亡人数、治愈率、重症或危重症人数、重症或危重症人数率五个评估指标建立综合评价模型对各国疫情防控水平进行分析，先利用层次分析法和序关系分析法、熵值权重法和变异系数法，从主观和客观两方面对八国疫情防控水平进行权重的计算，再运用综合集成赋权法得到指标最优权重，利用 TOPSIS 综合评价法得到八个国家的疫情防控水平排序[2]。

为进一步研究不同疫情防控水平对人口增长及经济恢复状况的影响，本文结合疫情防控水平排序，对排序首位和末位的两个具有代表性的国家建立预测模型进行分析。首先基于疫情爆发前两国的人口数据，利用灰色预测模型、ARIMA 模型对疫情爆发后的两年两国人口总数进行预测得出预测值，与真实值对比分析，研究在不同防疫程度对人口增长的影响。再基于两国疫情爆发前相关经济数据，利用 Eviews 软件建立多元线性回归预测模型，将得到的预测值与真实数据进行比较分析，研究在不同防疫程

度对经济恢复的作用。

（二）相关研究

目前对新冠疫情防控水平的研究主要集中于从医学、公共卫生学、经济学等方面研究国内外防控手段及其影响因素入手。如李寻昌、刘怡帆、王雅倩、杨威、许锐基于主成分分析法对新冠防控的影响因素进行研究，以英国、匈牙利、塞尔维亚等十四个疫情国家为例，提取相关数据中对新冠病毒感染和防控影响效果较显著的因素，构建新型冠状病毒感染防控评估模型。通过评估结果认为严格的管理控制手段和较高的疫苗接种率对疫情防控非常有效。在疫情出现反弹时严格采用管控手段能有效降低新冠病毒的扩散，而散漫的控制管理举措对疫情防控效果产生负向的影响[3]。刘学通、马静、王兆南、刘梦冉、邓晋琦等通过问卷调查的形式对省、市、县三级疾病预防控制中心防控胜任水平进行研究，得到早期疫情防控胜任水平平均自评分为 (7.68 ± 1.473) 分，为疫情防控能力的提升以及疾控系统的更新提供借鉴[4]。戴冉等以徐州市 2019-2021 年的疫情防控效率相关数据为基础，基于熵权-TOPSIS 法构建卫生应急管理评价模型，为区域性卫生应急管理与我国卫生应急体系提出针对性的改进建议[5]。

针对新冠疫情下不同防控手段或水平对各国人口增长和经济发展状况的影响，现有研究大多集中于疫情对中国或全球经济运行的影响。如徐光顺、张寅东基于 2015-2020 年山东省区县级金融有关数据，运用双重差分法探究了新冠疫情对区县经济发展产生的重大影响，着重研究数字金融对疫情经济冲击的调节原理与效果，发现数字金融对山东省经济发展受到的疫情负面冲击起到了显著的缓冲作用，从金融角度帮助完善相关疫情防控措施，为疫情防控常态下的区县经济起到一定的指导作用[6]。胡烜瑜基于对疫情危机和金融危机的定性研究，以及对供求体系与经济周期规律的预测，研究新冠疫情下的金融市场震荡与经济冲击，并对未来面对类似风险时的政策提出相关可行性建议。其研究认为，合理并适当进行管控、活用政策工具才能合作度过疫情下的经济危机[7]。张雪莹基于 2020-2021 年全国各省的季度和月度面板数据，利用空间杜宾模型量化分析新冠疫情对三次产业、农业工业，房地产业经济的负面影响，提出疫情常态化管理下的经济高质量发展路径。并进一步构建空间计量模型，探究疫情防控常态化管理的现实经济意义[8]。裴凯欣从具体行业入手，通过财务分析

方式来研究新冠对各行业经济的发展影响。根据研究结果推测出目前及未来整体经济和国家产业结构发挥的重要作用，认为新冠疫情对我国经济影响在总体上是消极的，但疫情也为某些行业带来了发展契机。从疫情带来的经济发展契机方面，为新冠疫情下国家经济的发展提供来可行性建议[9]。

本次研究参考相关研究采取的分析指标，查找疫情、人口、经济相关数据，建立综合评价模型与预测模型，对比分析不同的疫情防控水平及其对人口和经济两方面的影响，得到疫情防控成果及防控手段对人口和经济影响的相关结论。希望研究成果能为疫情防控防治提供借鉴和参考，控制疫情的发展蔓延，为今后各国应对重大传染病提供思路，并希望对疫情常态化管理下的人口增长和经济运行起到一定的借鉴作用。

二、模型建立与检验

（一）数据来源

本文选取的疫情防控水平相关研究数据来自世界卫生数据平台和中国疾病预防控制中心，为了使评估体系的结果更具说服力，本文依据地理位置从五大洲中分别选取经济水平较为突出，再研究疫情防控上能提供较多数据的国家，包括中国、美国、英国、俄罗斯、法国、土耳其、德国、西班牙共 8 个国家进行分析评价。选取的人口相关的数据，即中国和美国 1992-2019 年年末人口总数，来源于中国国家统计局和世界银行官网。选取的经济相关数据中的中国 2002 年至 2021 年的“国内生产总值”、“居民消费价格指数”、“居民人均可支配收入”、“进出口总额”、“货运量”这五个变量的数据均来源于《中国统计年鉴 2022》；美国 2002 年至 2021 年的“国民生产总值”和“居民可支配收入”为世界银行国民经济核算数据；“居民消费价格指数”来源于国际货币基金组织《国际金融统计》和数据文件；“进出口总额”来源于美国商务部，“货运量”世界银行，运输、水、信息和通讯技术部、交通部。

（二）指标选取

为综合评价 8 个有代表性的国家的疫情防控水平，本文根据收集到的数据利用 Excel 公式得到治愈率、重症或危重症率，最终确定通过各国累计确诊人数、累计死亡人数、治愈率、重症或危重症人数、重症或危重症率五个指标来评价其疫情防控水平。评价指标如下表 1 所示：

