

基于遗传算法优化反向传播神经网络的中国 二氧化碳排放量影响因素研究

田亚亚

(华北电力大学 经济管理系, 河北 保定 071003)

摘要:随着经济的发展与社会的进步,温室效应问题的愈加严重已经威胁到人类的正常生存与社会经济的可持续发展。为了能够更加准确的预测我国未来几年二氧化碳排放量,从而采取相应的节能减排政策,采用遗传算法(GA)优化反向传播神经网络(BPNN)的初始连接权值和阈值,克服了反向传播神经网络极易收敛于局部极小的缺点。通过我国1985—2015年的碳排放量和影响因素进行实证分析,结果证明GA优化过的BPNN模型(GA-BPNN)对二氧化碳排放量的预测有着更高的精度,更适用于我国当前二氧化碳排放量预测的需要。

关键词:二氧化碳排放量;遗传算法;BPNN;GA-BPNN

中图分类号:X-1 **文献标志码:**A **文章编号:**1671-1807(2018)01-0068-09

随着经济的进步,环境问题已经越来越成为制约经济发展的重要影响因素,温室气体作为主要的环境污染原因不可忽视。因此研究影响二氧化碳排放的影响因素,对二氧化碳排放进行更为精准的预测被国内外学者提上日程。国家“十三五”规划期间,强调要加快改善生态环境,推进资源节约集约再利用,健全生态安全保障机制。聚焦于实现单位国内生产总值能源消耗和二氧化碳排放大幅下降,减少主要污染物排放总量。因此,对影响二氧化碳排放的因素进行深入研究,并且建立更加精确的预测模型对于我国制定相应的节能减排政策有着重要的意义。目前国内外有很多对碳排放的影响因素进行分析及对碳排放量预测的方法,主要包括定性和定量分析。定性分析主要有:丁浩等^[1]运用以分维为基础的分形理论预测了我国2014—2016年的碳排放量,得出碳排放量将会持续上升的结论并提出了四点对策。杜强等^[2]运用K-均值聚类分析法将中国各省区碳排放分成5类,并构建了碳排放增长的Logistic预测模型,对全国碳排放放在省级层面上进行区分,提高了预测的准确性,明确了各省应承担的减排责任。赵息等^[3]应用DDEPM方法,借助Matlab软件,由1980—2009年碳排放数据计算出2020年碳排放量,并基于GDP数据计算

出2020年中国GDP数值。Rigoberto Pérez-Suárez^[4]等人基于扩展的环境库兹涅茨曲线和环境Logistic曲线对175个国家为样本进行研究,并对选择出的二次抽样样本未来几年的二氧化碳排放量做出预测,

得到很好的预测精度。Wang zhengxin^[5]等人采用非线性灰色多元模型分阶段预测了中国化石能源消费的二氧化碳排放量,证实了非线性灰色多元模型比传统的灰色模型(GM)和自回归移动平均模型(ARIMA)有着更高的预测精度。然而依赖于数值精确度的分析方法对信息资料和要求较高的,在选择影响二氧化碳排放量的因素方面又存在很多的不确定性,近年来很多研究者纷纷转入拥有较大灵活性的人工智能算法研究领域,但是人工智能算法对碳排放的研究相对较少,宋杰鲲等^[6]借鉴STIRPAT模型的影响因素选取了6项因素为自变量,用BP神经网络方法构建模型预测了2010—2015年我国碳排放,得出我国在“十二五”期间应降低GDP增长速度,实现低排放目标。Wen L等^[7]采用非惯性权重系数和粒子变异的粒子群算法去优化传统的BP神经网络的初始权值和阈值,通过对北京市1978—2012年二氧化碳排放量及影响因素的分析预测,表明IPSO-

收稿日期:2017-09-16

作者简介:田亚亚(1994—),女,河北邯郸人,华北电力大学经济管理系,技术经济及管理硕士研究生,研究方向:技术经济及管理理论。

BP 能够很好地提高粒子群算法的全局寻优能力。宋杰鲲^[8]选取人口、城镇化率、人均 GDP、服务业增加值比重、单位 GDP 能耗和煤炭消费比例 6 项因素作为自变量和 1980—2009 年碳排放数据作为样本,运用支持向量回归机方法构建中国碳排放预测模型,得到很好的效果。Sun Wei 等^[9]建立了改进的粒子群算法优化 BP 神经网络的模型,用河北省 1978—2012 年的二氧化碳排放量对网络进行训练与预测,结果表明,改进后的算法有效地克服了 BPNN 随机选择初始权值与阈值的缺点,有更好的预测效果。人工智能算法由于拥有较大的灵活性,已经越来越广泛的被各界学者使用,但是目前在碳排放领域内采用人工智能算法进行预测并不多,因此本文提出基于遗传优化算法的 BP 神经网络算法,首先通过遗传优化算法选择出 BP 神经网络的最优权值和阈值,然后用建立的 GA-BPNN 模型对我国 1980—2015 年的碳排放量与影响因素进行实证分析。

