

## 多尺度分形特征下碳排放权价格预测算法

杨 星<sup>1,2</sup>, 梁敬丽<sup>1</sup>, 蒋金良<sup>1</sup>, 米君龙<sup>2†</sup>

(1. 华南理工大学广州学院 经济学院, 广东 广州 510800; 2. 暨南大学 经济学院 金融系, 广东 广州 510630)

**摘要:** 在非线性的范式下, 本文构建了基于多贝西小波三层变换和单支重构的遗传算法径向基函数神经网络模型 (daubechies wavelet-genetic algorithm-radial basis function neural network model, Db3-GA-RBF), 探讨了欧盟碳排放权市场的价格预测问题. 研究表明: 1) 欧盟碳排放权交易市场配额三阶段的现货价格波动均具有局部尺度多样性特征, 且第3阶段碳价格序列多重分形特征最强, 本质上碳排放权市场是一个多重分形与混沌市场; 2) Db3-GA-RBF模型能有效地提高数据的准确性和模型的泛化能力, 使模型的预测精度更强; 3) 与其他预测模型效果相比, 基于施瓦茨信息准则(Schwartz's information criterion, SIC)的Db3-GA-RBF(SIC)模型的预测精度大约提高70%.

**关键词:** 欧盟碳排放权市场; 分形与混沌; 小波变换; 径向基函数网络; 预测分析

**引用格式:** 杨星, 梁敬丽, 蒋金良, 等. 多尺度分形特征下碳排放权价格预测算法. 控制理论与应用, 2018, 35(2): 224 - 231

中图分类号: F831.5 文献标识码: A

## Price forecasting algorithm of carbon emission rights under multiscale fractal characteristics

YANG Xing<sup>1,2</sup>, LIANG Jing-li<sup>1</sup>, JIANG Jin-liang<sup>1</sup>, MI Jun-long<sup>2†</sup>

(1. Economic School, Guangzhou College of South China University of Technology, Guangzhou Guangdong 510800, China;  
2. Department of Finance, College of Economics, Jinan University, Guangzhou Guangdong 510630, China )

**Abstract:** In this paper, the daubechies wavelet-genetic algorithm-radial basis function neural network (Db3-GA-RBF) model is constructed by nonlinear paradigm, and the price forecasting problem of European Union carbon emissions market (EU-ETS) is discussed. The research showed that: 1) The European Union allowance (EUA) spot price fluctuation in the three stage of the EU carbon emissions market has the characteristics of local scale diversity, and the third stage carbon price series has the strongest multifractal characteristics. Essentially, the carbon emission rights market is a multifractal and chaotic market; 2) The Db3-GA-RBF model can effectively improve the accuracy of data and the generalization ability of the model, and make the prediction accuracy of the model stronger; 3) Compared with other forecasting models, the prediction accuracy of Db3-GA-RBF (SIC) model based on Schwartz's information criterion (SIC) is improved by about 70%.

**Key words:** EU carbon emissions market; fractal and chaos; wavelet decomposition; radial basis function networks; predictive analytics

**Citation:** YANG Xing, LIANG Jiangli, JIANG Jinliang, et al. Price forecasting algorithm of carbon emission rights under multiscale fractal characteristics. *Control Theory & Applications*, 2018, 35(2): 224 - 231

### 1 引言(Introduction)

尽管关于金融资产价格的预测研究已经进行了大量有价值的工作, 但对金融市场能否预测及预测效果如何至今仍是一个极有争议的课题. 在传统的有效市场理论看来, 金融市场的价格波动是无法预测的, 原因是金融市场的价格遵循随机游走过程, 金融资产的收益率相互独立, 历史价格对今天的价格并不产生影

响. 其研究基础是基于线性的、完全理性的均衡范式. 但主流金融理论却很少关注到标度不变性在金融市场上的动力学行为. 事实上, 金融市场并非是一个简单的、有秩序的线性系统, 本质上, 它是一个耗散的、复杂的、非线性系统, 它以非线性的方式与外界发生作用.

作为一个新兴的金融市场, 碳排放权市场是一个

收稿日期: 2017-04-28; 录用日期: 2017-12-12.

†通信作者. E-mail: mijl@gcu.edu.cn; Tel.: +86 20-36903493.

