

选题：A

论文题目：

建模： 成宇琛 电气 1813 班
编程： 孙瑞 计算机 1806 班
论文： 邓思彤 电气 1813 班

2019 年 5 月 3 日

摘要

称为温室效应。

关键字— 温室效应，碳排放

目录

摘要	1
1 问题重述	3
1.1 问题背景	3
1.2 问题提出	3
2 问题分析	4
2.1 问题 1：我国能源结构分析	4
2.1.1 总结	5
2.2 问题 2：碳排放预测模型	5
2.2.1 碳排放影响因素分析	5
2.2.2 模型选择	5
2.3 问题 3：政策建议	6
3 符号说明和模型假设	6
3.1 符号说明	6
3.2 模型假设	6
4 模型建立及求解	7
4.1 数据预处理	7
4.1.1 初始数据计算	7
4.1.2 数据分类	7
4.1.3 数据归一化处理	7
4.2 BP 神经网络结构构建	8
4.2.1 结构展示	8
4.2.2 设计网络输入层和输出层	8
4.2.3 选取隐含层节点数	9
4.3 BP 神经网络函数、参数设定	9
4.3.1 选取激励函数	9
4.3.2 选取学习速率	9
5 模型评价与改进	10
6 政策建议	11
参考文献	12
索引	13

A 附件	14
A.1 附件 1	14
A.2 附件 2	14

1 问题重述

1.1 问题背景

受温室效应的影响，全球气候不断恶化，这严重影响了自然生态环境和人类生活环境。其中二氧化碳等温室气体大规模排放被认为是引起温室效应的主要原因。

当前，中国碳排放量已经超过美国和欧洲碳排放量之和，排放的温室气体，占全球温室气体排放总量的逾四分之一^[1]。中国作为全球最大的碳排放国家，面临着巨大的减排压力。

从可持续发展的角度来看，探索碳排放增长的内在因素、开展碳减排策略的研究，对我国实施碳排放减排政策和低碳经济发展战略，以至减缓全球温室效应增长，具有重要的理论和实际意义。

