### How Many Nuclear Bombs can Destroy the Earth?

鉴于核武器是属于毁灭性的武器，如何正确认识、合理预测核武器的发展趋势并探讨相关问题就显得尤为重要。本文采用层次分析法、时间序列预测模型、神经网络模型及聚类分析法对往年数据进行分析并对未来的核武器进行预测。

For question 1,在问题a中，利用Python和Excel处理数据，得出曾经拥有核武器的国家有China，France，India，Israel，North Korea，Pakistan，Russia，United Kingdom，South Africa，United States。在问题b中，将2002年和2022年的核武器储备量相减求差，排序得出United States核武器储备减少量最多。在问题c中，利用Python以五年为单位遍历核武试验次数，最终得到从1962到1966这五年间核弹实验次数最多，为422.0次。在问题d中，采用层次分析法，以核武增减量和试射次数为指标设定评价矩阵，得出过去十年North Korea是在核武器研究方面最活跃的国家。在问题e中，计算各个国家耗费的时间差，排序后得到United States是“不考虑武器”到“拥有核武器”的转变速度最快的国家。

For question 2, 首先采用时间序列预测法建立一个预测核武器数量的数学模型，然后利用贝叶斯网络分别对拥有核武器的国家数量和各国核武器数量进行神经网络训练，进而预测出未来100年内拥有核武器的国家数量、各国核武器数量变化趋势。再根据相关文献，推断出伊朗是未来100年拥有核武器的国家。根据模型预测出2123年全球核武器总数为5979。

For question 3, 收集相关数据，从摧毁地球表面和瓦解地球两方面考虑，第一方面，采用分析软件进行模拟试爆，得出合理爆炸半径，计算出摧毁地球表面需要151万颗沙皇炸弹。第二方面，利用物理公式推导得出瓦解地球的核弹数量约为1075万亿颗沙皇炸弹。然后结合相关文献，得出目前核弹并不能毁灭地球。其次采用聚类分析的方法对城市进行划分，利用分析软件得出摧毁各类城市需要的核弹数量，根据各类城市的数量和摧毁其需要的核弹数量进行计算，最终得到每个拥核国家最多应该拥有328颗核武器。

For question 4,我们撰写了一份非技术文章，向联合国描述我们的发现，并对所有国家提出一些可使用的建议。

Key words：层次分析法；时间序列预测模型；神经网络模型；聚类分析法；核武器预测

#### 1介绍

##### 1.1背景

核武器的威力相当强大，它能在瞬间释放出巨大的能量，其最大威力为一亿吨级。目前为止，沙皇炸弹是世界上威力最大的核弹，它也是人类目前为止制造的最强的核武器。当今世界以和平发展为时代主题，因此如何正确认识、合理预测核武器的发展趋势并探讨其相关问题已成为人们关注的焦点。

##### 1.2问题重述

自第二次世界大战结束后，核武器就一直是国际关系的焦点之一。在过去几年中，各个国家在不同的领域进行了大量的核武器试验。this paper mainly solves the following problems:

**Question 1**: a) 在给定的10个国家中，分析有哪些国家曾经拥有过核武器。b) 2002年至2022年期间，哪个国家核武器储备量变化最大。c)哪五年核武器试验次数最多。d) 根据已知数据，分析在最近10年内，哪个国家在核武器研究方面最活跃。e)从在给定10个国家中，哪个国家从“不考虑核武器”到“拥有核武器”的转变速度最快。

**Question 2:** a)建立预测核武器数量的数学模型，并预测未来100年拥有核武器的国家。b) 预测未来 100 年的核武器数量变化趋势，2123 年的核武器总数，以及每个国家的核武器数量。

**Question 3:** a)收集相关数据，选取合适的指标 ,建立数学模型,分析并计算至少需要多少枚核弹才能毁灭地球。b)分析现如今核弹的最大破坏力，判断这种威力能否能够毁灭地球。c)讨论全球的核弹量应限制在什么范围内，确定各个核拥国家的核武器安全数量。

**Question 4:** 准备一篇非技术文章，解释团队发现，并为所有国家提出几条建议。

#### 2假设

1. 假设各国核武器预测数量仅在客观情况下进行考虑，忽略了不可抗力因素。
2. 假设原子弹的爆炸半径确定，原子弹的对地面的爆炸近似为一个圆，且各国原子弹的爆炸范围相同。
3. 仅考虑核弹对地球的造成影响，忽略其他系统的干扰。
4. 假设相关数据网站提供的数据真实可靠，本文研究是基于真实的数据进行。

#### 3符号说明

|  |  |
| --- | --- |
| Symbols | Definitions |
| λ | 最大特征值 |
| CI | 一致性指标 |
| RI | 随机一致性指标 |
| CR | 核验系数 |
| E | 功 |
| G | 引力常量 |
| m | 地球质量 |
| r | 地球半径 |
|  | 第i行j列矩阵的值 |
|  | 归一化后第i行j列矩阵的值 |

#### 4问题分析

在进行问题分析之前，我们先进行了数据预处理，把官方数据分成了四个Excel表，分别是proliferation.xlsx、position.xlsx、stockpiles.xlsx、tests.xlsx。

##### 3.1基础数据分析

1. 在该问题中，需要结合已有数据，即“possession”这个指标利用Python和Excel进行判别。
2. 引入变化量指标，作为2022年的核武器储备数量与2002年的核武器储备数量的差值。这个差值即为过去的20年里各个国家的核武器储备的变化量。通过降序排列找出过去20年里核武器储备减少或增加最多的国家，并以图表形式呈现。
3. 利用python中的字典将年份与实验次数进行对应。再以五年为一单位进行年份的遍历，计算这五年核武器总值。通过迭代的算法依次判定大小，并将每次比较的较大值赋给最大值，以此来得到实验次数最多的五年。
4. 在这一问中，采用层次分析的方法。首先根据现实经验构建输入判别矩阵，需要计算特征向量和最大特征值，并利用python的numpy库分析构造的输入判别矩阵，返回最大特征值和其对应的特征向量。寻找最大特征值，进行一致性检验。之后输入数据，对数据进行归一化处理，运用判别矩阵的特征向量计算各个指标的权重，最后利用矩阵乘法得出结果。
5. 用Python统计每个国家第一次出现0和3的年份，将其放入Excel表中，然后利用Excel计算年份差。对年份差进行升序排列，第一名则是所求国家，之后进行可视化处理

##### 4.2预测核武器的数量

a)在该问题中，对预测核武器数量的模型采用时间序列预测法。而对预测拥有核武器国家数量的模型采用贝叶斯网络进行拟合，得出未来100年拥有核武器国家的数量，再根据相关文献推断出具体国家。

b)根据问题a）的模型预测未来100年核武器数量变化趋势和2123年的核武器总数。对预测各个国家核武器数量的模型采用贝叶斯神经网络进行拟合，将已知数据及问题二a）中得到的预测数据、国家往期拥核数据和该国核武器占全球核武器比重趋势纳入到输入参数，进行神经网络训练得到相关模型。训练完成后对各个国家核武器数量进行预测，最后输出可视化结果。

##### 4.3保护我们的地球

a）由于研究至少需要多少核弹，所以我们假设这个核弹是目前威力最强的沙皇炸弹。而在这个问题中我们共研究了两个层面。一是摧毁地球表面至少需要多少核弹，二是彻底瓦解地球至少需要多少核弹。对于第一个层面，我们通过利用分析软件进行试爆，选取三个具有代表性的地点进行模拟，得出合理爆炸半径，并根据核弹爆炸面积大于等于地球面积即可摧毁地球表面算出至少需要的核弹数量。对于第二个层面，我们利用物理公式和收集的数据，研究至少需要多少核弹才能让地球彻底解体。

b）根据相关模型和模拟的数据，可以得出核弹目前所拥有的最大破坏力，再结合当前数据，判断是否可以毁灭地球。

c）在这个问题中，首先我们查找到相关人口信息，进行聚类分析，分出五类城市。再通过试爆软件分析出摧毁各个类型的城市需要多少颗核武器，算出核武器的总数量，并最终进行汇总，均摊到每一个拥核国家，就是每个国家应该保持的核武器数量下限。