1 关于 GA-BPNN 模型

1.1 BPNN

BP(Back Propagation)网络是 1986 年由 Rumelhart 和 McClelland 为首的科学家小组提出,是一种按误差逆传播算法训练的多层前馈网络,是目前应用最广泛的神经网络模型之一,其典型的拓扑结构图如图 1 所示,结构包括输入层(input layer)、隐含层(hidden layer)和输出层(output layer),各层神经元之间相互连接,但是每一层之间的神经元之间没有连接关系,模型的输入矢量设定为 $X_1, X_2, \dots, X_n$,输出矢量设定为 $Y_1, Y_2, \dots, Y_m$,连接权值和阈值设定为 W_{ij}, W_{jk} 。 n 和 m 分别为输入层和输出层的节点数目,代表 n 个独立变量到 m 个独立变量的映射。BP 网络能学习和存贮大量的输入-输出模式映射关系,而无需事前揭示描述这种映射关系的数学方程。采用 BP 神经网络预测前首先要对网络进行训练,使其具有联想与预测能力,训练过程包括以下几个步骤:

1) 网络初始化。根据输入输出的序列 (X, Y) 确定网络输入层。隐含层和输出层的节点个数,初始化每一层的连接权值和隐含层和输出层的阈值 a, b ,给定学习速率和神经元激励函数。本文所选的隐含层激励函数为

$$f(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

2) 隐含层输出的计算。隐含层的输出 H 由输入变量 x , 连接权值 w_{ij} 和阈值 a 计算得来,公式如下:

$$H_j = f\left(\sum_{i=1}^n w_{ij}x_i - a_j\right), j = 1, 2, \dots, l \quad (2)$$

式中 l 为隐含层节点数, f 为隐含层激励函数。

3) 输出层输出值的计算。输出层的输出 O 由隐含层输出 H , 连接权值 w_{ij} 和阈值 b 计算,公式如下:

$$O_k = \sum_{j=1}^l H_j w_{kj} - b_k, k = 1, 2, \dots, m \quad (3)$$

4) 误差计算。网络预测误差为 e 。

$$e_k = Y_k - O_k, k = 1, 2, \dots, m \quad (4)$$

5) 根据预测误差 e 进行权值更新。

$$w_{ij} = w_{ij} + \eta H_j (1 - H_j) x(i) \sum_{k=1}^m w_{kj} e_k, i = 1, 2, \dots, n; j = 1, 2, \dots, l \quad (5)$$

$$w_{ij} = w_{ij} + \eta H_j e_k, j = 1, 2, \dots, l; k = 1, 2, \dots, m \quad (6)$$

式中 η 为学习速率。

6) 根据预测误差 e 进行阈值更新。

$$a_j = a_j + \eta H_j (1 - H_j) \sum_{k=1}^m w_{kj} e_k, j = 1, 2, \dots, l \quad (7)$$

$$b_k = b_k + e_k, k = 1, 2, \dots, m \quad (8)$$

7) 判断算法迭代是否结束,若没结束,返回(2)。

它的学习过程由正向传播和反向传播两个部分构成。在正向传播过程中,信息由输入层经过各中间隐层加权处理向输出层传播,于输出层获得网络的输入响应,输出值通过函数计算与目标输出值比较,若存在误差,误差则反向传播沿之前的路线返回,即从输出层经过各中间隐层逐层调整其连接权值,最终回到输出层,以减小误差,随着神经网络的误差逆向传播不断地对权值进行修正,整个网络对输入信息响应的准确率自然也有所提升。BPNN 的算法流程如图 2 所示。

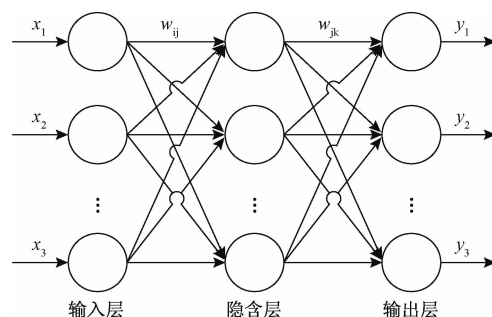


图 1 BPNN 拓扑结构图

但是,因为传统 BP 算法初始连接权值和阈值是

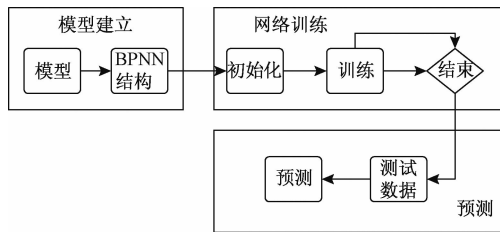


图 2 BPNN 算法流程图

随机设定的,所以如果初值选取不当会直接影响网络的收敛性能和最终的稳定状态,网络学习容易陷入局部极小。而且在 BP 神经网络学习的过程中,其主要更新的是网络的连接权值和阈值,对于传统的 BP 网络是通过不断的训练才能找到合适的值。然而,过度的训练或缺乏训练都得不到理想的结果,其间还会造成“过拟合”等现象的发生,从而影响网络的泛化能力。因此,本文针对以上问题,提出先利用遗传算法来优化 BP 神经网络的连接权值和阈值,可以充分发挥遗传算法的全局寻优能力和 BP 算法的局部搜索优势,较好地克服 BP 神经网络权值随机的问题,提高 BP 神经网络的泛化能力和神经网络的学习能力。

1.2 遗传算法(GA)