表 1 反映八个国家防疫水平的评价指标表

评价指标	一级指标
八个国家疫情防控水平	累计确诊人数
	累计死亡人数
	治愈率
	重症或危重症人数
	重症或危重症率

对反映八个国家的防疫水平的五个评价指标进行相关性分析，结果表明各指标之间具有较强的相关性，结果如图 1 所示。说明指标选取用于解释新冠疫情评价体系的信息重合度较为适中，其解释能力较高，指标的选取较为合适。



图 1 八个国家五个指标相关性

为进一步研究疫情防控对人口的影响，更好地预测中国和美国在 2020-2021 年的人口数量，对于收集到的数据，本文提取中国和美国 1992-2019 年的人口总数，结合该数据，利用 SPSSPRO 画出相应的折线图如图 2、图 3 所示，从图中可以看出，中国和美国的人口数量呈现出以一定速率增长的趋势。

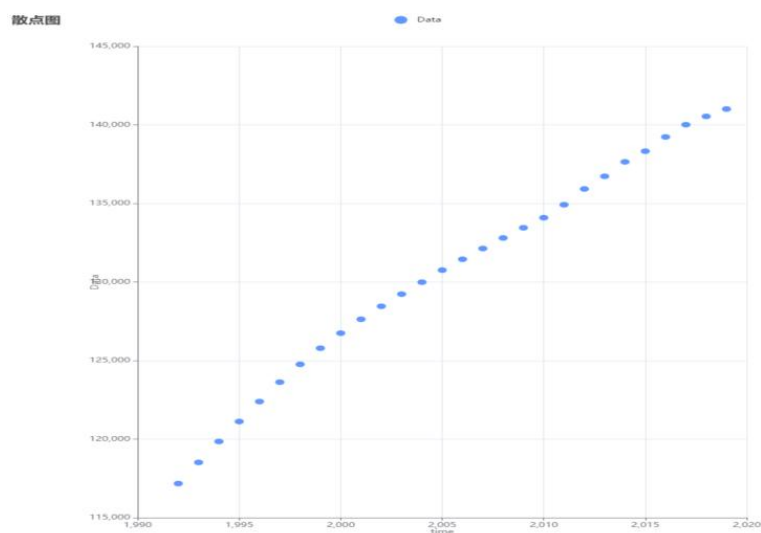


图 2 1992-2019 年中国人口数量变化图

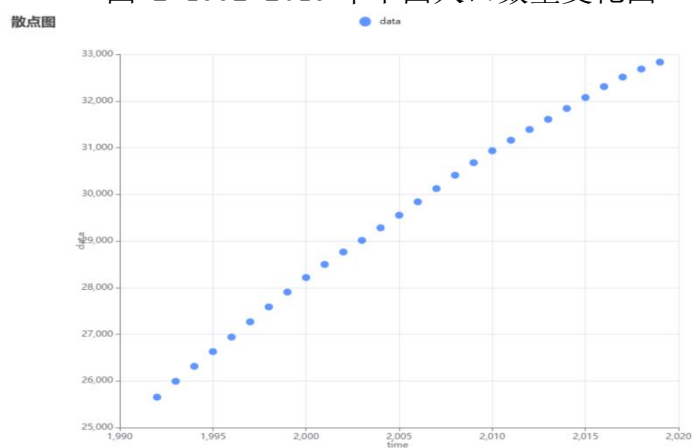


图 3 1992-2019 年美国人口数量变化图

本文选取 2002 年至 2019 年中美两国的国内生产总值、居民消费价格指数、居民人均可支配收入、进出口总额以及货运量这五个指标来建立多元线性回归模型。以中国为例，将“GDP”（单位：亿元）作为被解释变量（用 Y 表示），将“居民消费价格指数”（上年=100）、“居民人均可支配收入”（单位：元）、“进出口总额”（单位：亿元）、“货运量”（单位：百万吨-公里）作为解释变量（分别用 X1、X2、X3、X4 表示）。

（三）模型建立

1. 疫情防控水平综合评价模型

①主观赋权法

（1）层次分析法（AHP）

层次分析法（AHP）通过建立递阶层次结构，把人的主观判断转化成若干两两因素重要程度的判断上，从而把难以操作的定性判断量化为可操作的重要性程度判断上[10]。根据重要程度构造出的特殊矩阵得到各项指标的权重值,其步骤描述如下:

① 构建判断矩阵

将选取的二级指标两两比较，判断其相互的重要程度通过 1-9 维度法构造判断矩阵，重要程度的判断情况为：

若标度为奇数，当标度为 1 时认为两因素重要性相当。随着标度数字的增加，认为一个因素相较于另一个因素的重要程度越来越强烈，标度为 9 时极端重要。若标度为偶数，则取相邻两奇数标度判断的中值。

② 一致性检验

由判断矩阵计算被比较元素对于该准则的相对权重，并进行一致性检验，检验通过权重才能使用。一致性比例 $CR < 0.1$ ，则判断矩阵的一致性可以接受；否则需要对判断矩阵进行修正。其计算步骤如下：

Step1 带入已知数据计算一致性指标 CI，其公式如下：

$$CI = \frac{\lambda_{\max} - N}{N - 1} \quad \text{公式 (1)}$$

$CI > 0$ ，CI 越小一致性越大；

Step2 查找对应的平均随机一致性指标 RI；

Step3 计算一致性比例 CR，其公式如下：

$$CR = \frac{CI}{RI} \quad \text{公式 (2)}$$

(2) 序关系分析法 (GI)

序关系法在层次分析法的基础上进行改进，不需构造判断矩阵，减少了在确定个指标权重时的计算量，且无需检查一致性，可操作性增强。计算步骤如下所示：

Step1 确定指标之间的序关系：结合专家打分，根据 4 个一级评价指标的重要程度确定排序关系，若指标 X_i 相对于 X_j 重要，则记为 $X_i > X_j$ ；

Step2 对各项指标重要程度之比 $\frac{X_{k-1}}{X_k}$ 赋值为：

$$r_k = \frac{\omega_{k-1}}{\omega_k}, (k = n, n-1, \dots, 2) \quad \text{公式 (3)}$$