#### 5问题一的解决

##### 5.1 a）

通过“possession”指标对国家是否曾经拥有核武器进行度量，借助python，遍历各个国家每一年的“possession”指标，并对该指标进行检验，如果数值大于1，则说明曾经有过核武器。

通过对数据的处理，我们得到曾经拥有过核弹的国家是：China，France，India，Israel，North Korea，Pakistan，Russia，United Kingdom，South Africa，United States。

##### 5.2 b）

运用Python对各个国家在2002年和2022年核武器储备变化量进行统计分析，并运用Excel对2022年和2002年核武储备数量进行差值计算，最后进行降序排列，如表1所示。

表1核武器储备变化量

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Country | Abbreviation | Year | Stockpile | Change |
| Pakistan | PAK | 2022 | 165 | 139 |
| India | IND | 2022 | 160 | 137 |
| China | CHN | 2022 | 350 | 115 |
| North Korea | PRK | 2022 | 20 | 20 |
| Israel | ISR | 2022 | 90 | 14 |
| South Africa | ZAF | 2022 | 0 | 0 |
| France | FRA | 2022 | 290 | -60 |
| United Kingdom | GBR | 2022 | 180 | -100 |
| Russia | RUS | 2022 | 4477 | -5637 |
| United States | USA | 2022 | 3708 | -6749 |

通过对数据筛选、计算和排序，得到United States核武器储备减少量最多，进行可视化数据展示。

##### 5.3 c）

在这个问题中，只需要利用Python以五年为单位从1945年开始进行遍历，，统计每个五年的核武总量。之后通过依次比较，得到实验次数最多的五年，并进行可视化处理，如图1所示

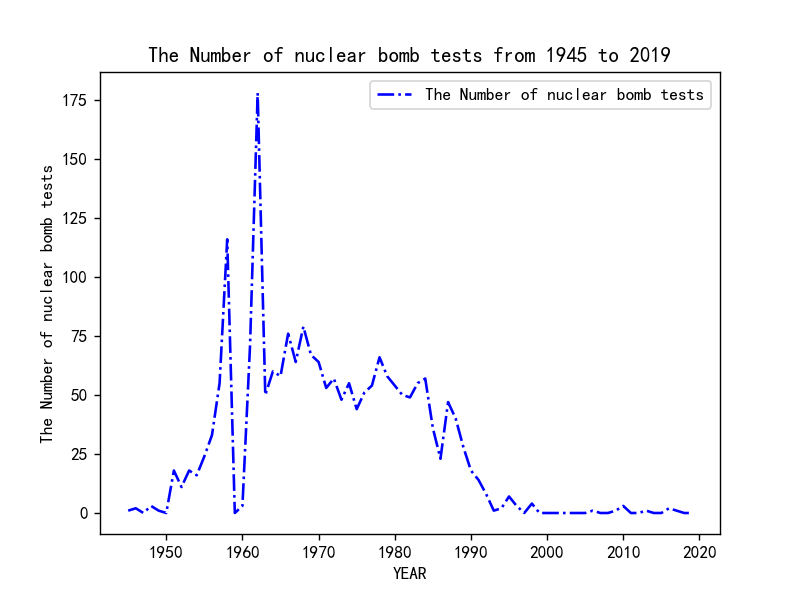


图1核试验次数

通过使用迭代法，得出从1962到1966这五年间核弹实验次数最多，为422.0次。

##### 5.4 d）

理论上可以说明一个国家核实验积极的变量有核武器增长的数量以及核武器试验的次数。而核武器试验的数量是一个国家进行核武器研究的重要指标，而核武器的增长相比而言则是一个附加的说明条件，故而我们选取核试验次数作为该国在核武器研究活跃度方面的重要因素并将核武器增长数量纳入考虑并最终通过层次分析法获得核武器研究积极度进行评分。

该问题采用层次分析法解决。层次分析法根据问题的性质和要达到的总目标，将问题分解为不同的组成因素，并按照因素间的相互关联影响以及隶属关系将因素按不同的层次聚集组合，形成一个多层次的分析结构模型。

首先进行数据预处理，通过Excel筛选以及排序得到各个国家2012与2022年的核武器数量，统计实验次数，得到相关数据。

之后设定一个评价矩阵，在这里假设评价矩阵是[[1,1/5],[5,1]]，其表示的是核武器试射的次数比核武器增加的数量显著重要。然后利用python中的np.linalg.eig函数求特征值和特征向量，并找最大特征值所在的行和列，及其对应的特征向量，处理特征向量为权向量。最后进行一致性检验，一致性指标用CI计算，CI越小，说明一致性越大，定义一致性指标为：

CI=

CI=0，有完全的一致性；CI 接近于0，有满意的一致性；CI 越大，不一致越严重。

为衡量CI 的大小，引入随机一致性指标 RI：

RI=

考虑到一致性的偏离可能是由于随机原因造成的，因此在检验判断矩阵是否具有满意的一致性时，还需将CI和随机一致性指标RI进行比较，得出检验系数CR，公式如下：

CR=

一般，如果CR<0.1 ，则认为该判断矩阵通过一致性检验，否则就不具有满意一致性。检验通过说明该评价矩阵可用。

读入数据进行处理，先对数据进行归一化，从而去量钢化。

然后利用权重向量进行计算得出两者之间的评价。最终结果如表2所示。

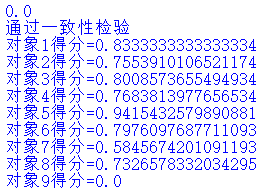


表2 核国家的核武器研发积极度评分

根据对象得分，我们可以知道核武研发积极的前三名为朝鲜、中国和印度。与此同时。同时对相关数据进行汇总形成图表3，在十年间只有朝鲜进行了核试验，进一步验证了我们的结果。

表3 核国家的增减量和试射次数

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Country | Abbreviation | 2012持有核弹数量 | 2022持有核弹数量 | 对象号码 | 核弹增量 | 十年间实验次数 |
| China | CHN | 240 | 350 | 1 | 110 | 0 |
| France | FRA | 300 | 290 | 2 | -10 | 0 |
| India | IND | 100 | 160 | 3 | 60 | 0 |
| Israel | ISR | 80 | 90 | 4 | 10 | 0 |
| North Korea | PRK | 0 | 20 | 5 | 20 | 4 |
| Pakistan | PAK | 110 | 165 | 6 | 55 | 0 |
| Russia | RUS | 4750 | 4477 | 7 | -273 | 0 |
| United Kingdom | GBR | 225 | 180 | 8 | -45 | 0 |
| United States | USA | 4881 | 3708 | 9 | -1173 | 0 |

通过表2以及层次分析可知North Korea在核武器研究方面最为活跃。

##### 5.5 e）

运用Python，对position.xlsx中的status进行遍历，找出每个国家之第一次出现0和3的时间段，并相互减去，之后导出，并进行可视化处理，得出如下图表：

表4 核国家拥核耗费时间

|  |  |  |
| --- | --- | --- |
| 国家 | 耗费时间 | 获得核武器年份 |
| United States | 7 | 1945 |
| Russia | 11 | 1949 |
| United Kingdom | 14 | 1952 |
| France | 22 | 1960 |
| China | 26 | 1964 |
| Israel | 29 | 1967 |
| South Africa | 41 | 1979 |
| India | 49 | 1987 |
| Pakistan | 49 | 1987 |
| North Korea | 68 | 2006 |