本文责任编辑: 张化光.

国家社科基金重点项目(15AGJ009)资助.

Supported by the National Social Science Foundation of China (15AGJ009).

复杂演化的非线性动力系统, 碳排放权价格序列本质上是一个高噪声、非线性、非平稳、带有确定性混沌的序列<sup>[1]</sup>. 相对于有效市场, 混沌与分形系统下的动态不平稳价格时间序列的预测似乎难度更大, 也更具有挑战性. 但价格预测对市场的稳健发展如此重要, 以至于无法跨过它而到达一个理想的彼岸, 这也正是本文研究的初衷.

针对国际碳市场价格的预测, 国内外学者已经进行了多方面探讨, 这些研究归纳起来主要有3个方面: 1) 从能源市场、金融市场和宏观经济与碳市场的相关性上实现对碳价格的预测; 2) 通过对碳价格序列的计量和统计特征实现对碳价格的预测; 3) 基于市场的混沌特征以非线性、非参数的方法实现对碳价格的预测.

最有代表性的从能源市场实现对碳市场价格的预测的研究当属Carolina G M等人<sup>[2]</sup>, 该研究认为在整个世界经济和能源部门不确定的条件下, 需要建立一个有效且强大的多变量预测模型, 其中价格波动特征通过条件异方差动态要素模型的方式提取, 并解决了广义自回归条件异方差模型(generalized autoregressive conditional heteroskedasticity, GARCH)估计多变量时经常出现的维数问题. 通过比较向量自回归积分移动平均模型(vector autoregressive integrated moving average, VARIMA)、单变量模型和多变量模型的预测误差, 发现在煤、石油、天然气及碳配额价格之间的动态交叉相关性可以有效地进行碳价格预测; Niblock S<sup>[3]</sup>通过皮尔逊相关系数检验、回归分析、邹检验和子阶段的回归分析研究了欧盟碳市场的资本特征, 发现碳市场与国际金融市场日趋协整, 这种联系恰好证明了碳市场收益率的可预测性; Marius C F等人<sup>[4]</sup>则是通过鉴别驱动碳市场的基本要素, 即天然气、石油和煤炭及股票指数, 建立了一个套利定价理论(arbitrage pricing theory, APT)模型, 实现了对碳价格的有效预测.

通过对碳价格序列计量和统计特征实现对碳价格的预测的研究主要包括: Byun S J, CHO H<sup>[5]</sup>比较分析了GARCH模型、K邻近算法与隐含波动率对碳期货收益的预测能力, 发现GARCH模型表现要比隐含收益率和K最邻近模型更优, 而带有正态分布的GJR-GARCH模型在所有的GARCH模型中预测效果最好; Chevallier<sup>[6]</sup>则首次将非参数模型应用于碳市场, 应用BlueNext交易所和欧洲气候交易所(ECX)碳期货日价格变化数据, 探讨非参数模型对碳价格行为发展趋势, 通过成对检验统计量验证了非参数模型的预测能力, 得出非参数模型在样本外的预测能力要强于线性的自回归(auto-regressive, AR)模型, 使得预测误差下降了15%的结论; Benz E和Truck S<sup>[7]</sup>比较了GARCH模型与Markov制度转换模型在碳配额价格波动预测时的有效性, 证明Markov制度转换模型无论是在样本内

的适用性和样本外的预测性上都要优于GARCH模型.