Step3 假设专家给出 r_k 理性赋值，计算指标权重 ω_k ；

最终利用 Matlab 序关系分析法相关程序求得本问各指标权重。

②客观赋权法

(1) 熵值权重法

熵值权重法是一种客观确定各个指标权重的方法，利用熵值权重法可以估算各个指标的权重，其本质是利用指标信息的差异来计算其权重以及熵值，其差异系数越高，对评价的重要性就越大或称权重越大，对评价的结果的贡献也就最大。利用熵值取权法，通过现有数据进行相关权重计算，具体操作如下。

Step1 构造原始指标数据矩阵：根据所选指标数据 A_1 GDP、 A_2 消费者价格指数（CPI）、 A_3 全国居民平均可支配收入、 A_4 进出口交易总额、 A_5 货运量，构造原始数据矩阵。用 A_{ij} 表示第 i 个样本关于指标变量 A_j 的取值，构造数据矩阵 $A=(A_{ij}) i \times 5$ 如下：

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{15} \\ A_{21} & A_{22} & \dots & A_{25} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & A_{i2} & \dots & A_{i5} \end{pmatrix}$$

Step2 对数据矩阵的各项值进行归一化处理，公式如下：

$$A'_{ij} = \frac{A_{ij} - A_{\min_i}}{A_{\max_i} - A_{\min_i}} \quad \text{公式 (4)}$$

Step3 根据归一化指标数据，计算第 j 项指标下第 i 个样本占该指标的比重 P_{ij} ，公式如下：

$$P_{ij} = \frac{A_{ij}}{\sum_{i=1}^{10} A_{ij}} \quad \text{公式 (5)}$$

Step4 计算第 j 项指标的熵值 E_j ，公式如下：

$$E_j = -\frac{1}{\ln n} \sum_{i=1}^n P_{ij} * \ln P_{ij} \quad \text{公式 (6)}$$

Step5 计算各指标的权重 W_j ，公式如下：

$$W_j = \frac{1-E_j}{\sum_{j=1}^5 1-E_j} \quad \text{公式 (7)}$$

(2) 变异系数法

变异系数法根据指标数据求权重指标，反应指标数据变化的客观信息，是一种客观的求权重的方法。根据每个指标带来的信息量大小确定其权重[11]。其计算步骤如下：

Step1 构造原始指标数据矩阵 $B = (x_{ij})$

Step2 计算指标的均值和标准差，其公式如下：

$$\begin{cases} \bar{x}_j = \frac{1}{n} \sum_{i=1}^n x_{ij} \\ S_j = \sqrt{\frac{\sum_{i=1}^n (x_{ij} - \bar{x}_j)^2}{n-1}} \end{cases} \quad \text{公式 (8)}$$

Step3 计算第 j 项指标的变异系数：

$$v_j = \frac{S_j}{\bar{x}_j}, (j = 1, 2, \dots, n) \quad \text{公式 (9)}$$

Step4 计算确定各权重：

$$\omega_j = \frac{v_j}{\sum_{j=1}^n v_j}, (j = 1, 2, \dots, n) \quad \text{公式 (10)}$$

③综合集成赋权法

设 w^j 和 w^i 分别为主观赋权法和客观赋权法对第 j 项指标所确定的权重系数，将 w^j 和 w^i 进行加权求和作为第 j 项指标和新的权重系数，代入数值可得即 $w^j = k_1 w_j + k_2 w_i (j=1, 2, 3, \dots)$ ，其中 k_1 和 k_2 为待定常数，且满足 $k_1 + k_2 = 1$ ， k_1, k_2 分别表示主观权重与客观权重的相对重要程度[2]。

1.4 TOPSIS 综合评价法

TOPSIS 综合评价法通过对有限个评价对象与理想化目标的距离来进行检测从而对其进行排序。将各个备选方案与正理想解和负理想解作比较，当其中有一个方案最

接近理想解，同时又远离负理想解时，这个方案即为备选方案中到目前为止最好的方案[7]。TOPOSI 综合评价法算法描述下：

Step1 构造决策矩阵 $C = (c_{ij})$ ，i 为各个国家，j 为各个指标

Step2 对决策矩阵进行规范化处理

Step3 构建权重 ω_j ，此权重由上述最优权重确定

Step4 计算加权规范矩阵 $M = (m_{ij})_{5 \times 8}$

Step5 确定正理想解 M^* 和负理想解 M^0 ，设正理想解 M^* 的第 j 个属性值为 m_j^* ，负理想解 M^0 的第 j 个属性值为 m_j^0

Step6 计算个方案到正负理想解的距离。

方案到正理想解的距离为：

$$S_i^* = \sqrt{\sum_{j=1}^8 (m_{ij} - m_j^*)^2} \quad \text{公式 (11)}$$

方案到负理想解的距离为：

$$S_i^0 = \sqrt{\sum_{j=1}^8 (m_{ij} - m_j^0)^2} \quad \text{公式 (12)}$$

Step7 计算各方案排序指标值：

$$f_i^* = \frac{S_i^0}{S_i^0 + S_i^*} \quad \text{公式 (13)}$$

2. 新冠疫情下的人口预测模型

① 灰色预测模型

灰色预测模型是利用所产生的一系列数据，通过累积产生的原始数据，获得一个近似的指数法则，然后再对其进行建模。灰色预测法是一种能够精确地预测未来的相关性的预测方法，本文利用规律再进行建模的方法，将分散在时间轴上的离散数据视为一组不断变化的序列，利用累加和累减的方法，弱化灰色系统中的未知因素，加强已知因素的影响，最终构造出一个以时间为变量的连续微分方程。通过数学方法确定

方程中的参数，从而实现预测中国 2026-2030 年人口数量的目的。而灰色预测模型的一般表达方式为 GM (n, x) 模型，其含义是：用 n 阶微分方程对 x 个变量建立模型。模型具体算法步骤如下：

Step1 对原始数据 $x^{(0)}$ 作首次累加得 $x^{(1)}$

Step2: 构造数据矩阵 B 及数据向量 Y

Step3: 计算：

$$\hat{u} = \begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (B^T B)^{-1} B^T Y \quad \text{公式 (14)}$$