表5 核国家拥核耗费时间图表展示

根据表3可知从“不考虑武器”到“核试验，拥有核武器”的转变速度最快的国家是United States。

#### 6问题二的解决

##### 6.1 预测核武器数量模型的建立

根据以下思路建立时间序列分析模型。先进行ADF检验，分析p值，分析其是否可以显著性地拒绝序列不平稳的假设(P<0.05)。若呈现显著性(P<0.05)，则说明拒绝原假设，该序列为一个平稳的时间序列，反之则说明该序列为一个不平稳的时间序列。ADF检验结果如表4所示。

表1 ADF检验表

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ADF检验表 | | | | | | | |
| 变量 | 差分阶数 | t | P | AIC | 临界值 | | |
| 1% | 5% | 10% |
| Stockpile | 0 | -2.164 | 0.220 | 1102.297 | -3.519 | -2.9 | -2.587 |
| 1 | -2.139 | 0.229 | 1088.27 | -3.519 | -2.9 | -2.587 |
| 2 | -6.751 | 0.000\*\*\* | 1068.378 | -3.522 | -2.901 | -2.588 |
| 注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平 | | | | | | | |

该序列检验的结果显示，基于变量Stockpile:在差分为2阶时，显著性P值为0.000\*\*\*，水平上呈现显著性，拒绝原假设，这表明该序列为平稳的时间序列。最佳差分序列图如图2所示。

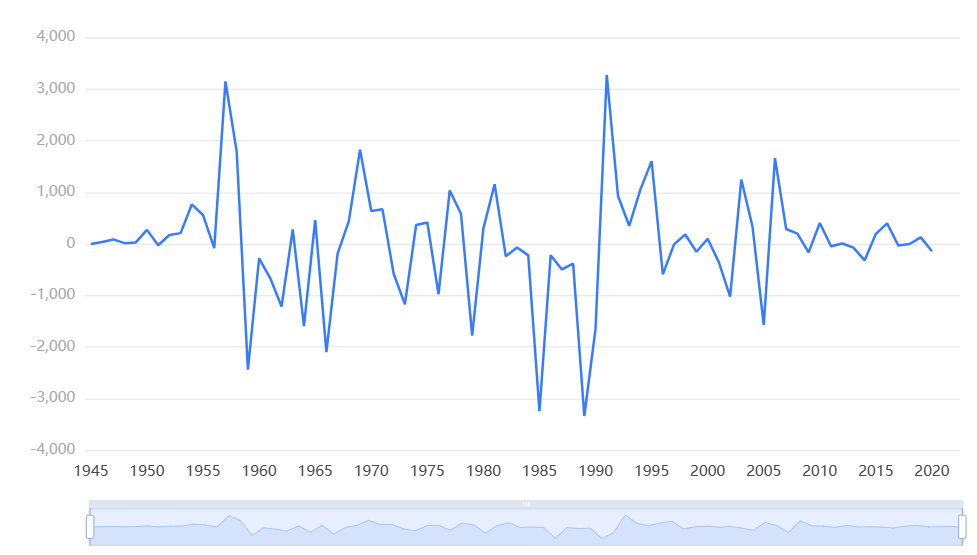


图2最佳差分序列图

查看差分前后数据对比图，判断是否平稳，同时对时间序列进行偏（自）相关分析，根据截尾情况估算其p、q值。最终差分数据自相关图如图3所示，最终差分数据偏自相关图如图4所示。

图3到图6：

green：Upper bound of ACF 95% confidence interval

yellow：Lower bound of ACF 95% confidence interval

blue：Autocorrelation coefficient

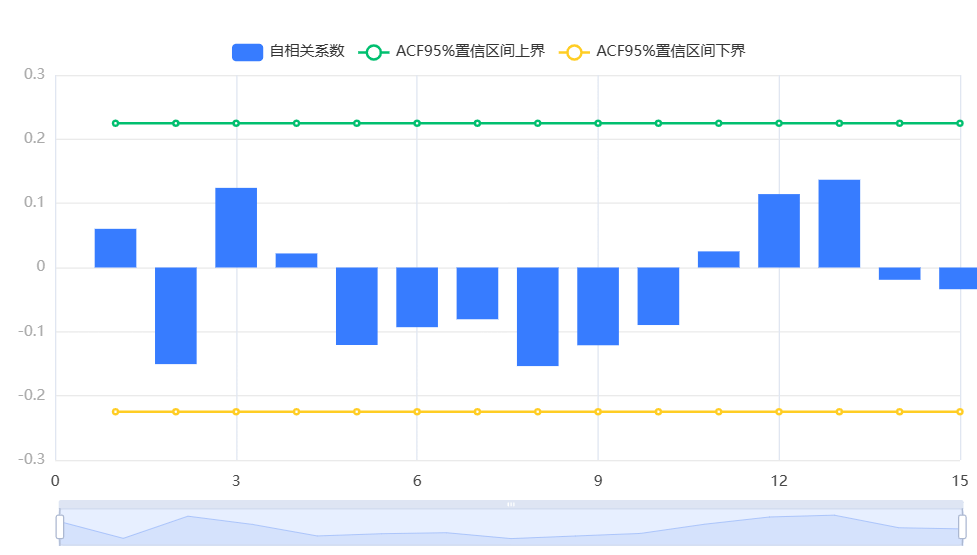


图3最终差分数据自相关图

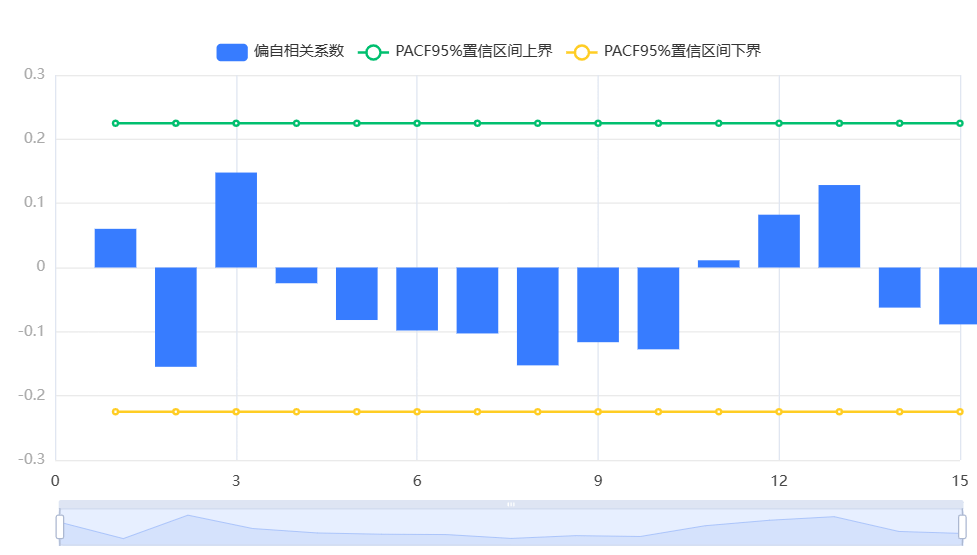


图4最终差分数据偏自相关图

图3展示了自相关图(ACF)，包括系数，置信上限和置信下限。横轴代表延迟数目，纵轴代表自相关系数。图4展示了偏自相关图(PACF)，包括系数，置信上限和置信下限。横轴代表延迟数目，纵轴代表偏自相关系数。

由图3和图4可以看出都是拖尾，并且原始数据差分后是平稳的时间序列，所以基本适合ARIMA。

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，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。