从混沌的角度探讨碳价格的预测问题主要是Fan等人<sup>[8]</sup>, 他们通过3个经典的指标——最大李雅普诺夫指数、关联维数及柯尔莫哥洛夫熵, 验证了欧盟排放交易体系下第3阶段碳市场的混沌特征, 确定了碳价格序列貌似随机的不规则运动, 并基于相空间重构恢复了原动力系统, 利用多层感知器(multi layered perceptron, MLP)的神经网络模型实现了对欧盟碳价格的短期预测; Bangzhu Zhu和Yiming Wei<sup>[9]</sup>则认为碳价格同时包含线性和非线性模式, 所以单一的模型并不能实现准确的预测, 因此提出以自回归积分移动平均模型(autoregressive integrated moving average, AR-IMA)和最小二乘支持向量机(least squares support vector machine, LSSVM)复合的方法来预测碳价格, 其线性与非线性结合技术对碳价格的有效预测提供了很好的启示; 高杨、李健<sup>[10]</sup>针对碳价格的非线性及非平稳性等不规则特性, 提出了基于经验模态分解-粒子群算法-支持向量机的国际碳价格误差校正模型(empirical mode decomposition-particle swarm optimization-support vector machine, EMD-PSO-SVM), 解决了误差序列的随机性和相邻频带的干扰性、预测结果的滞后性和拐点误差大等问题, 实现了预测值与误差预测值趋势的一致性, 有效地提高了模型的预测精度; 朱帮助、魏一鸣<sup>[11]</sup>提出采用数据分组处理方法-粒子群算法-最小二乘支持向量机(group method of data handling-particle swarm optimization-least squares support vector machine, GMDH-PSO-LSSVM)的非线性、非参数方法对碳价格进行预测, 给出了一种基于知识抽取和智能优化的解决方法, 并证明了该模型良好的泛化预测能力.

有别于上述研究, 本文将构建一个基于多贝西小波3层变换和单支重构的遗传算法径向基函数神经网络模型(daubechies wavelet-genetic algorithm-radial basis function neural network model, Db3-GARBF)模型来预测碳现货价格的变化趋势, 其研究思路如下: 首先, 验证建模的市场基础——一个适应于非线性范式的多重分形市场是否存在; 其次, 构建一个预测模型以对未来碳价格进行预测; 第三, 对模型的预测精度进行检验和比较; 最后, 给出研究结论.

## 2 碳排放权市场多重分形特征的提取与检验(Extraction and test of multiple-scale fractal features of carbon emission right market)

### 2.1 表征多重分形特征的指标(The indicators of multiple-scale fractal features)

多重分形特征通常通过局部霍尔德指数(Holder)和豪斯多夫分形维数(Hausdoff)来表征. 局部霍尔德指数刻画市场波动的不规则程度, 即波动的奇异性,

而与其相适应的奇异指数波动点集的豪斯多夫维数则刻画局部霍尔德指数的概率分布。

局部Holder指数描述的是时间序列的局部奇异性的变化,豪斯多夫维数反映的是具有相同局部Holder指数的点集的分形维数,二者共同构成奇异性谱(即多重分形谱)能够全面细致地刻画系统的动力学特征。

### 2.2 多重分形特征的分析方法 (Analytical methods of multiple-scale fractal features)

现存的多重分形分析方法可分为数值分析法和小波分析法两大类:相对于数值分析法,小波分析法在研究信号的多重分形特征时具有无可比拟的优势.小波分析法又分为小波变换模极大值法(wavelet transform modulus maxima, WTMM)和小波领袖法(wavelet leaders, WL),其中WL采用离散小波变换,在信号包含震荡奇异性或chirp-type类型的奇异性时均适用,分解算法较快,计算成本也比WTMM更低.故本文采用WL分析欧盟碳市场的多重分形特征,主要测度指标为:多重分形的结构函数与标度指数函数(或尺度函数)、辅助指标为:多重分形谱的宽度以及豪斯多夫维数的方差。

#### 1) 多重分形的结构函数与标度指数函数。

传统上多重分形的结构函数通常用小波变换得到,标度指数函数通过勒让德(Legendre)变换得到<sup>[12]</sup>。

多重分形的结构函数为

$$S^L(j, q) = \frac{1}{n_j} \sum_{k=1}^{n_j} L_x(j, k)^q, \quad (1)$$

其中 $n_j$ 为尺度 $2^j$ 上小波领袖系数 $L_x(j, k)$ 的个数。

多重分形的标度指数函数(或尺度函数).在 $2^j \rightarrow 0$ 时,结构函数与标度指数函数 $\zeta(q)$ 满足:  $S^L(j, q) = F_q 2^{j\zeta(q)}$ ,由此可以从结构函数 $\log S_L(j, q)$ 与 $\log 2^j$ 的线性回归中得出标度指数函数 $\zeta(q)$ :

$$\zeta(q) = \liminf_{j \rightarrow 0} \frac{\log S_L(j, q)}{\log 2^j}. \quad (2)$$