碳排放核算方法采用联合国提供的 IPCC 方法：

$$C = \sum E_i \times F_i$$

其中 C 为总碳排放量， E_i 为能源 i 消耗量， F_i 为能源 i 的碳排放系数。各系数取值参见表 1 所示。

表 1: 各化石能源碳排放系数

能源	煤炭	石油	天然气
碳排放系数	0.7476	0.5825	0.4435

附件 1 中列出了中国近些年来的一些经济数据。

1.2 问题提出

问题 1. 分析我国能源消费结构。

问题 2. 确定碳排放量影响因素，并建立碳排放预测模型。

问题 3. 根据模型，提出节能减排政策建议。

2 问题分析

2.1 问题 1：我国能源结构分析

我们将煤炭、石油、天然气和一次电力及其他等能源占能源消费总量的比重，按时间序列作出折线图（见图 1）。

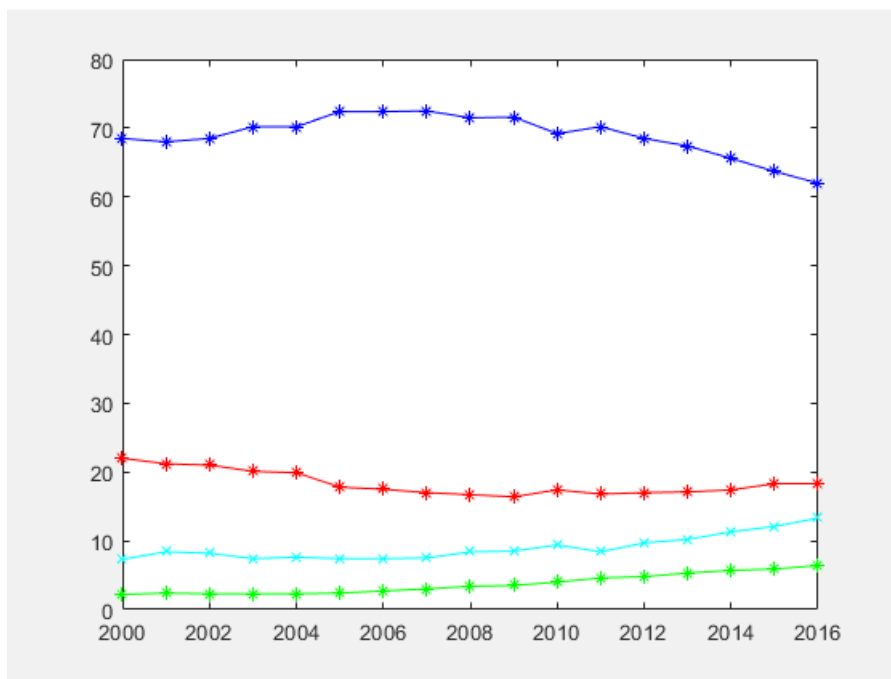


图 1: 蓝色曲线为煤炭占总能源的比例，红色为石油，绿色为天然气，青色为一次电力及其他能源。

由图 1，沿 Y 轴可以分析出以下几点：

1. 煤炭消费占我国能源消费的首要地位，其比例始终居于 60% 之上。
2. 石油消费为最主要的辅助能源，其比例一直在 20% 附近。
3. 一次电力能源与天然气在总能源中的比例相对较低，大体上低于 10% 。

再沿时间轴可以分析出：

1. 煤炭消费在波动中整体呈现缓步下降趋势，但仍然占据我国能源消费的首要地位。
2. 石油消费在参考年限中呈稳定趋势，略有波动和下滑。但与一次电力及其他的距离不断拉近。
3. 一次电力及其他能源和天然气能源消费是呈逐年上升趋势，且天然气的上升速度随年份增加而增加。但所占比例仍然不足。

2.1.1 总结

我国的能源消费结构始终为以煤炭消费为主导地位，并以石油消费为主要辅助消费，但也存在天然气、一次电力及其他能源等能源占据市场。能源消费市场呈现“一超多强”局面。

天然气和一次电力这样的清洁能源与石油、煤炭消费所占的比重不断拉近，但仍未占据可观消费比例。这说明我国的能源消费结构在不断优化的过程中还有很大的发展空间。

2.2 问题 2：碳排放预测模型

2.2.1 碳排放影响因素分析

为建立模型，需要对碳排放的影响因素进行分析。根据有关文献，碳排放影响因素一般包括：

1. 人口因素；
2. 城镇化率；
3. 经济发展水平：人均 GDP 或者消除价格波动影响的国内生产总值；
4. 人均碳排放量；
5. 能源消费强度：能源消费量与 GDP 之比；
6. 能源消费结构：各种能源所占比例，可以用煤炭比例来表示；
7. 产业结构：三类产业占比，可以用第二产业占比表示；
8. 国际贸易：出口额占 GDP 比重。

我们选取，如表 2 中变量，作为主要影响因素并确定其计算方法。

2.2.2 模型选择

由上述可知，影响中国碳排放量的因素繁多，此外碳排放量与影响因素之间是非线性关系。

BP神经网络是一种按误差反向传播（简称误差反传）训练的多层前馈网络，其算法称为BP算法，它的基本思想是梯度下降法，利用梯度搜索技术，以期使网络的实际输出值和期望输出值的误差均方差为最小。

BP神经网络模型具有较强的自组织、自适应与自学习能力，非线性映射能力，泛化能力以及容错能力，因此我们选用BP神经网络预测模型对我国碳排放量进行预测。

表 2: 模型变量说明

变量	计算方法
碳排放量	各能源按碳排放系数求和（参见问题重述中： IPCC 方法 ）
总人口数	城乡人口之和
GDP	各产业生产总值之和
产业结构	第二产业与 GDP 比值
城镇化率	城镇人口与总人口比值
经济发展水平	GDP 与人口比值
国际贸易	出口额与 GDP 比值
人均碳排放量	碳排放量与总人口比值
能源消费强度	能源消费量与 GDP 比值