得 \hat{a} , \hat{b} ；建立模型

$$\frac{d x^{(1)}}{dt} + \hat{a} x^{(1)} = \hat{b} \quad \text{公式 (15)}$$

解得 Y

Step4:

$$\hat{x}^{(1)}(k+1) = \left(x^{(0)}(1) - \frac{\hat{b}}{\hat{a}} \right) e^{-\hat{a}k} + \frac{\hat{b}}{\hat{a}} \quad \text{公式 (16)}$$

Step5: 求生成序列预测值 $\hat{x}^{(1)}(k+1)$ 及模型还原值 $\hat{x}^{(0)}(k+1)$

利用 SPSSPRO 软件导入中国和美国 1992-2019 年二十七年人口数量可得到灰色预测模型拟合结果表，详细表格见支撑材料。从表分析可以得到，中国人口数据原序列的所有级比值都位于区间 (0.938, 1.067))，美国原序列的所有级比值都位于区间 (0.933, 1.071)) 内，说明原序列适合构建灰色预测模型。，通过级比检验说明原序列适合构建灰色预测模型。对于中国数据后验差比值为 0.014<0.35，该模型精度高。对于美国数据后验差比值为 0.009< 0.35，模型精度也很高。中国和美国数据的灰色预测模型平均相对误差为 0.496% ， 0.573%，意味着模型拟合效果良好。故可用该模型进行中国和美国 2020-2021 年人口数量的预测。

本文根据系统预测拟合检测方法进行运算，可以得出对于中国和美国在 2020-2021 年的人口数量不断增加，通过以上数据可知，随着年份的推移，中国和美国的人口数量不断增加，通过分析可得，原始数据与预测的人口数量数据之间有必然的联系，为了检测预测的准确性与精确值，本文对预测的实际值进行拟合，得出以下结果，结果如下图所示：

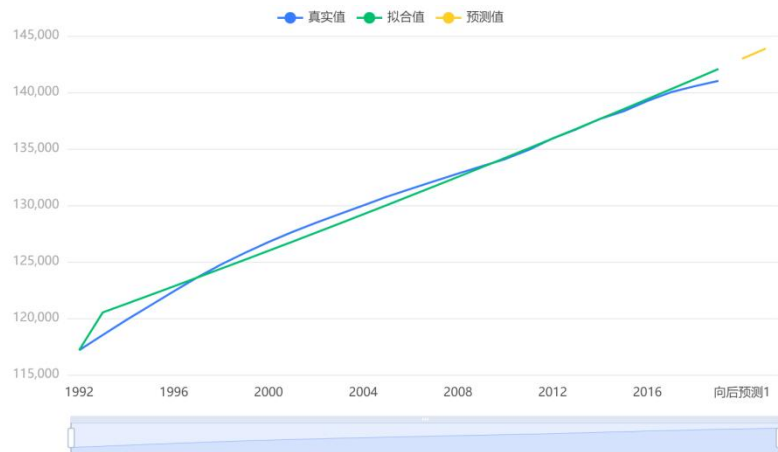


图 4 中国灰色预测模型拟合预测图

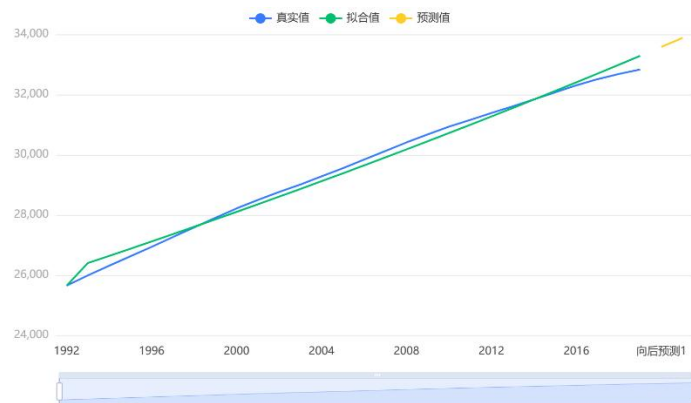


图 5 美国灰色预测模型拟合预测图

②ARIMA 预测模型

1. ARIMA 模型需要序列满足平稳性，通过对 ADF 检验的结果进行分析，从 t 值的角度来看，ARIMA 模型能否有效地排除序列非平稳性的假设两组比较差异有统计学意义 ($P < 0.05$)。

2. 查看差分前和之后的数据对比，判断其是否稳定（波动的范围很小），并对时间序列做一些（偏）自相关分析，根据截尾情况估算其 p、q 值。

3. ARIMA 模型要求模型具备纯随机性，即模型残差为白噪声，查看模型检验表，根据 Q 统计量的 P 值 ($P > 0.05$) 对模型白噪声进行检验，也可以结合信息准则 AIC 和 BIC 值进行分析（越低越好），也可以通过模型残差 ACF/PACF 图进行分析根据模型参数表，得出模型公式结合时间序列分析图进行综合分析，得到向后预测的阶数结果。

Step1: AR(p) 模型的定阶原理

AR 模型是一个线性模型，p 阶自回归模型的一般表达式为：

$$x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \dots + \phi_p x_{t-p} + \varepsilon_t \quad \text{公式 (17)}$$

AR(p)的平稳序列模型:

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-2} + \dots + \phi_k \rho_{k-p} (k \geq p) \quad \text{公式 (18)}$$

因此符合 AR(p) 的平稳序列模型，其自相关系数在 p 阶之后一直不会为 0，存在拖尾现象。

Step2: MA(q) 模型被称为移动平均模型，一个 q 阶的移动平均模型可以用数学式表达为

$$X_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q} \quad \text{公式 (19)}$$

MA(q) 模型

$$\rho_1 = \frac{\theta_1 + \theta_1 \theta_2 + \theta_2 \theta_3 + \dots + \theta_{q-1} \theta_q}{1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2} \quad \text{公式 (20)}$$

MA(q) 模型的自相关系数 q 阶截尾

Step3: 将 AR(P) 模型和 MA (q) 相结合得出 ARIMA (p, q) 模型

$$x_t = \phi_0 + \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_q \varepsilon_{t-q} \quad \text{公式 (21)}$$