由于勒让德变换实际计算的复杂性,故本文中多重分形的结构函数与标度指数函数采用Chhabra算法进行. Chhabra(1989)根据Shannon, Eggelston和Billingsley理论提出了基于样本数据计算标度指数函数、奇异性指数和豪斯多夫维数的经验公式:

$$\begin{cases} \hat{\zeta} = \sum_{j=j_1}^{j_2} w_j \log_2 S^L(j, q), \\ \hat{D}(q) = \sum_{j=j_1}^{j_2} w_j U^L(j, q), \\ \hat{h}(q) = \sum_{j=j_1}^{j_2} w_j V^L(j, q), \end{cases} \quad (3)$$

其中:

$$U^L(j, q) = \sum_{k=1}^{n_j} R_X^q(j, k) \log_2 R_X^q(f, k) + \log_2 n_j,$$

$$V^L(j, q) = \sum_{k=1}^{n_j} R_X^q(j, k) \log_2 L_X(j, k),$$

$$R_X^q(j, k) = L_X(j, k)^q / \sum_{k=1}^{n_j} L_X(j, k)^q,$$

而权重 $w_j$ 必须满足 $\sum_{j_1}^{j_2} j w_j \equiv 1$ 且 $\sum_{j_1}^{j_2} w_j \equiv 0$ ,其可以表示为

$$w_j = b_j(V_0 j - V_1)/(V_0 V_2 - V_1^2),$$

且 $V_i = \sum_{j_1}^{j_2} j^i b_j$ ,  $i=0, 1, 2$ ,而 $b_j$ 为对 $\log_2 S^L(j, q)$ 与 $\log_2 2^j$ 进行线性回归估计标度指数函数 $\zeta(q)$ 值时所选择的置信度水平。

Chhabza算法可以根据标度指数函数 $\zeta(q)$ 与相应的多分辨量 $q$ 之间的关系来判断欧盟碳市场的分形特征:若标度指数函数 $\zeta(q)$ 与 $q$ 之间呈现线性关系,则说明碳收益序列仅具有单分形特征;若标度指数函数 $\zeta(q)$ 与 $q$ 之间呈现非线性关系,则说明碳收益序列具有多重分形的特征。

#### 2) 定义多重分形谱的宽度。

多重分形谱的宽度为 $\Delta h$ ,  $\Delta h = h_{\max} - h_{\min}$ ,  $\Delta h$ 所衡量的是市场价格波动不均匀性,  $\Delta h$ 越大,表明市场波动奇异性之间的差异越大,即波动越不均匀,多重分形特征越强;反之,则多重分形特征越弱。

#### 3) 定义豪斯多夫维数的方差。

豪斯多夫维数的方差为 $S_D$ ,  $S_D = \text{Var}(D(h))$ ,它反映市场波动的奇异性强度.作为对多重分形谱宽度的补充,豪斯多夫维数方差从一个更全面和细致的角度反映了碳收益序列波动奇异性的复杂程度。

### 2.3 EU-ETS多重分形特征检验 (The test of the multiple-scale fractal features of EU-ETS)

本文选择BlueNext交易所和欧洲气候交易所(ECX)碳交易作为研究对象.交易品种为碳配额(european union allowance, EUA),价格数据均来源于彭博数据库(Bloomberg database),具体如表1所示。

表1 样本数据

Table 1 Sample data

时期	交易所	样本量	样本区间
第1阶段	BlueNext交易所	500	2005年6月27日至2007年6月29日
第2阶段	BlueNext交易所	1186	2008年2月26日至2012年12月5日
第3阶段	欧洲气候交易所ECX	622	2012年12月7日至2015年5月8日

**注1** 由于第1阶段碳配额不能储存到第2阶段,所以第1阶段末碳配额价格趋于零,考虑到数据的有效性,本文剔除了2007年后半年的数据。

考虑到碳价格数据的波动特征,本文首先将碳现货价格序列转换为对数收益率序列:

$$\{R_t | R_t = \ln P_t - \ln P_{t-1}, t = 2, 3, \dots, N\}, \quad (4)$$

其次,选择基于消失矩  $N_\psi = 3$  的 Daubechies 小波 (Db3) 对碳收益率序列进行离散小波变换,根据小波领袖系数函数式得到碳收益率序列的小波领袖系数;为尽量保证所得多重分形谱的完整性,通过多次试验将三阶段的多分辨量矩阶数  $q$  分别选为  $[-15, 15]$ ,  $[-8, 8]$  和  $[-12, 12]$ ;最后,根据 Chhabza 算法得到不同多分辨量矩阶数下的标度指数函数值  $\zeta(q)$ , 据此判断欧盟三阶段碳收益率序列的分形性。