2.3 问题 3：政策建议

参见 [7 政策建议](#)。

3 符号说明和模型假设

3.1 符号说明

表 3: 符号说明

符号	说明
Y	碳排放总量
$X1$	总人口数
$X2$	GDP
$X3$	产业结构
$X4$	城镇化率
$X5$	经济发展水平
$X6$	国际贸易
$X7$	人均碳排放量
$X8$	能源消费强度

3.2 模型假设

1. 没有外在的、突发的影响或变化，如：能源革命，能源枯竭……即总体碳排放是以某种趋势变化的，总体能源结构稳定；

2. 限定碳排放主要影响因素在总人口，GDP，产业结构，城镇化率，经济发展水平，国际贸易，人均碳排放量，能源消费强度之中；
3. 不考虑给定的数据的资金时效性，及给定的 GDP 已消除价格波动影响。
4. 使用的数据真实有效。

4 模型建立及求解

4.1 数据预处理

4.1.1 初始数据计算

根据问题分析中问题二的[变量计算方法](#)，我们通过 Excel 处理[附件 1](#)中数据，得到、整理出了 2000 ~ 2016 年，碳排放量及其八个变量的具体值。（数据如表 4 所示，具体文件参见 [附件 2](#)）

表 4: 2000-2016 年各变量具体值

年份	碳排放量	城镇化率	GDP	产业结构	总人口数	经济发展水平	国际贸易	人均碳排放量	能源消费强度
2000	95528.51053	0.362197518	100280.1	0.455372502	126743	0.791208193	0.205763656	0.753718237	1.465535036
2001	99939.25859	0.376597428	110863.2	0.447945757	127627	0.868650051	0.198659249	0.783057336	1.403053493
2002	109314.6759	0.390897838	121717.4	0.444517382	128453	0.9475637	0.221398091	0.851009131	1.393202615
2003	128517.4696	0.405302298	137422	0.456239903	129227	1.06341554	0.264061795	0.994509426	1.43414446
2004	149897.5482	0.417600086	161840.1	0.459014175	129988	1.245038773	0.30340441	1.153164509	1.422892102
2005	171351.2686	0.429899966	187318.9	0.470237654	130756	1.432583591	0.334445697	1.310465819	1.395315689
2006	187685.8448	0.443430102	219438.5	0.475585642	131448	1.669393981	0.353617073	1.427833401	1.305454603
2007	203788.9583	0.458892446	270232.3	0.468610155	132129	2.04521566	0.346233962	1.542348449	1.152497314
2008	207400.2101	0.469895032	319515.6	0.469324815	132802	2.40595473	0.314209827	1.56172505	1.003428315
2009	217249.6957	0.48341701	349081.4	0.458837681	133450	2.615821656	0.234987295	1.627948263	0.962887166
2010	229528.7214	0.499496611	413030.4	0.463960522	134091	3.080224624	0.25911614	1.711738456	0.873175437
2011	248898.1417	0.512702713	489300.5	0.464006883	134735	3.631576799	0.251870987	1.847316152	0.791012885
2012	254319.7118	0.525700866	540367.5	0.452735037	135404	3.990779445	0.239390785	1.878228943	0.744193535
2013	261402.5332	0.537296431	595244.5	0.440081513	136072	4.374481892	0.230378273	1.921060418	0.700406304
2014	262747.8929	0.547703645	643973.9	0.431029581	136782	4.708031027	0.22343095	1.920924485	0.661216239
2015	261805.7824	0.560998676	689052.1	0.409316364	137462	5.012673321	0.20487101	1.904568407	0.623907829
2016	260834.8387	0.573496973	744127.2	0.398098605	138271	5.381657759	0.186015634	1.886403068	0.585678094