3. 新冠疫情下的经济预测模型

为了研究中国和美国疫情防控手段对国民经济的影响，本文通过建立多元线性回归模型，由 2022-2019 年的数据对之后的各变量数据进行预测得出 2020-2021 年的各变量数据，将 2020-2021 年的各变量预测值与实际值进行比较，从而得出政府在新冠疫情期间实施的手段对我国经济的影响。

多元线性回归模型如下：

设随机变量 Y 与一般变量 X_1, X_2, \dots, X_k 的理论线性回归模型为：

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_k X_{ki} + \mu_i \quad (i=1, 2, \dots, n) \quad \text{公式 (22)}$$

其中， $\beta_0, \beta_1, \dots, \beta_k$ 是 $k+1$ 个未知参数， β_0 称为回归常数， $\beta_0, \beta_1, \dots, \beta_k$ 称为偏归系数。Y 称为被解释变量（因变量），而 X_1, X_2, \dots, X_k 是 k 个可以精确测量并可控制的一般变量，称为解释变量（自变量）， μ_i 为随机误差项。

三、模型求解与分析

（一）疫情防控水平综合评价模型的求解

1. 最优权重求解

运用综合集成赋权法法，结合层次分析法、序关系分析法、熵值权重法、变异系

数法求得最优权重，最终得到五个指标的最优权重如表 2 所示：

表 2 八个国家五个指标的最优权重

指标	主观赋值法		客观赋值法		最优权重
	层次分析法	序关系分析法	熵值取值法	变异系数法	
累计确诊人数	0.224	0.2020	0.236	0.262	0.23
累计死亡人数	0.2327	0.2424	0.262	0.268	0.25
治愈率	0.1199	0.1203	0.117	0.018	0.09
重症或危重症人数	0.3345	0.2909	0.315	0.308	0.31
重症或危重症率	0.089	0.1443	0.07	0.144	0.12

最优权重图如下所示：

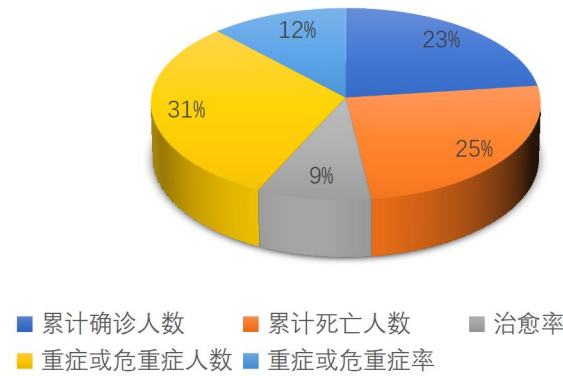


图 6 最优权重图

2. TOPSIS 综合评价法

根据八国防疫水平相关数据，建立 TOPSIS 评价模型，得出在熵权法下八国防疫水平的综合得分排序如表 3 所示，仅在客观评价法的基础上进行综合评价法有失一般性，下面在问题求解中将综合客观评价法和主观评价法给出最优权重得出 TOPSIS 综合评价法最终结果。

表 3 TOPSIS 评价法计算结果

国家	综合得分指数	排序
土耳其	0.60162511	1
中国	0.54260987	2
俄罗斯	0.5379255	3
德国	0.49847629	4
英国	0.49787896	5
法国	0.47864151	6
西班牙	0.46499872	7
美国	0.45253514	8

3. 结果分析

将影响各国家疫情防控水平的五个评价指标数据的原始数据及最优权重代入 TOPSIS 综合评价模型，即可得出八个国家疫情防控水平的综合水平得分，其分数和排序如下表 4 及图 7 所示。

表 4 八个国家五个指标的综合水平得分

国家	综合得分指数	排序
中国	0.94656095	1
土耳其	0.89624585	2
俄罗斯	0.75921246	3
西班牙	0.75774806	4
英国	0.75642076	5
法国	0.7353445	6
德国	0.69924965	7
美国	0.13349086	8