表5 ARIMA模型（0,2,0）检验表

|  |  |  |
| --- | --- | --- |
| ARIMA模型（0,2,0）检验表 | | |
| 项 | 符号 | 值 |
|  | Df Residuals | 75 |
| 样本数量 | N | 78 |
| Q统计量 | Q6(P值) | 0.287(0.592) |
| Q12(P值) | 5.344(0.501) |
| Q18(P值) | 11.251(0.508) |
| Q24(P值) | 15.963(0.595) |
| Q30(P值) | 25.005(0.405) |
| 信息准则 | AIC | 1284.436 |
| BIC | 1289.097 |
| 拟合优度 | R² | 0.979 |
| 注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平 | | |

系统基于AIC信息准则自动寻找最优参数，模型结果为ARIMA模型（0,2,0）检验表，基于变量：Stockpile，从Q统计量结果分析可以得到：Q6在水平上不呈现显著性，不能拒绝模型的残差为白噪声序列的假设，同时模型的拟合优度R²为0.979，模型表现优秀，模型基本满足要求。

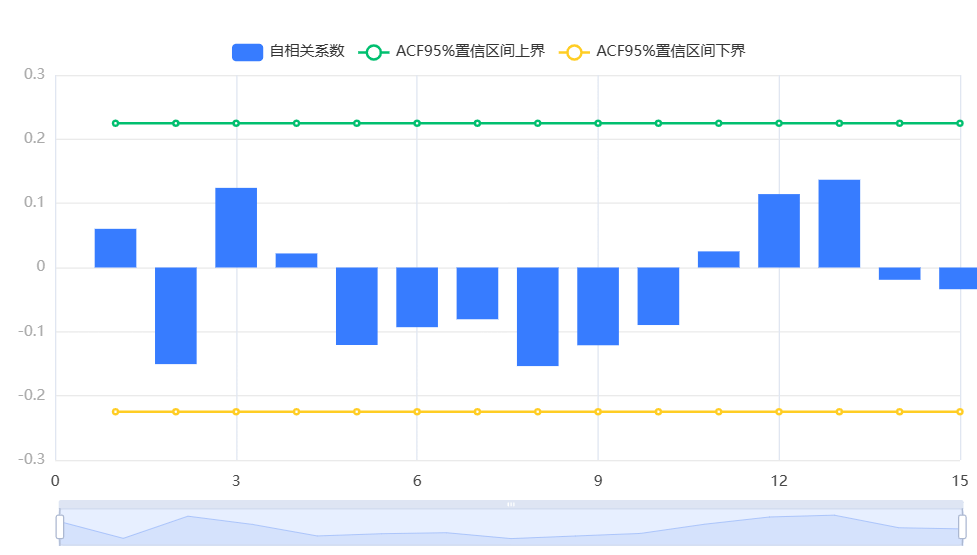


图5残差自相关图

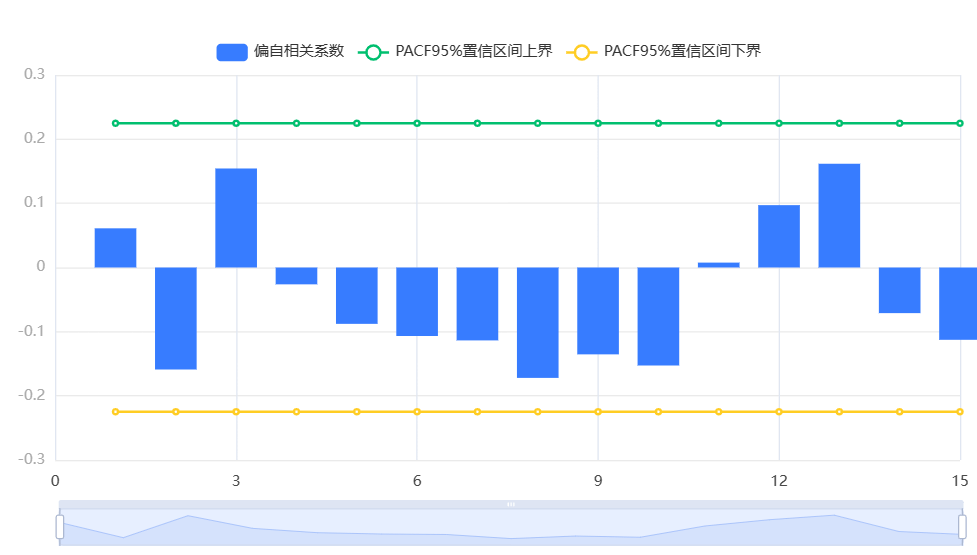


图6残差偏自相关图

图5展示了模型的残差自相关图(ACF)，包括系数，置信上限和置信下限。横轴代表延迟数目，纵轴代表自相关系数。图6展示了模型的残差偏自相关图(PACF)，包括系数，置信上限和置信下限。横轴代表延迟数目，纵轴代表偏自相关系数。

据图5和图6可知相关系数都位在置信区上下限之内，是白噪声序列，可直接用于时间序列分析。

表6 模型参数表

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 模型参数表 | | | | | | |
|  | 系数 | 标准差 | t | P>|t| | 0.025 | 0.975 |
| 常数 | -0.329 | 126.421 | -0.003 | 0.998 | -248.109 | 247.451 |
| 注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平 | | | | | | |