4.1.2 数据分类

之后我们把数据分类，分别作为 BP 神经网络的训练数据集、检验数据集和最后的预测数据集。详细如表 5。

4.1.3 数据归一化处理

此外，为去除数据量纲，方便数据处理，并使 BP 神经网络收敛加快，我们将数据集进行了归一化处理。

表 5: 数据分类

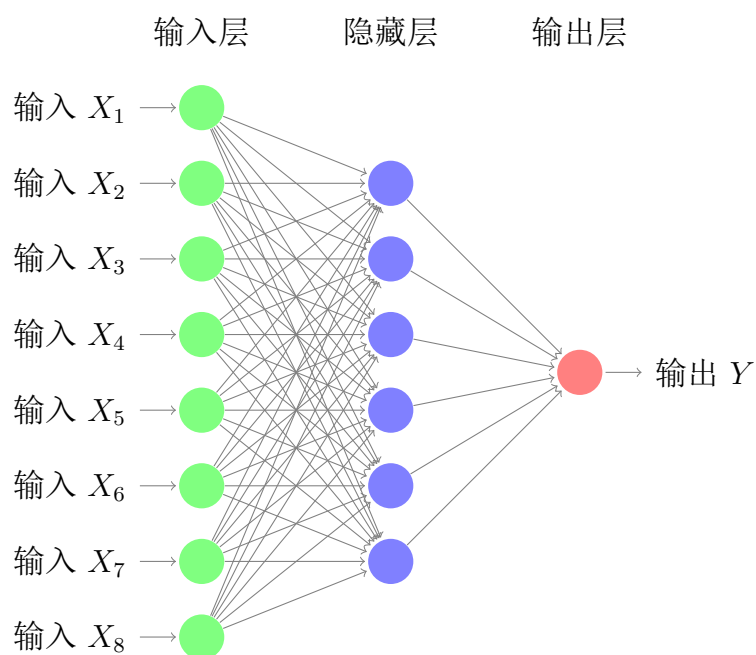
数据集	年份
训练数据集	2000–2013（包括 2013）
检验数据集	2014–2015（包括 2015）
预测数据集	2016

4.2 BP 神经网络结构构建

4.2.1 结构展示

最终网络结构示意图如图 2 所示：

图 2: 神经网络结构图



4.2.2 设计网络输入层和输出层

首先设计网络输入、输出层。具体如表 6 所示：

表 6: BP神经网络输入、输出层

输入层								输出层
X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	Y

4.2.3 选取隐含层节点数

隐含层节点数的选取是决定神经网络训练精度的关键，过多过少都会有很大影响。

隐含层节点过多，能有效减少系统误差，但也会导致诸如：网络训练时间延长、训练容易陷入局部极小点的问题。从而降低网络的容错性和泛化能力；而隐含层节点数过少，则又可能造成网络性能差或者网络根本不能被训练的问题。

因此隐含层节点数的选取，我们参考 Kolmogorov 定理，并采取下列经验公式：

$$J = \sqrt{m + n} + n \quad \text{其中 } m \text{ 为输入节点数, } n \text{ 为输出节点数, } a \text{ 的取值范围为 } 1-10$$

我们最终选取 6 为隐含层节点数。

4.3 BP 神经网络函数、参数设定

具体设定及初始化参见表 7 所示：

表 7: 函数设定及参数初始化

函数	设定
隐层激励函数	tansig
输出层激励函数	logsig
网络训练函数	traingdx
网络性能函数	mes
参数	初始化
期望误差最小值 err - goal	0.0000001
最大循环次数 max - epoch	5000
修正权值的学习速率 lr	0.01

4.3.1 选取激励函数

BP 神经网络通常采用 Sigmoid 可微函数和线性函数作为网络的激励函数。

我们选择 S 型正切函数 tansig 作为隐层神经元的激励函数。

而由于网络的输出归一到 $[-1, 1]$ 范围内，因此预测模型选取 S 型对数函数 logsig 作为输出层神经元的激励函数

4.3.2 选取学习速率

学习速率对 BP 神经网络具有重要影响作用。

学习速率太小，网络学习缓慢，需要增加训练次数；学习速率太快，容易导致网络不收敛，影响训练的精度。

我们最终选取学习速率为 0.01，最大循环次数为 5000 次。

5 模型评价与改进

6 政策建议


参考文献

索引

温室效应, [1](#)

A 附件

A.1 附件 1

中国近些年来的一些经济数据¹。 

A.2 附件 2

第一次数据处理：得到 2000 – 2016 年，八个影响因素的具体值。 

¹来源：中国统计年鉴及中国能源统计年鉴