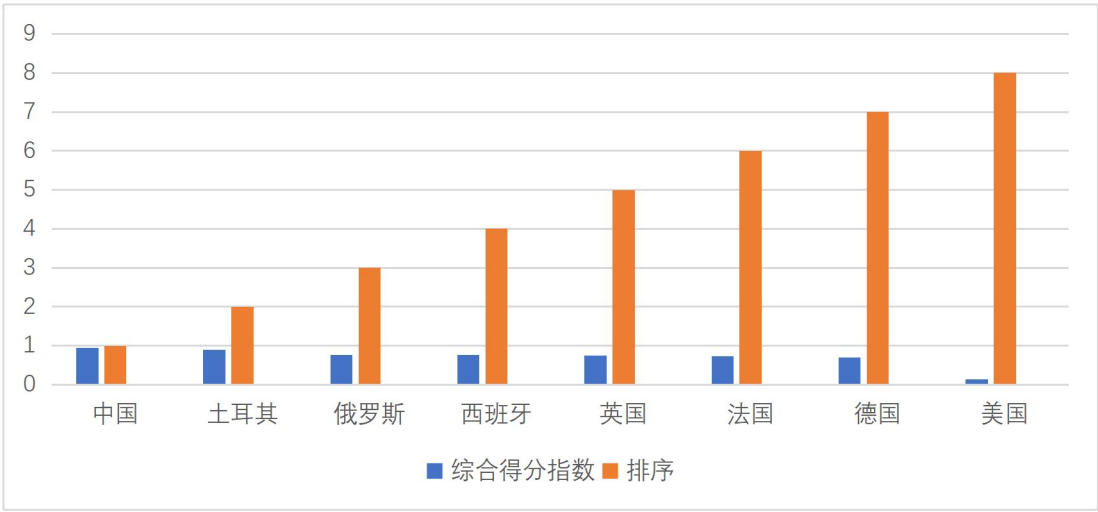


图 7 八个国家得分图

从结果可知：中国、土耳其的疫情防控水平得分最高，疫情防控效果显著；俄罗斯、西班牙、英国、法国，疫情防控水平居中；德国、美国疫情防控水平低下。

（二）新冠疫情下的人口预测模型的求解

1. 灰色预测模型

应用 SPSSPRO 的 GM (1, 1) 模型，对 2020-2021 年中国和美国的人口数量进行预测，最终得到如表 5、表 6 所示预测结果。

表 5 中国灰色预测模型预测结果表

年份	中国人口数量（万）
2020	142963.405
2021	143870.737

表 6 美国灰色预测模型预测结果表

年份	美国人口数（万）
2020	33582.915
2021	33883.425

2. ARIMA 预测模型

应用 Rstudio 程序计算 2020-2021 年的预测值如下表：

表 7 中国 ARIMA 预测模型预测结果表

年份	Forecast
2020	141475
2021	141942

表 8 美国 ARIMA 预测模型预测结果表

年份	Forecast
2020	32977.44
2021	33132.34

表 9 中美 2020-2021 年年末人口总数预测值与真实值对比表

年份	预测中国人口数（万）	真实中国人口数（万）	预测美国人口数（万）	真实美国人口数（万）
2020	142220	141212	33280	33150.1
2021	142910	141260	33508	33189.4

虽然疫情给中国和美国的人口数带来了消极影响，预测人口数远大于真实人口数，但根据两国不同的防疫措施和手段对比得出，中国的防疫措施对人口的负增长有一定的抑制作用，与美国的年末人口数差额负增长抑制效果对比可以看出中国采取及时有效的防疫措施可以预防在疫情环境下的人口负增长。

（三）新冠疫情下的经济预测模型的求解

1. 模型求解

以中国为例，假定所建模型及其中的随机扰动项 ε 满足各项古典假定，可以用 OLS 法估计其参数，得出回归结果如下 11 所示：

表10 OLS回归结果

Dependent Variable: Y				
Method: Least Squares				
Date: 05/25/23 Time: 19:26				
Sample: 2002 2019				
Included observations: 18				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
	-200387.			
C	9	111200.4	-1.802043	0.0948
	1503.37			
X1	2	1111.293	1.352814	0.1992
	28.9872			
X2	2	0.882457	32.84830	0.0000
	0.53424			
X3	5	0.204697	2.609934	0.0216
	-0.00637			
X4	9	0.011380	-0.560536	0.5846
	0.99946			476953
R-squared	9	Mean dependent var		.9
	0.99930			278020
Adjusted R-squared	6	S.D. dependent var		.0
	7324.10			20.865
S.E. of regression	8	Akaike info criterion		86
	6.97E+0			21.113
Sum squared resid	8	Schwarz criterion		19
	-182.792			20.899
Log likelihood	8	Hannan-Quinn criter.		97
	6120.69			1.4279
F-statistic	8	Durbin-Watson stat		92
	0.00000			
Prob(F-statistic)	0			

由表可知：货运量 X4 的偏回归系数 $\hat{\beta}_4$ 为负值，货运量每增加 1 百万吨公里，国民生产总值平均减少 0.006379 亿元，不符合定性分析，说明模型可能存在多重共线

性问题。同时，模型的可决系数 R^2 为 0.999469，修正的可决系数 \bar{R}^2 为 0.99306，极接近于 1，说明模型拟合程度较好。在 0.05 的显著性水平下，回归模型的 F 值为 $6120.698 > F_{0.05}(4, 15) = 3.06$ ，说明回归模型具有良好的显著性，但是居民消费价格指数 X1 的 P 值为 $0.0948 > 0.05$ ，货运量 X4 的 P 值为 $0.5846 > 0.05$ ，说明 X1 和 X4 对 Y 无显著性影响，可以初步判定该回归模型存在多重共线性。

为进一步确认模型中是否存在多重共线性问题以免对模型结果造成影响，本文采用简单系数法和利用 X2 变量进行辅助回归求方差扩大因子进行确认。相关系数表如表 11 所示，辅助回归结果见附件。

表 11 相关系数表

	X1	X2	X3	X4
X1	1.000000	0.011509	0.117203	0.082243
X2	0.011509	1.000000	0.966113	0.962245
X3	0.117203	0.966113	1.000000	0.990973
X4	0.082243	0.962245	0.990973	1.000000

由辅助回归结果表可知：X2 和 X4 的相关系数为 $0.962245 > 0.8$ ，X3 与 X4 的相关系数为 $0.990973 > 0.8$ ，说明模型中存在着严重的多重共线性。 $R_j^2=0.943975$ ，根据公式：

$$VIF_j = \frac{1}{1-R_j^2} \tag{23}$$

可以求得 $VIF_j=26.4866 >> 10$ ，说明模型中确实存在严重的多重共线性。

2. 模型修正

为了避免多重共线性对模型造成严重影响，本文采用对数变换的方式对模型进行修正，得到回归结果如下，如表 12 所示。

表 12 对数变换回归结果

Dependent Variable: LNY				
Method: Least Squares				
Date: 05/25/23 Time: 19:34				
Sample: 2002 2019				
Included observations: 18				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.16738	1.267906	-2.498123	0.0267

	4			
	0.72281			
LN1	3	0.266290	2.714382	0.0177
	0.80055			
LN2	5	0.045762	17.49386	0.0000
	0.36455			
LN3	7	0.079486	4.586413	0.0005
	0.04799			
LN4	3	0.093173	0.515091	0.6151
	0.99948			12.881
R-squared	9	Mean dependent var		97
	0.99933			0.6766
Adjusted R-squared	1	S.D. dependent var		78
	0.01749			-5.0235
S.E. of regression	6	Akaike info criterion		02
	0.00398			-4.7761
Sum squared resid	0	Schwarz criterion		76
	50.2115			-4.9893
Log likelihood	2	Hannan-Quinn criter.		99
	6353.74			1.6975
F-statistic	5	Durbin-Watson stat		31
	0.00000			
Prob(F-statistic)	0			