标度指数函数  $\zeta(q)$  与多分辨量矩阶数  $q$  之间的关系如图1所示。

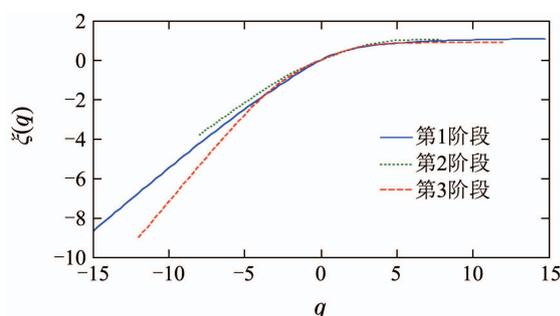


图1 欧盟碳市场三阶段的标度指数函数  $\zeta(q)$  与多分辨量矩  $q$  之间的关系

Fig. 1 The relationship between the scale index function  $\zeta(q)$  and the multi-resolution moment  $q$  in the three phases of EU-ETS

图1显示: 欧盟碳市场三阶段的标度指数函数  $\zeta(q)$  与多分辨量矩  $q$  之间均呈现显著的非线性关系且为凸的递增函数,表明欧盟三阶段碳市场均具有多重分形特征。

用欧盟三阶段碳收益序列的多重分形谱宽度和豪斯多夫维数方差表征的多重分形特征如表2所示。

表2 表征欧盟碳市场的多重分形特征的宽度和方差  
Table 2 The width and variance of the multiple-scale fractal features of EU-ETS

指标	第1阶段	第2阶段	第3阶段
$\Delta h$	0.6361	0.5400	0.9123
$S_D$	0.1145	0.0907	0.4299

表2显示: 第3阶段碳收益序列的多重分形谱宽度和豪斯多夫维数方差远大于第1-2阶段,说明该阶段波动最不均匀、多重分形强度最强且奇异性最为复杂;与之相反,第2阶段碳收益序列无论在谱宽度和维数方差上都小于其他两个阶段,说明价格序列波动较为均匀、多重分形程度较弱,市场效率略高。

### 3 碳价格预测算法 (The prediction algorithm of the carbon price)

上述研究表明: 欧盟碳排放权市场价格波动具有很强的多重分形特征,即显著的多尺度变化特征,碳价格预测模型必须能够准确捕捉这种局部尺度的多样性。鉴于此,本文构建了一个能够刻画碳价格多尺度变化特征的预测模型——基于多贝西小波三层变换单支重构的遗传算法径向基函数神经网络模型 (daubechies wavelet-genetic algorithm-radial basis function neural network model, Db3-GA-RBF) 来预测欧盟碳现货价格。不失其一般性,该模型将实现对碳价格序列的一步向前预测,即

$$P_t = f(P_{t-1}, P_{t-2}, \dots, P_{t-m}), \quad (5)$$

其中  $m$  为常数,表示滞后阶数。

Db3-GA-RBF 模型的具体建模步骤分为3步: 1) 确定最优滞后期  $m$ , 即确定输入到模型中的原始序列和滞后序列; 2) 对原始序列和滞后序列进行小波变换和单支重构,获取不同尺度上的分量组合; 3) 初始化RBF神经网络,确定网络结构和网络参数,完成建模。