表6展示了本次模型参数结果，包括模型的系数、标准差，T检验结果等，用于分析模型公式。

基于变量Stockpile，系统基于AIC信息准则自动寻找最优参数，模型结果为**ARIMA模型（0,2,0）检验表且基于2差分数据**，模型公式如下：

y(t)=-0.329

通过上述分析，对该模型的拟合值、预测值进行可视化展示，如图7所示。

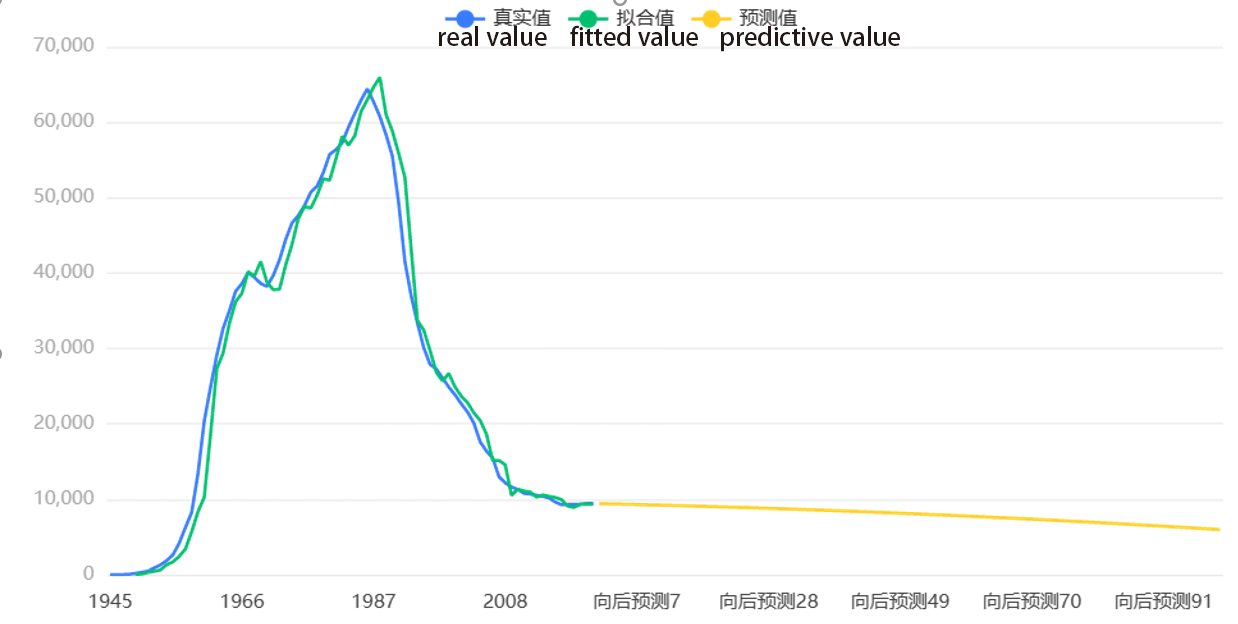


图7 时间序列模型的拟合与预测

根据预测模型可知道具体数据，具体数据由于篇幅原因加在论文后，2123年的核武器总数将达到5978.815789473637，约为5979。

##### 6.2预测未来100年拥有核武器的国家

###### 6.2.1预测拥有核武器国家数量的模型的建立

具体的算法步骤如下：

1.首先要对数据进行处理。将所给表转化为元胞数组，并将Year到Possession中的内容全部写入新列表。

2.其次进行模型的训练。默认的lmh函数训练时间序列效果很差，采用贝叶斯正则化算法，所用时间更多但是准度更高。建立非线性自结合的时间神经网络，并采用preparents函数进行数据的初始化，对preparents转化后的数据进行训练，70%的输入向量与目标向量用于训练，15%用于防止过拟合，15%用于测试，测试模拟的结果与实际的结果之间的误差，然后保存训练的模型。

3.然后进行预测。读入模型，预测过去的数值进行拟合，尝试一下拟合效果。再依次把已知数据放入模型中，迭代预测后续，重复之前的步骤。

4.最后显示预测结果。将得到的数据进行可视化展示，如图8所示。

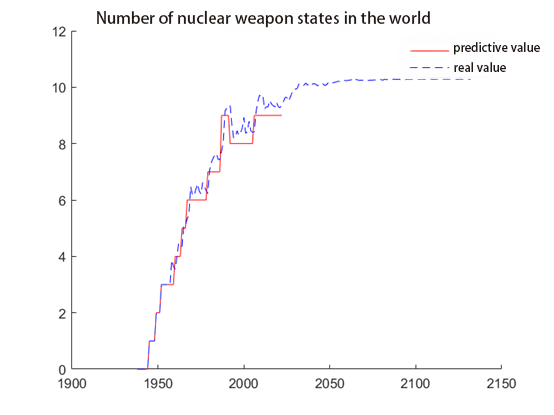


图8全球拥核国家数量

根据模型预测的结果可知，未来100年拥有核武器的国家将达到10个。所给数据中，已知有9个国家拥有核武器，那么新增的一个国家需要结合相关文献进行分析。

###### 6.2.2 根据文献推断国家

通过收集相关文献，我们认为未来100年拥有核武器的新增国家可能是伊朗。根据文献记载，从2006年开始伊朗的“绿盐项目”、高能炸药试验以及“流星计划”等都有可能涉及发展核武器。截至2015年５月，伊朗在纳坦兹和福尔多的两座燃料浓缩厂中共装有19138台离心机，其中9156台IＲ－１型投产运行，其余的已安装完毕等待通料运行。2021 年 1 月，伊朗政府发言人阿里 · 拉比伊宣布，伊朗已开始把浓缩铀丰度提高至 20%。尽管伊朗目前在技术方面仍有很长的路要走，但是未来100内伊朗还是很有可能成为拥有核武器的国家。

所以，结合上述预测和已知数据，我们认为未来100年拥有核武器的国家有伊朗、美国、俄罗斯、英国、中国、法国、印度、朝鲜、巴基斯坦、以色列这十个国家。

##### 6.3预测每个国家的核武器数量

因为考虑到各个国家未来的核武器变化数量不能简单地认为和世界核武器总数变化完全相同（但是趋势大致相同）。并且由于国际形势和世界背景（核武器呈现逐年减少的态势、世界和平与发展占据主流地位，以及联合国宪章等等因素的制约），我们认为应将国家往期拥核数据、全球整体核武器数量以及该国核武器占比比重趋势一起纳入到神经网络输入参数中，以便对各个国家进行针对性的预测。

具体的算法步骤如下：

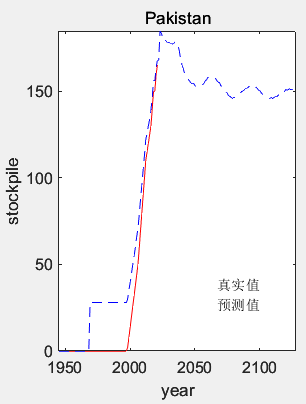
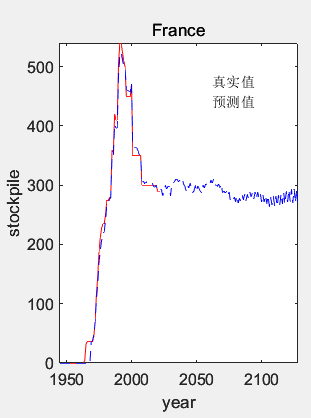
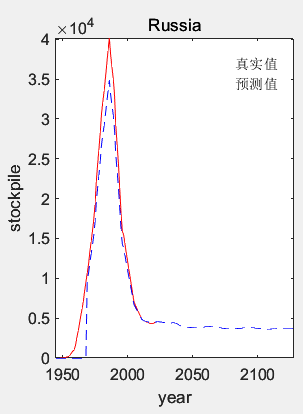
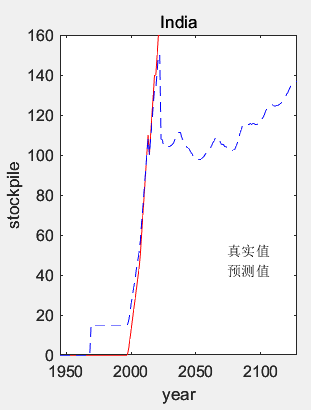
1.首先初始化数据。将国家标签转为数字编号，并将其变成一个categorical类型的数组，即分类数组。然后去重统计国家数，并统计每个国家每一年的核武器数量和全球核武器总量。

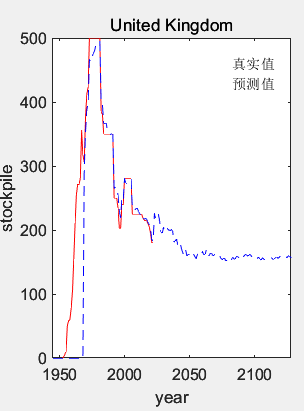
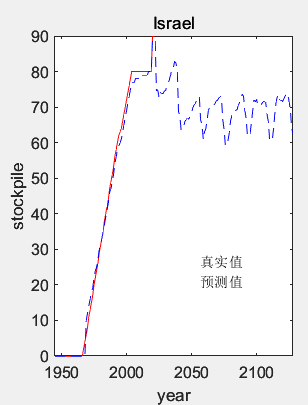
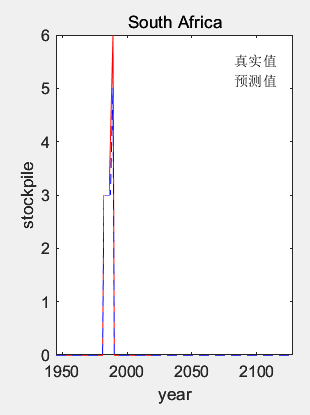
2.进行模型的训练。将国家往期核武器拥有量、全球整体核武器数量以及该国核武器占比纳入考虑范围，给定输入参数，保留上文预测全球核武数量的结果，采用贝叶斯正则化算法对每个国家进行训练，并保留对应的训练模型。