遗传算法(Genetic Algorithms)是由美国 Michigan 大学的 Holland 教授于 1975 年在他的专著《自然界和人工系统的适应性》^[10]中首先提出的,它是一类借鉴生物界自然选择和自然遗传机制的随机搜索算法。遗传算法在优化参数形成的串联体中引入“优胜劣汰,适者生存”的生物进化原理,按照所选择的适应度函数并通过遗传中的选择、交叉、变异对个体进行筛选,适应度值好的个体将被保留下来,这样往复循环,直到满足条件。因此在应用遗传算法时,选择合适的适应度函数至关重要,本文应用遗传算法优化 BPNN 的初始权值和阈值,因此将预测输出与期望输出间的误差绝对值和 E 作为个体适应度值 F ,

计算公式为: $F = k(\sum_{i=1}^n abs(y_i - O_i))$, 式中 n 为网络节点数; y_i 和 o_i 分别代表第 i 个节点的期望输出和预测输出, k 为系数。接下来是遗传算法的选择、交叉和变异操作。选择操作有多种方法,本案例选择基于适应度比例的轮盘赌法,由于适应度值越小越好,因此在个体选择前对适应度值求倒数,则每个个体 i 的选择概率为

$$p_i = \frac{f_i}{\sum_{j=1}^N f_j}, f_i = k/F_i \quad (9)$$

式中 F_i 为个体 i 的适应度值, N 为种群个体数, k 为系数。交叉操作选择实数交叉法,第 k 个染色体 a_k 和第 l 个染色体 a_l 在 j 位的交叉操作为:

$$a_{kj} = a_{kj}(1-b) + a_{lj}b, a_{lj} = a_{lj}(1-b) + a_{kj}b \quad (10)$$

b 为 $[0,1]$ 间的随机数。对于 BP 神经网络初始权值和阈值的优化,变异选用非均匀变异算子,选取第 i 个个体的第 j 个基因 a_{ij} 进行变异,方法如下:

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{\max})f(g)r > 0.5 \\ a_{ij} + (a_{\min} - a_{ij})f(g)r \leq 0.5 \end{cases} \quad (11)$$

$$f(g) = r(1 - g/G_{\max})^2 \quad (12)$$

式中, $a_{\max}$ 和 $a_{\min}$ 分别为基因 a_{ij} 的上界和下界, r_2 是一个随机数值, g 为当前迭代次数, $G_{\max}$ 为最大进化次数, r 为 $[0,1]$ 之间的随机数。通过以上选择、交叉和变异的一次次的迭代进化,当满足迭代终止条件时,遗传算法输出的末代种群对应的就是最优解或近优解,至此筛选出 BPNN 的最优权值和阈值。GA 优化 BP 神经网络的算法流程图如图 3 所示。

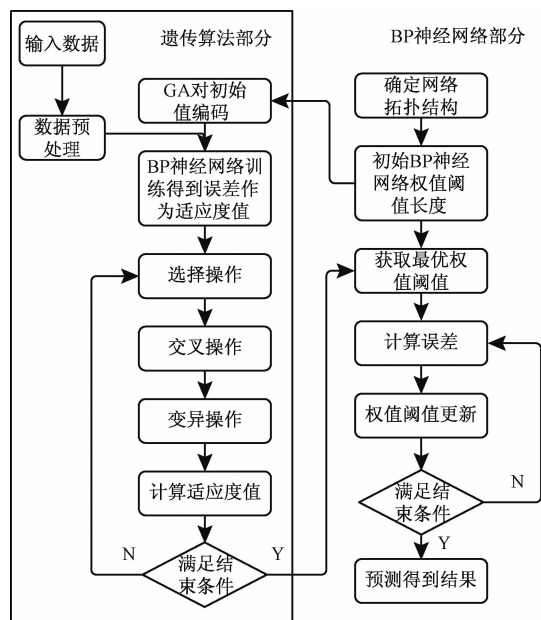


图 3 遗传算法优化 BP 神经网络的算法流程

2 数据分析

2.1 数据来源与换算

本文选用我国 1980 年至 2015 年的能源消费数据和相关数据进行分析,所选数据时间上限是目前最新的统计年鉴公布的最后年份。由于中国统计年鉴没有历年二氧化碳排放量的直接数据,因此需要通过我国主要能源消费数据与不同能源的碳排放系数将数据折算为二氧化碳排放量数据,首先

从《中国统计年鉴 2016》^[11] 得到主要能源消费数据,见表 1。然后通过不同能源二氧化碳排放系数将能源消费数据折算为二氧化碳排放量如图 3,其中碳排放系数见表 2。