由表 12 可知：货运量 X4 的偏回归系数符号 $\hat{\beta}_4$ 已经修正，多重共线性问题得到改善，但是 X4 的 P 值为 0.6151>0.05，说明在 0.05 的显著性水平下，X4 对 Y 无显著影响。因此本文剔除 X4 变量，再次进行回归分析，得出结果如表 13 所示。可用规范的形式将回归方程写为：

$$\hat{Y}_t = -2.874398 + 0.714121X_1 + 0.813148X_2 + 0.392924X_3$$

表 13 剔除 X4 后的回归结果

Dependent Variable: LNY				
Method: Least Squares				
Date: 05/25/23 Time: 19:49				
Sample: 2002 2019				
Included observations: 18				
Variable	Coefficient	Std. Error	t-Statistic	Prob.

nt				
	-2.87439			
C	8	1.103024	-2.605925	0.0207
	0.71412			
LNx1	1	0.258688	2.760551	0.0153
	0.81314			
LNx2	8	0.037655	21.59450	0.0000
	0.39292			
LNx3	4	0.055794	7.042404	0.0000
	0.99947			12.881
R-squared	8	Mean dependent var		97
	0.99936			0.6766
Adjusted R-squared	7	S.D. dependent var		78
	0.01703			-5.1144
S.E. of regression	1	Akaike info criterion		09
	0.00406			-4.9165
Sum squared resid	1	Schwarz criterion		49
	50.0296			-5.0871
Log likelihood	9	Hannan-Quinn criter.		27
	8940.75			1.8639
F-statistic	8	Durbin-Watson stat		80
	0.00000			
Prob(F-statistic)	0			

按照上述方法，同理可得出美国 2002 年至 2019 年国民生产总值及其影响因素的回归方程为：

$$\hat{Y}_t = 3.777929 + 0.332344X_1 + 1.039340X_2$$

3. 模型检验

(1) 经济意义检验

由回归结果可知：所估计的参数 $\hat{\beta}_0 = -2.874398$ ， $\hat{\beta}_1 = 0.714121$ ， $\hat{\beta}_2 = 0.813148$ ， $\hat{\beta}_3 = 0.392924$ ，说明居民消费价格指数每增加 1%，国民生产总值将平均增加 0.714121%；居民人均可支配收入每增加 1%，国民生产总值将平均增加 0.813148%；进出口总额每增加 1%，国民生产总值将平均增加 0.392924%，与预期的经济意义相符。

(2) 拟合优度和统计检验

拟合优度的度量：由表 13 可知，模型的可决系数 R^2 为 0.999478，修正的可决系

数 \bar{R}^2 为 0.999367, 极接近于 1, 说明所建模型整体上对样本数据拟合较好, 即解释变量“居民消费价格指数”、“居民人均可支配收入”、“进出口总额”对被解释变量“国民生产总值”的绝大部分差异作出了解释。

对回归系数的 t 检验: 针对 $H_0: \beta_j = 0 (j = 1, 2, 3)$ 和 $H_1: \beta_j \neq 0 (j = 1, 2, 3)$, 由图 5 还可以看出, 估计的回归系数 $\hat{\beta}_1$ 的标准误差和 t 值分别为: $SE(\hat{\beta}_1) = 0.258688$, $t(\hat{\beta}_1) = 2.760551$; $\hat{\beta}_2$ 的标准误差和 t 值分别为: $SE(\hat{\beta}_2) = 21.59450$, $t(\hat{\beta}_2) = 21.59450$ 。取 $\alpha = 0.05$, 查 t 分布表得自由度为 $n-2=18-2=16$ 的临界值 $t_{0.025}(16) = 2.120$ 。因为 $|t(\hat{\beta}_1)| = 2.760551 > t_{0.025}(16) = 2.120$, 应拒绝原假设 $H_0: \beta_1 = 0 (j = 1, 2, 3, 4)$; $|t(\hat{\beta}_2)| = 21.59450 > t_{0.025}(16) = 2.120$, 应拒绝原假设 $H_0: \beta_2 = 0 (j = 1, 2, 3, 4)$; $|t(\hat{\beta}_3)| = 7.042404 > t_{0.025}(16) = 2.120$, 应拒绝原假设 $H_0: \beta_3 = 0 (j = 1, 2, 3, 4)$ 。对回归系数的显著性检验表明, “居民消费价格指数”、“居民人均可支配收入”、“进出口总额”对“国民生产总值”有显著影响。

4. 模型预测

如果 2020 年中国的居民消费价格指数为 102.5, 居民人均可支配收入为 32189 元, 进出口总额为 322215.24 亿元, 利用所估计的模型可预测中国在 2020 年国民生产总值可能到达的水平, 点预测值的计算方法为:

$$\hat{Y}_f = -2.874398 + 0.714121 \times 102.5 + 0.813148 \times 32189 + 0.392924 \times 322215.24 = 152850.8449 \text{ (亿万元)}$$

2020 年国民生产总值的真实值 1005451.3 亿万元与该预测值相比, 增加约 852600.4551 亿万元。

同理, 如果 2021 年中国的居民消费价格指数为 100.9, 居民人均可支配收入为 35128 元, 进出口总额为 391009 亿元, 可预测中国在 2021 年国民生产总值为:

$$\hat{Y}_f = -2.874398 + 0.714121 \times 100.9 + 0.813148 \times 35128 + 0.392924 \times 391009 = 182274.1622 \text{ (亿万元)}$$

2021 年国民生产总值的真实值 1133239.8 亿万元与该预测值相比, 增加约 950965.6378 亿万元。

按照上述方法, 同理可得出美国 2020 年与 2021 年国民生产总值的预测值分别为 65602.7157 亿万美元和 71980.9277 亿万美元。与 2020 和 2021 年国民生产总值的真实值相比, 分别增加 209948.7089 亿万美元和 232430.9963 亿万美元。

表 14 中美 2020-2021 年国民生产总值预测值与真实值对比表

年份	预测中国国民生产总值（亿万元）	真实中国国民生产总值（亿万元）	预测美国国民生产总值（亿万美元）	真实美国国民生产总值（亿万美元）
2020	152850.8449	1005451.3	65602.7157	21060473.61
2021	182274.1622	1133239.8	71980.9277	23315080.56