3.预测各个国家的核武器数量。读入训练模型，然后储存每个国家对应训练的神经网络，读入每个国家的核武器总量，运用各自对应的神经网络预测其未来的核武器数量。

4.可视化展示。利用Matlab将各个国家预测的拥核数量进行可视化展示。如下图10所示。

The Blue one is predictive value，the red one is real value





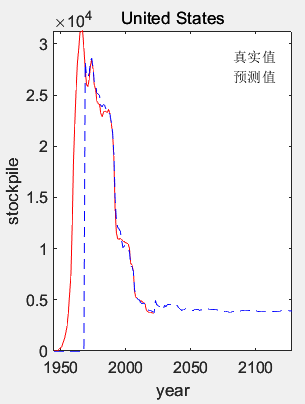
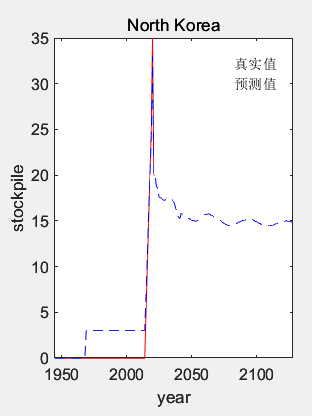


图9各国拥有核武器数量预测

#### 7问题三的解决

##### 7.1研究多少枚核弹才能毁灭地球

###### 7.1.1层面一

在本题中，由于我们所研究的是至少需要多少枚核弹才能毁灭地球，所以我们假设所用核弹是目前世界上威力最大的核弹，即沙皇炸弹。根据上述假设，核弹的爆炸半径确定，核弹的对地面的爆炸近似为一个圆，且各国核弹的爆炸范围相同。所以我们认为核弹爆炸不受地形等外力影响，因此核弹的爆炸点及波及周边为平原。并且遵循层面一的核心思想：核弹爆炸面积大于等于地球面积即可摧毁地球表面。所以我们每次所取试爆点为最大杀伤半径，任意两个爆炸范围不相互重叠。

选取平原（东京，莫斯科，圣保罗）进行模拟试炸，取平均值结果如下：

1. 火球半径：4.62 km（67.1 km²）。核火球的最大尺寸，与地面损伤的相关性取决于爆炸高度。如果它接触到地面，放射性沉降物的数量会显著增加。火球内的任何东西都会有效地蒸发。可忽略沉降物的最小爆炸高度：4.16km。
2. 辐射半径（500雷姆）：5.05 km（80.2 km²）。500雷姆电离辐射剂量，可能致命，约1个月；15%的幸存者最终会因暴露而死于癌症。
3. 重爆炸损伤半径（20 psi）：10.4 km（339 km²）。在20磅/平方英寸的超压下，严重建造的混凝土建筑遭到严重破坏或拆除，死亡人数接近100%。通常用作城市严重破坏的基准。使这种效应最大化的最佳爆破高度为6.7千米。
4. 热辐射半径（三度烧伤）：60.1 km（11350 km²）。三度烧伤遍及皮肤的各个层，通常是无痛的，因为它们会破坏疼痛神经。它们会造成严重的疤痕或残疾，并可能需要截肢。在此产量下，三度烧伤的100%概率为13.6cal/cm2。

在这里，由于我们讨论的是毁灭地球，故而就以重爆炸损伤半径为基准进行计算，随后通过收集数据得到地球表面积为510072000 km²，经过简单除法运算可知，如果核弹爆炸无缝衔接，那么510072000/339= 1504637，也就是大约151万个沙皇导弹才能摧毁地球表面。

###### 7.1.2层面二

物质都是靠引力结合在一起，必须通过做功才能使其分开。而同一物体不同部分结合在一起的能量就是结合能。通过沙皇炸弹爆炸所产生的能量与结合能相除进而得出毁灭地球所需要的核弹数量。利用简单物理公式进行推导得：

E=·G·

通过带入引力常量G、地球质量m、地球半径r进行计算，可得要使得地球的物质移动到无限远处需做功E=2.25·1032（J）。

与此同时，每公斤三硝基甲苯可产生4184000焦耳的能量，1吨三硝基甲苯相等于4184000000焦耳，而沙皇炸弹所产生的能量是5000万吨三硝基甲苯的量，经过乘法计算可以得到，一颗沙皇炸弹所产生的能量是2.092e+17焦耳，再经过除法计算，可知需要1075525812619502颗沙皇炸弹才能彻底瓦解地球。

##### 7.2目前核弹能否摧毁地球

以威力最大的沙皇导弹为例，根据上文模拟试炸试验结果表明，重爆炸损伤半径为10.4 km。在20磅/平方英寸的超压下，严重建造的混凝土建筑遭到严重破坏或拆除，死亡人数接近100%。这种效应最大化的最佳爆破高度为6.7千米，需要上百万颗才能把地球表面摧毁，以人类现在的科技水平和地球资源，是完全制造不出这么多的导弹，因此不足以毁灭地球。

##### 7.3研究核弹数量应限制的范围

###### 7.3.1聚类分析

通过查找相关资料得到人口信息并采用聚类分析的方法对城市人口进行划分。为了方便研究，我们借鉴了城市规模划分标准将其分成了5类城市，分别是小城市、中等城市、大城市、特大城市、超大城市。利用SPSS对其进行聚类分析，如图11所示。