表 1 1980 年至 2015 年中国主要能源消费

年份	能源消费总量(万吨标准煤)	煤炭	石油	天然气	一次电力
1980	60 275.00	43 518.55	12 476.93	1 868.53	2 411.00
1981	61 364.00	44 427.54	14 481.90	1 656.83	797.73
1982	64 686.00	47 544.21	14 619.04	1 617.15	905.60
1983	68 877.00	51 037.86	15 152.94	1 653.05	1 033.16
1984	75 493.00	56 393.27	16 381.98	1 660.85	1 056.90
1985	76 682.00	58 124.96	13 112.62	1 687.00	3 757.42
1986	80 850.00	61 284.30	13 906.20	1 859.50	3 799.95
1987	86 632.00	66 013.58	14 727.44	1 819.27	4 071.70
1988	92 997.00	70 863.71	15 809.49	1 952.94	4 370.86
1989	96 934.00	73 669.84	16 575.71	1 938.68	4 749.77
1990	98 703.00	75 211.69	16 384.70	2 072.76	5 033.85
1991	103 783.00	78 978.86	17 746.89	2 075.66	4 981.58
1992	109 170.00	82 641.69	19 104.75	2 074.23	5 349.33
1993	115 993.00	86 646.77	21 110.73	2 203.87	6 031.64
1994	122 737.00	92 052.75	21 356.24	2 332.00	6 996.01
1995	131 176.00	97 857.30	22 955.80	2 361.17	8 001.74
1996	135 192.00	99 366.12	25 280.90	2 433.46	8 111.52
1997	135 909.00	97 039.03	27 725.44	2 446.36	8 698.18
1998	136 184.00	96 554.46	28 326.27	2 451.31	8 851.96
1999	140 569.00	99 241.71	30 222.34	2 811.38	8 293.57
2000	146 964.00	100 670.34	32 332.08	3 233.21	10 728.37
2001	155 547.00	105 771.96	32 975.96	3 733.13	13 065.95
2002	169 577.00	116 160.25	35 611.17	3 900.27	13 905.31
2003	197 083.00	138 352.27	39 613.68	4 532.91	14 584.14
2004	230 281.00	161 657.26	45 825.92	5 296.46	17 501.36
2005	261 369.00	189 231.16	46 523.68	6 272.86	19 341.31
2006	286 467.00	207 402.11	50 131.73	7 734.61	21 198.56
2007	311 442.00	225 795.45	52 945.14	9 343.26	23 358.15
2008	320 611.00	229 236.87	53 542.04	10 900.77	26 931.32
2009	336 126.00	240 666.22	55 124.66	11 764.41	28 570.71
2010	360 648.00	249 568.42	62 752.75	14 425.92	33 900.91
2011	387 043.00	271 704.19	65 023.22	17 803.98	32 511.61
2012	402 138.00	275 464.53	68 363.46	19 302.62	39 007.39
2013	416 913.00	280 999.36	71 292.12	22 096.39	42 525.13
2014	425 806.00	279 328.74	74 090.24	24 270.94	48 116.08
2015	430 000.00	275 200.00	77 830.00	25 370.00	51 600.00

表 2 不同能源的碳排放系数

种类	煤炭	石油	天然气	一次电力
二氧化碳排放/(t/t)	0.747 6	0.582 5	0.443 5	0

表 3 1980 年至 2015 年中二氧化碳排放量 万吨

年份	总二氧化碳排放量	年份	总二氧化碳排放量
1980	40 630.97	1998	89 771.32
1981	42 384.54	1999	93 044.46
1982	44 776.85	2000	95 528.51
1983	47 715.62	2001	99 939.26
1984	52 438.70	2002	109 314.68
1985	51 840.51	2003	128 517.47
1986	54 741.19	2004	149 897.55
1987	58 737.34	2005	171 351.27
1988	63 052.87	2006	187 685.84
1989	65 590.73	2007	203 788.96
1990	66 691.61	2008	207 400.21
1991	70 302.72	2009	217 249.70
1992	73 831.37	2010	229 528.72
1993	78 051.54	2011	248 898.14
1994	82 292.89	2012	254 319.71
1995	87 577.05	2013	261 402.53
1996	90 091.48	2014	262 747.89
1997	89 781.40	2015	262 327.09

图 4 表示 1980 年到 2015 年中国能源消费量，从图中可以看出，煤炭消耗量和能源消费总量基本呈同步增长趋势，可以看出煤炭消费是造成二氧化碳排放的首要因素，而一次电力和天然气的占比很小且增长相对缓慢。因此，对于我国二氧化碳排放量的控制，主要应从减少煤炭消费方面入手，可以大力推广一次电力能源消费，从而达到减少二氧化碳排放量的目的。由图 5 可以看出我国的二氧化碳排放量的趋势，在 2001 年中国加入世界贸易组织(WTO)后有了大幅上涨。但经济的增长不能以破坏环境作为代价，因此对二氧化碳排放量预测的研究被提上日程。

这篇文章首先对影响我国二氧化碳排放量的因素进行了初步筛选。预选的 7 个影响因素分别为：总人口，国民总收入，国内生产总值，煤炭消耗量，能源生产总量，第三产业 GDP 贡献率和石油消耗量。各影响因素数据值来源于中国统计年鉴 2016^[11]。数据如表 4 所示。