5. 结果分析

结果表明：美国虽然国民经济有所增长，但是由于疫情期间消极应对疫情防控工作，所以国民经济涨幅较低。而中国政府于美国政府相比，由于防疫效果显著，防控手段实施前后，国民经济涨幅较大，有效缓解了疫情对国民经济造成的冲击。

四、模型评价

（一）模型的优点

（1）采用 TOPSIS-RSR 模型能巧妙的结合已知标准，能同时对于多个对象进行评价。

（2）运用变异系数以及层次分析法等四个模型进行综合求权，再使用 TOPSIS 综合分析法求得 8 个国家在确诊人数、痊愈人数、死亡人数等方面的得分，更具有客观性与综合性，增强了结果的说服力。

（3）对各指标进行了相关性分析来更准确地描述变量之间的相关程度，使结果更具客观性。

（4）运用灰色序列预测模型，相比其他需要较大样本的预测方法来说，该方法所需建模信息少，运算方便，建模精度高。利用灰色预测法和 ARIMA 模型求预测值，使得预测值更贴合实际，误差较小。

（5）采用多元线性回归模型，用多个变量的最优组合来预测和估计因变量，是使结果符合实际情况，具有实用意义。

（二）模型的不足

（1）采用的 TOPSIS 法只能反映各评价指标之间内部的相对接近度，不能反映与期望的最佳方案的相对接近度。

(2) 评价世界各国的疫情防控水平时，评估指标应尽量全面，还应该考虑疫苗接种率、人流量等因素。

(3) 在做数据特征分析时，由于查找数据的主观性，可能会存在人为的、非客观的剔除有效的数据，即需预测 2020-2021 年的人口数量数据，使得预测值可能会跟实际数据出现一定的偏差。

五、结论与建议

(一) 主要研究结论

本文通过使用截至 2021 年 12 月 31 日八国防疫水平相关数据及 2002 年至 2019 年中美两国的经济指标数据，运用评价和预测模型探究疫情防控对人口和经济方面的影响，得出如下结论：

(1) 在疫情期间，实现疫苗全覆盖，严格执行减少人员流动、采取物理隔离、免费治疗患者的中国，疫情防控水平综合得分最高，疫情防控效果显著；疫情防控措施较为的全面土耳其和俄罗斯、西班牙，疫情防控水平居中；欧洲发达国家德国、法国、英国采取了一定作用的防疫措施，疫情防控水平偏下，而综合得分最低，刻意淡化疫情警告、消极应对疫情防控工作的美国疫情防控水平和防控意识最低。

(2 美国作为世界上最发达的国家，人口不到中国的 1/4，但是对比中国和美国 2020-2021 年末人口总数的真实值可以发现美国预测人口数与真实人口数差额比更大。面对公共卫生危机，美国并未做出公众所需要的、公共卫生应有的反应。在早期美国应对新冠疫情的关键时刻，基于其对选举政治、政党政治的考量，美国未采取有效防控措施导致疫情给人口带来了不可磨灭的消极影响。

(3) 新冠疫情爆发对我国经济造成了严重的负面影响，中国先是处于疫情控制的五阶段，后又处于处于重启经济的五阶段。中国采用典型的“两手抓”的中国方案，“倒 V 形”疫情防控措施，实现了人力资本、经济损失最小化及经济复苏最大化，有效缓解了疫情对经济造成的冲击。

(二) 相关建议

在防疫常态化的背景下，各国应继续严格执行防疫战略，在本文论证的基础之上，提出如下建议：

(1) 弥补公共卫生空白，升级公共卫生基础设施，推进公共卫生理念，建立强有力的公共卫生治理结构，坚决保障世界各国人民的福祉。

(2) 采取积极的财政政策,更加灵活适度稳健的货币政策。

(3) 主动开展国际合作,打造人类卫生健康命运共同体理念和行动。

参考文献

- [1]李国平, 杨艺, 徐祯, 等. 全球重大疫情对经济系统的影响及其响应的比较研究[J]. 世界地理研究, 2020, 29(6):11.
- [2]黄敏;周德红;樊旭岩;陈锦辉;李晓柯;李嘉琦. 基于两种组合赋权法的天然气长输管道安全风险评估比较[J]. 工业安全与环保, 2023, (02):42-47.
- [3]李寻昌;刘怡帆;王雅倩;杨威;许锐. 基于主成分分析的新型冠状病毒感染防控影响因素研究[J]. 甘肃科学学报, 2023, 35(01):49-56+63.
- [4]刘学通;马静;王兆南;刘梦冉;邓晋琦;裴迎新;崔少颖;戴政. 省、市、县三级疾病预防控制中心新冠肺炎早期疫情防控胜任水平研究[J]. 卫生软科学, 2022, 36(12):73-78.
- [5]戴冉;温宗良;徐驰浩. 基于熵权-TOPSIS 的疫情防控效率评价——以徐州市 2019-2021 年为例[J]. 数据, 2023, (03):33-34.
- [6]徐光顺;张寅冬. 数字金融对新冠疫情经济冲击的调节效应研究——来自山东区县层面的经验证据[J]. 现代商业, 2022, (27):97-102.
- [7]胡烜瑜. 新冠疫情下的金融市场震荡与经济冲击[J]. 中国商论, 2020, (15):27-31+44.
- [8]张雪莹. 新冠肺炎疫情对产业经济的影响研究[C]. 青岛理工大学, 2022.
- [9]裴凯欣;杨晓丹. 新冠疫情对我国经济影响的研究[J]. 产业与科技论坛, 2021, (18):52-54.
- [10]宋婷. 层次分析法视角下美国应对新冠疫情失败的原因探析[C]. 外交学院, 2021.
- [11]陶磊, 肖广哲, 宋林. 基于变异系数法与 TOPSIS 的充填管道失效风险评价[J]. 2019. 05. 006.

附录

符号	说明
P_{ij}	第 j 项指标下第 i 个样本占该指标的比重
E_j	第 j 项指标的熵值第 j 项指标的熵值
W_j	各指标的权重
CI	一致性指标
λ_{\max}	判断矩阵的最大特征值
CR	一致性比例
$\overline{x_j}$	平均值
S_j	标准差
f_i^*	各方案排序指标值