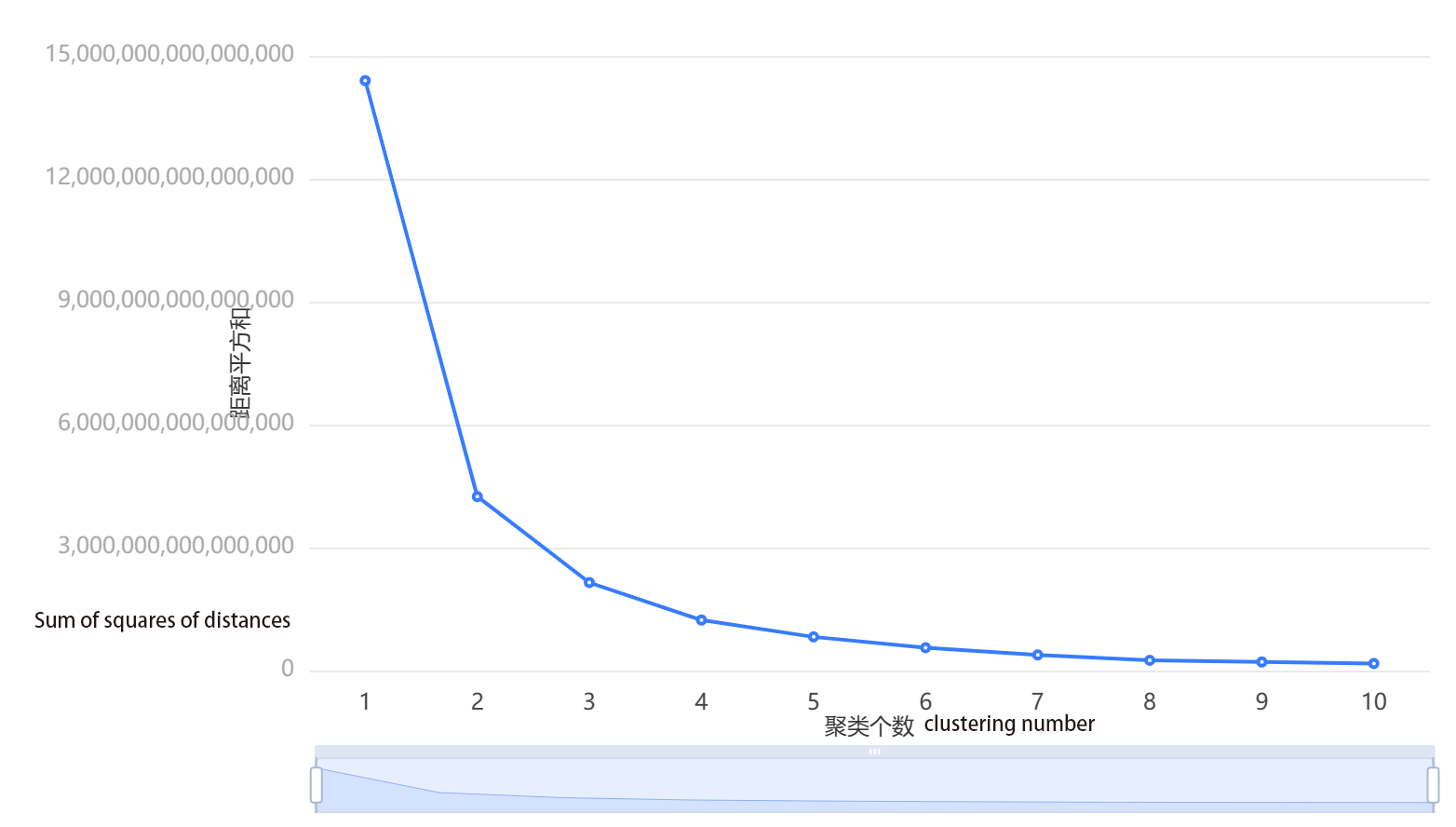


图11 聚类数对比图

该图横坐标是聚类个数，纵坐标是K均值聚类的损失函数，是所有样本到类别中心的距离平方和，也就是误差平方和（值越小说明聚类效果越好），从图中可以看出聚类个数在4之后效果较好。

然后进行字段差异性分析，如表7所示。

表7 字段差异性分析

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 聚类类别（平均值±标准差） | | | | | F | P |
| 类别3(n=758) | 类别1(n=143) | 类别5(n=43) | 类别2(n=19) | 类别4(n=9) |
| Population Estimate | 1049296.834±491403.487 | 3882370.629±1049850.611 | 8981976.744±1974328.608 | 17235000.0±2221086.421 | 27858222.222±5358298.722 | 3994.543 | 0.000\*\*\* |
| Square Kilometers | 305.439±335.098 | 946.315±1017.119 | 1975.814±1889.251 | 3208.105±2670.117 | 3588.222±2087.7 | 165.194 | 0.000\*\*\* |
| 注：\*\*\*、\*\*、\*分别代表1%、5%、10%的显著性水平 | | | | | | | |

对于变量Population Estimate，显著性P值为0.000\*\*\*，水平上呈现显著性，拒绝原假设，说明变量Population Estimate在聚类分析划分的类别之间存在显著性差异；对于变量Square Kilometers，显著性P值为0.000\*\*\*，水平上呈现显著性，拒绝原假设，说明变量Square Kilometers在聚类分析划分的类别之间存在显著性差异。

根据聚类汇总分析各聚类类别的频数，如表8所示。

表8聚类汇总

|  |  |  |
| --- | --- | --- |
| 聚类类别 | 频数 | 百分比% |
| 聚类类别\_1 | 143 | 14.712 |
| 聚类类别\_2 | 19 | 1.955 |
| 聚类类别\_3 | 758 | 77.984 |
| 聚类类别\_4 | 9 | 0.926 |
| 聚类类别\_5 | 43 | 4.424 |
| 合计 | 972 | 100.0 |

聚类类别\_1的频数为143，所占百分比为14.712%；聚类类别\_2的频数为19，所占百分比为1.955%；聚类类别\_3的频数为758，所占百分比为77.984%；聚类类别\_4的频数为9，所占百分比为0.926%；聚类类别\_5的频数为43，所占百分比为4.424%。

根据数据集聚类标注可以知道每一个样本数据被分到哪个类别，其次根据聚类中心坐标分析各样本与中心点的距离，最后通过中心值排序我们可以判断样本类别，如表9所示。

表9聚类中心点坐标

|  |  |  |
| --- | --- | --- |
| 聚类种类 | 中心值\_Population Estimate | 中心值\_Square Kilometers |
| 1 | 3882370.629370628 | 946.3146853146851 |
| 2 | 17234999.999999996 | 3208.105263157894 |
| 3 | 1049296.8337730826 | 305.43931398416896 |
| 4 | 27858222.222222216 | 3588.2222222222217 |
| 5 | 8981976.744186044 | 1975.813953488372 |

通过图中的中心值排序我们可以判断，4为超大城市、2为特大城市、5为大城市、3为小城市、1为中等城市。

最后进行相关指标的评价。如表10所示。

表10评价指标

|  |  |  |
| --- | --- | --- |
| 轮廓系数 | DBI | CH |
| 0.726 | 0.525 | 3994.54 |

对于一个样本集合，它的轮廓系数是所有样本轮廓系数的平均值。轮廓系数的取值范围是[-1,1]，本题中轮廓系数为0.726，分数较高，聚类效果较好。 DBI指标用来衡量任意两个簇的簇内距离之后与簇间距离之比。本题中该指标较小，聚类效果较好。CH指标由分离度与紧密度的比值得到，本题中CH较大，DBI适中，聚类效果良好。

###### 7.3.2限制结果

通过NUKEMAP by Alex Wellerstein模拟并计算出摧毁各个类型的一个城市需要多少颗核武器，最终进行汇总。而我们的爆炸半径则全部采用重爆炸损伤半径，即重度爆炸损伤半径所产生的伤亡。摧毁各类城市的核武器如表11所示，各类城市数量及其对应的核武器数量如表12所示。

表11摧毁各类城市核武器汇总

|  |  |
| --- | --- |
| 超大城市 | 沙皇导弹（50Mt，249平方千米，14颗） |
| 特大城市 | 沙皇导弹（50Mt，249平方千米，13颗） |
| 大城市 | 沙皇导弹（50Mt，249平方千米，8颗） |
| 中等城市 | castle bravo（15MT，90.5平方千米，2颗）沙皇导弹（50Mt，249平方千米，3颗） |
| 小城市 | castle bravo（15MT，90.5平方千米，1颗）沙皇导弹（50Mt，249平方千米，1颗） |

表12 各类城市数量及其核武器总量汇总

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 城市类别 | 人口数量平均值 | 城市面积平均值（立方千米） | 数量 | 总数 |
| 小城市 | 1049296.83 | 305.44 | 758 | 1516 |
| 中等城市 | 3882370.63 | 946.31 | 143 | 715 |
| 大城市 | 8981976.74 | 1975.81 | 43 | 344 |
| 特大城市 | 17235000.00 | 3208.11 | 19 | 247 |
| 超大城市 | 27858222.22 | 3588.22 | 9 | 126 |

最后算出重度爆炸损伤的情况下，只需要2948颗导弹就可以毁灭世界的基本主要城市，考虑到人类技术水平的进步所带来的武器杀伤能力增强，以及当今时代以和平发展为主题，未来的核武器应当限制的较少为好，所以平均到每个拥核国家之后，这些国家中每个国家最多应该拥有328颗核武器。

#### 8研究发现及建议

为了了解核武器在过去几十年及未来一百年发展情况，我们从核国家的数量、核武器的储备量及核武器试验次数三方面进行了分析，采用了时间序列分析模型对数据进行预测，通过对比得出，核弹数量在逐渐减少趋于平稳，在2123年的核武器总数将达到5979颗。