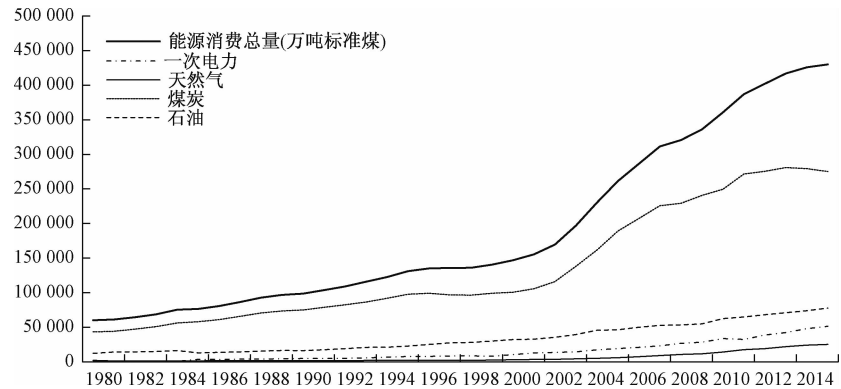


图 4 1980 年到 2015 年中国能源消费量

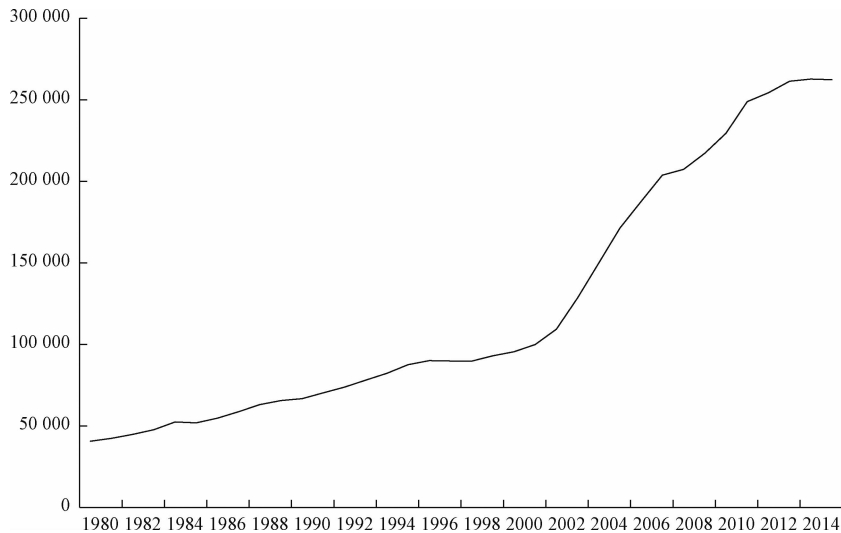


图 5 1980 年到 2015 年中国二氧化碳排放量趋势图

表 4 中国二氧化碳排放影响因素的数据

年份	总人口 (万人)	国民总收入 (亿元)	国内生产总值 (亿元)	煤炭消耗量 (万吨标准煤)	能源生产总量 (万吨标准煤)	第三产业 GDP 贡献率(%)	石油消耗量 (万吨标准煤)
1980	98 705.00	4 587.60	4 587.60	43 518.55	63 735.00	19.20	12 476.93
1981	100 072.00	4 933.70	4 935.80	44 427.54	63 227.00	41.80	14 481.90
1982	101 654.00	5 380.50	5 373.40	47 544.21	66 778.00	32.60	14 619.04
1983	103 008.00	6 043.80	6 020.90	51 037.86	71 270.00	32.70	15 152.94
1984	104 357.00	7 314.20	7 278.50	56 393.27	77 855.00	31.70	16 381.98
1985	105 851.00	9 123.60	9 098.90	58 124.96	85 546.00	34.80	13 112.62
1986	107 507.00	10 375.40	10 376.20	61 284.30	88 124.00	36.90	13 906.20
1987	109 300.00	12 166.60	12 174.60	66 013.58	91 266.00	34.80	14 727.44
1988	111 026.00	15 174.40	15 180.40	70 863.71	95 801.00	33.40	15 809.49
1989	112 704.00	17 188.40	17 179.70	73 669.84	101 639.00	40.10	16 575.71
1990	114 333.00	18 923.30	18 872.90	75 211.69	103 922.00	20.00	16 384.70
1991	115 823.00	22 050.30	22 005.60	78 978.86	104 844.00	32.20	17 746.89
1992	117 171.00	27 208.20	27 194.50	82 641.69	107 256.00	28.70	19 104.75
1993	118 517.00	35 599.20	35 673.20	86 646.77	111 059.00	28.00	21 110.73
1994	119 850.00	48 548.20	48 637.50	92 052.75	118 729.00	27.40	21 356.24
1995	121 121.00	60 356.60	61 339.90	97 857.30	129 034.00	28.50	22 955.80
1996	122 389.00	70 779.60	71 813.60	99 366.12	133 032.00	28.50	25 280.90
1997	123 626.00	78 802.90	79 715.00	97 039.03	133 460.00	34.50	27 725.44
1998	124 761.00	83 817.60	85 195.50	96 554.46	129 834.00	33.00	28 326.27
1999	125 786.00	89 366.50	90 564.40	99 241.71	131 935.00	37.40	30 222.34
2000	126 743.00	99 066.10	100 280.10	100 670.34	138 570.00	36.20	32 332.08
2001	127 627.00	109 276.20	110 863.10	105 771.96	147 425.00	49.00	32 975.96
2002	128 453.00	120 480.40	121 717.40	116 160.25	156 227.00	46.50	35 611.17
2003	129 227.00	136 576.30	137 422.00	138 352.27	178 299.00	39.00	39 613.68
2004	129 988.00	161 415.40	161 840.20	161 657.26	206 108.00	40.80	45 825.92
2005	130 756.00	185 998.90	187 318.90	189 231.16	229 037.00	44.30	46 523.68
2006	131 448.00	219 028.50	219 438.50	207 402.11	244 763.00	45.90	50 131.73
2007	132 129.00	270 844.00	270 232.30	225 795.45	264 173.00	47.30	52 945.14
2008	132 802.00	321 500.50	319 515.50	229 236.87	277 419.00	46.20	53 542.04
2009	133 450.00	348 389.50	349 081.40	240 666.22	286 092.00	43.70	55 124.66
2010	134 091.00	411 265.20	413 030.30	249 568.42	312 125.00	39.00	62 752.75
2011	134 735.00	484 753.20	489 300.60	271 704.19	340 178.00	43.80	65 023.22
2012	135 404.00	539 116.20	540 367.40	275 464.53	351 041.00	44.90	68 363.46
2013	136 072.00	590 422.40	595 244.40	280 999.36	358 784.00	47.20	71 292.12
2014	136 782.00	644 791.10	643 977.00	279 328.74	361 866.00	47.50	74 090.24
2015	137 462.00	682 635.10	685 505.80	275 200.00	362 000.00	53.70	77 830.00