#### 3.1 最优滞后期的确定 (The determination of the optimal lag period)

由于碳价格预测模型是从过去的滞后序列中实现对碳价格向前预测,所以在构建碳价格预测模型之前首先要确定碳价格序列的最优滞后期。相对于其他确定时间序列最优滞后期的方法,施瓦茨信息准则 (Schwartz's information criterion, SIC) 法不仅对于样本内及样本外预测表现均适用,而且无论在平稳还是非平稳的系统中施瓦茨信息准则都能得到一致性的估计。所以本文选择施瓦茨信息准则作为确定最优滞后期的标准,施瓦茨信息准则的对数形式为

$$\ln \text{SIC} = \frac{k}{N} \ln N + \ln \frac{\text{RSS}}{N}, \quad (6)$$

其中:  $k$  表示滞后的阶数或回归元个数,  $N$  表示样本量,  $\text{RSS}$  表示自回归的残差平方和,  $\frac{k}{N} \ln N$  是自由度丧失的惩罚因子,  $\ln \text{SIC}$  越低说明模型效果越好,所以选择使  $\ln \text{SIC}$  达最小值的滞后期作为最优滞后期。由于目前普遍认为由 Huang W 等人<sup>[14]</sup> 所提出的自相关准则 (autocorrelation criterion, AC) 能够有效地确定序列的最优滞后期,所以,本文在模型的分析中同时给出基于该准则的价格预测模型,通过比较来检验本文所构建的价格预测模型的有效性。

#### 3.2 小波变换和单支重构 (Wavelet transform and single reconfiguration)

1) 小波变换. 将原始信号  $f(t)$  与母小波  $\psi(t)$  做平移伸缩后所得的函数族做内积:

$$Wf(a, b) = \langle f, \psi_{a,b} \rangle = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt, \quad (7)$$

其中:  $\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right)$  是由母小波  $\psi(t)$  经过伸缩平移而得到的函数族,  $a$  为伸缩因子,  $b$  为平移因子. 通过将原始信号进行小波变换便可以得到在不同尺度上的小波系数<sup>[15]</sup>.

然而, 在小波变换将信号分解为逼近信号和细节信号时, 小波细节信号的长度将随分解层数的增加而逐层减半, 即小波细节信号的时域分辨率将会逐层降低. 为使每一个小波细节信号都能单独表征原始信号在相应分辨率(尺度)下波动特征, 需要对每一个小波分解信号进行单支重构.

2) 单支重构. 在令其他小波细节信号和小波逼近信号均为零的条件下, 仅利用某一个尺度下小波细节信号分量进行原始信号的重构, 是小波变换的逆运算:

$$f(t) = D_1 + D_2 + \dots + D_L + A_L, \quad (8)$$

其中:  $L$  为分解层数,  $A_L$  为表征信号结构特征的逼近信号的重构,  $D_j (j = 1, 2, \dots, L)$  为表征原信号细节特征的不同分辨率下细节信号的重构.

在碳价格序列经过小波变换和单支重构之后, 可以获得能够表征其不同分辨率下波动特征的独立分量序列, 从而提高神经网络模型对碳收益率序列的局部尺度多样性的刻画能力, 实现对碳现货价格更准确的预测.

### 3.3 初始化RBF神经网络(Initialize the RBF neural network)

径向基函数(radial basis function, RBF)神经网络是一种性能良好的三层前馈型神经网络, 其基本原理是: 将输入向量映射到由径向基函数作为“基”的隐含层空间, 实现在低维空间线性不可分的问题在高维空间的线性可分, 然后线性映射得到输出. 由于高斯形式的径向基函数具有局部响应的能力, 其逼近函数是闭集合, 使得RBF神经网络具有唯一最佳逼近的性质且不存在局部最小值问题<sup>[16]</sup>.

RBF神经网络必须经过初始化之后才能进行预测, 初始化过程是确定网络结构和网络参数的过程. 初始化过程需要解决两方面的关键问题: 1) 确定隐含层核函数数目(或节点数), 即最优网络结构的确定; 2) 确定径向基函数的中心、宽度和输出权值, 即最优参数的确定.

#### 1) 最优网络结构的确定.

在RBF神经网络结构中, 若隐含层核函数数目过少, 会使得神经网络的训练不够充分, 从而导致学习算法陷入局部最小值; 反之, 若隐含层核函数数目过多, 又可能会引起“过拟合”(overfitting)问题. 本文将

“最优”核函数数目的标准确定为: 训练误差持续下降而预测误差由下降转为上升时的节点数作为最优的隐含层节点数. 即: 从最小的隐含层节点数(即1个)开始尝试, 然后逐渐增加隐含层节点数, 在获得每个模型的训练误差和泛化误差的基础上依据“最优标准”来确定神经网络的核函数数目.