然后我们采用贝叶斯正则化算法，建立了非线性自结合的时间序列网络，发现未来100年拥有核武器的国家将达到10个，其中伊朗是新增的核武器国家。

最后我们利用沙皇炸弹对平原进行模拟测试得到其最大破坏半径并得出摧毁地球表面需要151万个沙皇炸弹，又根据物理公式推导得出彻底瓦解地球需要1075万亿颗沙皇炸弹。因此得出结论人类目前无法通过核弹毁灭地球。最后利用聚类分析并进行相应计算得出每个拥有核武器的国家最多应该拥有328颗核武器。

在此，我们提出以下几点建议:

（1）在不损害人类利益和国际安全的前提下，有效使用和管理武器系统；

（2）根据国际社会发展及普遍要求进行裁军；

（3）推动试验技术，加快以和平为目的的核试验，促进核技术现代化发展；

（4）通过外交、政治、经济乃至军事手段对进行核武器扩散的行为予以制止；

（5）限制核武器扩散。

#### 9优点和缺点

##### 9.1优点

1.分析世界核武数量变化表所使用的时间序列模型，应用了国内先进的建模工具SPSSPRO云软件，经过了单位根检验与偏自相关性分析，自动进行参数寻优，最终拟合优度R方达到了0.979，稳定性比较高。

2.在对各国核武进行预测的过程中，运用贝叶斯正则化算法训练时间序列，建立神经网络，采取在这一高精度的算法，所得效果较佳。

3.在第三大问中，对于毁灭地球这一定义思考得较为全面，分析了两种结果，通过美国的模拟核弹网站进行核弹的破坏力预测，精度比较高。

4.3C问中聚类分析较为准确，轮廓系数与DBI均在可接受的程度，分类较为明显。同时的话该模型建立较为新颖，具有一定的创新性。

5.三大问算法总体思路较为简便，有的问运用Excel等软件方便得出结果，综合利用Python、MATLAB等处理数据，模型与思路设计简洁实用效率高。

##### 9.2缺点

1.由于“黑匣子”问题的存在，神经网络在学习的过程中较为难以纠偏，需要耗费较多的精力，该算法虽然精度较高但是需要大量的时间进行训练，使用MATLAB运行运行时间耗费比较多，可能需要1分钟左右。

2.判断核武器研究最活跃的国家采取了层次分析法，这一算法的判别矩阵由个人主观构建，带有一定的主观性。

3.在第三问核武器的计算中，本模型考虑的因素较为理想，降低了该模型的普适性和推广能力。

#### 10 参考文献

[1] Scientific Platform Serving for Statistics Professional 2021. SPSSPRO. (Version 1.0.11)[Online Application Software]. Retrieved from https://www.spsspro.com.  
[2] 于宁莉,易东云,涂先勤.时间序列中自相关与偏相关函数分析[J].数学理论与应用,2007(01):54-57.

[3] 王燕．应用时间序列分析[M]．北京：中国人民大学出版社 2005.

[4] Scientific Platform Serving for Statistics Professional 2021. SPSSPRO. (Version 1.0.11)[Online Application Software]. Retrieved from https://www.spsspro.com.  
[5] Saroj,Kavita.Review:study on simple k mean and modified K mean clustering technique[J].International Journal of Computer Science Engineering and Technology,2016,6(7)：279-281.

[6]史建斌,朱剑钰.试析“联合全面行动计划”对伊朗核能力的影响.国际论坛,2016(3):1-6.

[7]杨晨,陈扩军. 伊朗已能够制造核弹 美国持续增兵海湾 美国制裁伊朗不得人心，中国不惧压力保持与伊经贸交往[J]. 祖国,2019,(12):40-43.

相关数据：

1.核弹数量预测时间序列数据

|  |  |
| --- | --- |
| 预测值 | |
| 阶数（时间） | 预测结果 |
| 1 | 9421.671052631578 |
| 2 | 9403.013157894737 |
| 3 | 9384.026315789473 |
| 4 | 9364.710526315788 |
| 5 | 9345.065789473683 |
| 6 | 9325.092105263157 |
| 7 | 9304.78947368421 |
| 8 | 9284.157894736842 |
| 9 | 9263.197368421052 |
| 10 | 9241.907894736842 |
| 11 | 9220.28947368421 |
| 12 | 9198.342105263157 |
| 13 | 9176.065789473683 |
| 14 | 9153.460526315788 |
| 15 | 9130.526315789471 |
| 16 | 9107.263157894735 |
| 17 | 9083.671052631576 |
| 18 | 9059.749999999998 |
| 19 | 9035.499999999998 |
| 20 | 9010.921052631576 |
| 21 | 8986.013157894735 |
| 22 | 8960.776315789471 |
| 23 | 8935.210526315786 |
| 24 | 8909.315789473681 |
| 25 | 8883.092105263155 |
| 26 | 8856.539473684206 |
| 27 | 8829.657894736838 |
| 28 | 8802.447368421048 |
| 29 | 8774.907894736838 |
| 30 | 8747.039473684206 |
| 31 | 8718.842105263153 |
| 32 | 8690.31578947368 |
| 33 | 8661.460526315785 |
| 34 | 8632.276315789468 |
| 35 | 8602.763157894731 |
| 36 | 8572.921052631573 |
| 37 | 8542.749999999993 |
| 38 | 8512.249999999993 |
| 39 | 8481.421052631571 |
| 40 | 8450.26315789473 |
| 41 | 8418.776315789466 |
| 42 | 8386.960526315781 |
| 43 | 8354.815789473676 |
| 44 | 8322.34210526315 |
| 45 | 8289.5394736842 |
| 46 | 8256.407894736833 |
| 47 | 8222.947368421042 |
| 48 | 8189.157894736832 |
| 49 | 8155.0394736842 |
| 50 | 8120.592105263147 |
| 51 | 8085.815789473672 |
| 52 | 8050.710526315777 |
| 53 | 8015.2763157894615 |
| 54 | 7979.513157894724 |
| 55 | 7943.4210526315655 |
| 56 | 7906.999999999986 |
| 57 | 7870.249999999985 |
| 58 | 7833.171052631564 |
| 59 | 7795.763157894721 |
| 60 | 7758.026315789458 |
| 61 | 7719.960526315773 |
| 62 | 7681.565789473667 |
| 63 | 7642.84210526314 |
| 64 | 7603.789473684192 |
| 65 | 7564.4078947368225 |
| 66 | 7524.6973684210325 |
| 67 | 7484.657894736822 |
| 68 | 7444.289473684189 |
| 69 | 7403.592105263136 |
| 70 | 7362.565789473661 |
| 71 | 7321.210526315766 |
| 72 | 7279.52631578945 |
| 73 | 7237.513157894712 |
| 74 | 7195.171052631554 |
| 75 | 7152.499999999974 |
| 76 | 7109.499999999973 |
| 77 | 7066.171052631551 |
| 78 | 7022.5131578947085 |
| 79 | 6978.526315789444 |
| 80 | 6934.210526315759 |
| 81 | 6889.565789473653 |
| 82 | 6844.5921052631265 |
| 83 | 6799.289473684178 |
| 84 | 6753.657894736809 |
| 85 | 6707.697368421019 |
| 86 | 6661.407894736807 |
| 87 | 6614.7894736841745 |
| 88 | 6567.842105263121 |
| 89 | 6520.565789473647 |
| 90 | 6472.960526315751 |
| 91 | 6425.026315789434 |
| 92 | 6376.763157894697 |
| 93 | 6328.171052631538 |
| 94 | 6279.249999999958 |
| 95 | 6229.999999999957 |
| 96 | 6180.4210526315355 |
| 97 | 6130.513157894692 |
| 98 | 6080.276315789428 |
| 99 | 6029.710526315743 |
| 100 | 5978.815789473637 |