2.2 SPSS 分析

用 SPSS 软件对 7 个预选影响因素分别与二氧化碳排放量进行双侧显著性检验并计算 Pearson 相关系数。分析结果如表 5 所示。可以看出,各因素与二氧化碳排

放的 Pearson 相关系数均在 0.7 以上,甚至煤炭消耗量的相关系数高达 0.999,且双侧显著性检验均为 0,表明 7 个影响因素与碳排放量呈极显著相关,可以作为影响我国二氧化碳排放量的因素进行进一步研究。

表 5 双侧显著性检验

因素	Pearson 系数	显著性检验(双侧)	因素	Pearson 系数	显著性检验(双侧)
总人口	0.873**	0.000	能源生产总量	0.999**	0.000
国民总收入	0.963**	0.000	第三产业 GDP 贡献率	0.724**	0.000
国内生产总值	0.962**	0.000	石油消耗量	0.990**	0.000
煤炭消耗量	0.999**	0.000			

注:**表示在 0.01 水平上(双侧)显著相关。

3 GA-BPNN 模型应用

3.1 参数设定

将 GA-BPNN 应用于本实例中,步骤如下:

Step1 本文使用经典的 3 层 BP 神经网络,根据实际情况及不同隐含层节点数的反复测试,设置 BP 神经网络的结构为 7-6-1,即输入层、隐含层和输出层节点数分别为 7、6、1。迭代次数为 100 次,误差精确度与学习率分别设置为 0.000 1 和 0.05。选取 1980 年至 2009 年的 7 个影响二氧化碳排放的因素的数据以及二氧化碳排放量作为训练集,以 2010 年至 2015 年相对应的数据作为测试集。

Step2 遗传算法的参数初始化,种群规模设置为 10,迭代次数、选择概率和变异概率分别设置为 10、0.3 和 0.15。

Step3 确定适应度函数,遗传算法中个体由神经网络全部权值和阈值编码而成,在本文中,采用训练数据训练 BP 神经网络后预测系统输出和期望输出之间的误差绝对值 E 作为个体适应度值 F ,计算公式为:

$$F = k \left(\sum_{i=1}^n \text{abs}(y_i - o_i) \right) \quad (13)$$

式中 n 为网络节点数; y_i 和 o_i 分别代表第 i 个节点的期望输出和预测输出, k 为系数。

Step4 选择操作中选用基于适应度比例的轮盘赌法,由于个体适应度值越小越好,所以在选择个体前对适应度值求倒数,即每个个体的选择概率为 $p_i = \frac{f_i}{\sum_{j=1}^N f_j}$, $f_i = k/F_i$ 式中 F_i 为个体 i 的适应度值, N 为种群个体数, k 为系数。

Step5 由于遗传算法中个体由神经网络全部权值和阈值实数编码而成,所以交叉操作选择实数交叉法,第 k 个染色体 a_k 和第 l 个染色体 a_l 在 j 位的交叉操作为:

$$a_{kj}(1-b) + a_{lj}b, a_{lj}(1-b) + a_{kj}b \quad (14)$$

b 为 $[0,1]$ 间的随机数。

Step6 变异选用非均匀变异算子,选取第 i 个个体的第 j 个基因 a_{ij} 进行变异,方法如下:

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{\min})f(g)r > 0.5 \\ a_{ij} + (a_{\max} - a_{ij})f(g)r \leq 0.5 \end{cases} \quad (15)$$

式中, $a_{\max}$ 和 $a_{\min}$ 分别为基因 a_{ij} 的上界和下界; $f(g) = r(1 - g/G_{\max})^2$, r_2 是一个随机数值, g 为当前迭代次数, $G_{\max}$ 为最大进化次数, r 为 $[0,1]$ 之间的随机数。

Step7 根据以上确立的适应度函数计算出个体适应度值,如果满足结束条件,则终止迭代,将该个体作为最优个体,即遗传算法选出的 BPNN 的最优权值和阈值。如果不满足条件,继续迭代循环,直到选出最优个体。

Step8 将遗传算法选择出的最优个体作为 BPNN 的初始权值和阈值,选取本文已选定的训练样本进行 BP 网络训练。

Step9 将训练好的网络应用于测试样本,得出测试结果并进行分析。

BPNN 和 GA 的初始化参数设置分别如表 6,表 7 所示。

表 6 BPNN 的参数设定

参数	值	参数	值
输入层节点数	7	迭代次数	100
隐含层节点数	6	误差精确度	0.000 1
输出层节点数	1	学习率	0.05

表 7 GA 的参数设定

参数	种群规模	迭代次数	选择概率	变异概率
值	10	10	0.3	0.15

3.2 预测结果分析

本文运用 Windows10 系统下的 MATLAB2016a 软件进行网络训练与预测,GA 经过迭代寻优后选择的神经网络的最优连接权值和阈值列于表 8,并将 GA-BPNN 算法与 BP 算法的训练及预测准确度进行了比较,两个算法经过程序训练及预测后得到的预测值列于表 9,且二者算法预测值误差对比如表 10 所示,本文选取了 MAPE, MaxAPE, MdAPE 和误差总和作为判断两个模型的预测误差指标,可以看出单纯的 BPNN 预测的平均绝对百分误差值为 12.25,而用 GA 优化过的 BPNN 的预测平均绝对百分误差值降为 1.05,累计误差也由之前的 1 326.7 降低为 42.4,最大相对误差与中位数百分误差都有大幅度的降低。两个模型的预测结果对比图如图 6 所示。可以看出,经过遗传算法优化过的 GA-BPNN 算法模型的拟合效果远远优于传统的 BPNN。表明遗传算法很好地克服了 BP 神经网络权值阈值选择的随机性和容易陷入局部极小值的问题,使得 GA-BPNN 有较好的预测效果。

表 8 BPNN 最优参数

参数	值
连接权值	2.172 134,1.151 548,-0.808 33,-2.186 78,0.926 692,-1.345 52,-1.192 86,1.072 462,-2.922 91,-0.837 39,2.340 291,2.697 898,2.978 582,-0.024 83,0.595 123,-0.136 94,-1.871 41,-0.217 5,0.603 124,-1.018 7,-2.247 48,2.753 925,0.823 069,2.281 815,2.490 428,-1.522 7,-2.646 92,-0.419 7,-0.325 01,2.273 832
连接权值	-1.403 86,-0.727 48,2.787 286,1.928 013,0.473 135
隐含层阈值	1.218 973,1.166 466,-0.892 29,-1.734 28,2.326 207
输出层阈值	2.973 866

表 9 模型预测结果与真实值对比

年份	2010	2011	2012	2013	2014	2015
真实值	229 528.72	248 898.14	254 319.71	261 402.53	262 747.89	262 327.09
BPNN	198 697.60	203 115.83	209 397.89	230 951.65	241 088.40	251 459.15
GA-BPNN	229 526.14	239 960.50	251 531.18	262 151.56	259 953.22	261 597.32

表 10 BPNN、GA-BP 算法预测值误差对比

算法	MAPE (%)	MaxAPE (%)	MdAPE (%)	Errorsum
BPNN	12.25	18.39	12.5	1 326.7
GA-BPNN	1.05	3.59	0.67	1 042.4

参考文献

[1] 丁浩, 鲍雨. 基于分形理论的我国碳排放量预测研究[J]. 河南科学, 2014, 32(9): 1815—1820.

[2] 杜强, 陈乔, 杨锐. 基于 Logistic 模型的中国各省碳排放预测[J]. 长江流域资源与环境, 2013, 2(2): 143—150.

[3] 赵息, 齐建民, 刘广为. 基于离散二阶差分算法的中国碳排放预测[J]. 干旱区资源与环境, 2013, 27(1): 63—69.

[4] RIGOBERTO PÉREZ-SUÁREZ, ANA JESÚS LÓPEZ-MENÉNDEZ. Growing green? forecasting CO₂ emissions with environmental kuznets curves and logistic growth models[J]. Environmental Science & Policy, 2015, 54: 428—437.

[5] ZHENG-XIN WANG, DE-JUN YE. Forecasting Chinese carbon emissions from fossil energy consumption using non-linear grey multivariable models[J]. Journal of Cleaner Production, 2017, 142: 600—612.

[6] 宋杰鲲, 张宇. 基于 BP 神经网络的我国碳排放情景预测[J]. 科学技术与工程, 2011, 11(17): 4108—4111.

[7] WEN LEI, LIU YANJUN. A research about Beijing's carbon emissions based on the IPSO-BP model[J]. Environmental Progress & Sustainable Energy, 2016, 36(2): 428—434.

[8] 宋杰鲲. 基于支持向量回归机的中国碳排放预测模型[J]. 中国石油大学学报, 2012, 36(1): 182—187.

[9] WEI SUN, YANFENG XU. Using a back propagation neural network based on improved particleswarm optimization to study the influential factors of carbon dioxide emissions in Hebei province, China[J]. Journal of Cleaner Production, 2016, 112: 1282—1291.

[10] HOLLAND J H. Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence[M]. 2nd ed. Cambridge: MIT Press, 1992.

[11] 国家统计局. 中国统计年鉴 2016[J]. 北京: 中国统计出版社, 2016.

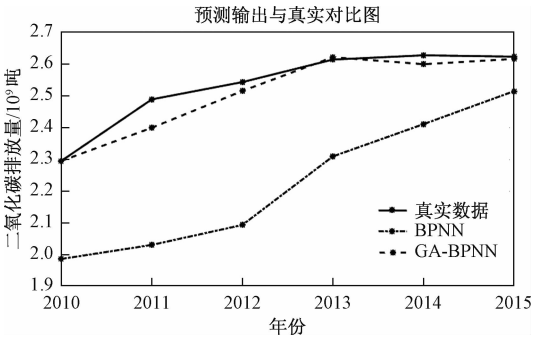


图 6 预测值与真实值对比图

4 结论

本文引进遗传算法(GA)优化 BP 神经网络初始连接权值和阈值,充分发挥了遗传算法的全局寻优能力和 BP 算法的局部搜索优势,可以很好地克服 BP 神经网络权值阈值选择的随机性和容易陷入局部极小值的问题。然后选取了我国 1980—2015 年的二氧化碳排放量与影响因素进行实证分析,采用算法 GA-BPNN 进行了模型的验证,通过与单纯的 BPNN 网络拟合预测结果的对比分析,证实了 GA-BPNN 有着更好的精度,可应用于我国二氧化碳排放量的预测研究。

Using a Back Propagation Neural Network Based on Genetic Algorithm to Study Influence Factors of Carbon Dioxide Emissions in China

TIAN Ya-ya

(North China Electric Power University, Baoding Hebei 071003, China)

Abstract: With the development of the economic, the problem of the global greenhouse effect has been attached more importance than before. In order to predict the emission of the carbon dioxide in the coming years more precisely and take corresponding energy conservation and emission reduction measures in time. In this paper, the genetic algorithm (GA) is used to optimize the initial connection weights and thresholds of the tradition Back Propagation Neural Network (BPNN) named GA-BPNN model, which can overcome BPNN's shortcoming of converging to the local minimum obviously. The data of China during the period 1985—2015 including carbon emission and influence factors are selected to perform the carbon dioxide emission prediction with the established model. The results indicates that the GA-BPNN model established in this paper has a higher accuracy, which is more applicable to the current prediction of carbon emissions.

Key words: carbon dioxide emissions; genetic algorithm; back propagation neural network; GA-BPNN

(上接第22页)

的需求。高耗能产业群作为“三高”产业,在技术创新过程中,不仅要注重人财物的投入量,还要注重投入结构,保证资源利用最大化。创新环境机制是保证网络持续运行的基础。因此,应建立健全创新环境机制,促进高耗能产业群协同创新网络创新水平的持续提升。

第三,充分发挥政府作用,实现创新治理的协同效应。政府角色在高耗能产业群协同创新过程中具有关键性作用。政府不仅可以在人财物和创新环境供给等方面为高耗能产业群发展奠定基础,而且可以根据国家发展重点需要和行业需求,有目的、有针对性地实施政策引导,有效集聚创新资源,加速提高高耗能产业群协同创新网络创新水平。

参考文献

- [1] 李克杰. 产业集群区域创新网络研究[D]. 广州:暨南大学,2006.
- [2] 郑季良,陈卫萍. 论高耗能产业群在循环经济中的协同发展[J]. 科技进步与对策,2010,27(2):53—56.
- [3] SCHIUMA G, LERRO A. Knowledge-based capital in building regional innovation capacity[J]. Journal of Knowledge Management, 2008, 12(5):121—136.

- [4] 刘汉文. 科技园区协同创新网络演化机制研究[D]. 北京:中国矿业大学,2015.
- [5] 黄海霞,陈劲. 创新生态系统的协同创新网络模式[J]. 技术经济,2016,35(8):31—37.
- [6] 吴荻. 产业集群绿色协同创新网络的运作模式研究[J]. 经济研究导刊,2016(19):24—26.
- [7] 全利平,蒋晓阳. 协同创新网络组织实现创新协同的路径选择[J]. 科技进步与对策,2011,28(9):15—19.
- [8] 王海峰. 协同演化视角下环境技术创新与环境治理制度耦合机制研究[J]. 系统科学学报,2014(4):49—52.
- [9] 韩周,秦远建,王蓼祥. 中国企业协同创新网络治理研究[J]. 科学管理研究,2016(1):75—78.
- [10] 张运生,邹思明. 高科技企业创新生态系统治理机制研究[J]. 科学学研究,2010,28(5):785—792.
- [11] 郑季良. 高耗能产业群循环经济协同发展模式研究[M]. 北京:科学出版社,2015:116—120.
- [12] 陈盼. 基于序参量的高耗能产业群循环经济协同发展评价研究[D]. 昆明:昆明理工大学,2013.
- [13] 李晓钟,王莹. 我国物联网产业协同发展机制及系统协同度评价研究[J]. 软科学,2015(1):42—46.
- [14] 郑季良,周旋. 高耗能企业科技创新与节能减排协调发展评价实证研究——基于宝钢2005—2014年可持续发展报告[J]. 科技与经济,2016,29(3):19—23.

Study on Synergy Evaluation of Collaborative Innovation Networks in Energy Intensive Industries

ZHENG Ji-liang, WANG Shao-fang

(Faculty of Management & Economics, Kunming University of Science & Technology, Kunming 650093, China)

Abstract: Under the new situation, traditional high energy-consuming industries are facing the problem of overcapacity, and it is urgent to upgrade the industrial structure through collaborative innovation networks. Based on technological innovation mechanism, innovation environment mechanism and innovation management mechanism, builds up structural model and evaluation index system of collaborative innovation networks operation mechanism in energy intensive industries, and use the synergetic degree model to analyze its collaborative process. Through an empirical study of the operation mechanism of collaborative innovation networks in energy intensive industries from 2008 to 2016 in Shenhua enterprise, the article made conclusions that the three mechanisms influence each other and promote each other, promote the innovative ability of collaborative innovation networks in energy intensive industries by synergy.

Key words: energy intensive industries; collaborative innovation networks; operation mechanism; synergetic degree