#### 2) 最优参数的确定.

在众多网络优化算法中, 遗传算法(genetic algorithm, GA)是一种全局并行、随机搜索, 具有很强的鲁棒性及收敛到全局最优能力的算法. 建模中引用此方法主要是用来优化RBF神经网络中的参数, 包括: 径向基函数的中心、宽度和输出权值<sup>[17]</sup>.

最优参数的确定包括: 编码、初始种群的确定、适应度函数的构造、遗传操作、权值计算.

① 编码. 编码分二进制编码和实数编码, 本文由于数据量较大, 故采用实数编码;

② 初始种群的确定. 本文采用迭代法取初始种群  $m = 50$ ;

③ 适应度函数的构造. 为获得逼近能力最强即训练误差最小的神经网络, 本文采用的适应度函数如下:

$$\text{Fit} = \frac{1}{\text{ESS} + \mu}, \quad (9)$$

其中: ESS表示个体误差的平方和,  $\mu$ 表示一个极小的常数防止出现ESS = 0的情况, 设  $\mu = 1.0E - 10$ . 个体的适应度值越大, 则被选中的可能越大.

④ 遗传操作. 遗传操作包括: 选择操作、交叉操作和变异操作.

选择操作采用基于适应度比例的轮盘赌注法进行随机采样, 对个体  $x_i$ , 其被选择的概率为  $p_s = f_i / \sum f_i$ , 取种群数目  $m$  乘以  $p_s$ , 即确定了第  $i$  个个体在下一代中应自我选择复制的数目; 交叉操作采用两点交叉的方法, 取交叉概率  $p_c$  在 [0.5 - 1.0] 之间; 变异操作取变异概率  $p_m$  在 [0.01 - 0.2] 之间.

⑤ 权值计算. 对隐含层与输出层之间的线性权值利用下式计算:

$$w_i(n+1) = w_i(n) + \xi \cdot \frac{\partial E(n)}{\partial w}, \quad (10)$$

式中:  $E(n)$  表示平均误差平方和;  $n$  表示迭代次数;  $\xi$  表示学习速度,  $\xi \in (0, 1]$ . 若遗传代数(迭代次数)达到给定的允许值或其他收敛条件满足时, 则停止遗传, 即当  $E$  小于指定  $\varepsilon$  时, 计算停止; 否则转回步骤③, 本文中取  $\varepsilon = 0.00001$ .

重复③-⑤对RBF神经网络进行训练, 以不断修正进化初始权值分布, 求得全局最优解. 将遗传算法最终所得最优参数保存为RBF神经网络的参数值.

在确定了最优的网络结构和最优的网络参数之后,

便完成了RBF神经网络模型的初始化过程. 然后可根据此模型对未来碳现货价格进行预测.

#### 4 基于经验数据的实证分析 (The empirical analysis based on the empirical data)

##### 4.1 数据来源与预处理 (Data sources and preprocessing)

本文所选择的数据均来自彭博(Bloomberg)数据库, 具体见前文表1. 鉴于欧盟碳排放权交易市场的前两个阶段都已完结, 所以本文的碳价格预测主要针对欧盟碳市场的第3阶段进行, 第3阶段的数据来源于欧洲气候交易所, 数据区间为2012年12月7日至2015年5月8日, 总计622个EUA现货价格数据. 为了保证数据的充分性和有效性<sup>1</sup>, 对第1阶段碳价格数据进行分析时剔除2007年6月29日之后的数据.

数据处理及模型构建均采用MATLAB R2012b软件. 为方便后续数据处理并加快模型收敛速度, 在建模之前首先对碳现货价格序列及其滞后序列进行归一化处理:

$$y_t = (0.99 - 0.01) \cdot \frac{[x_t - \min_t(x_t)]}{[\max_t(x_t) - \min_t(x_t)]} + 0.01. \quad (11)$$

利用上式, 可将碳价格数据变换到[0.01, 0.99]范围内. 欧盟碳市场第3阶段EUA价格序列归一化处理之后的结果如图2所示(原始数据单位: 欧元/吨二氧化碳).

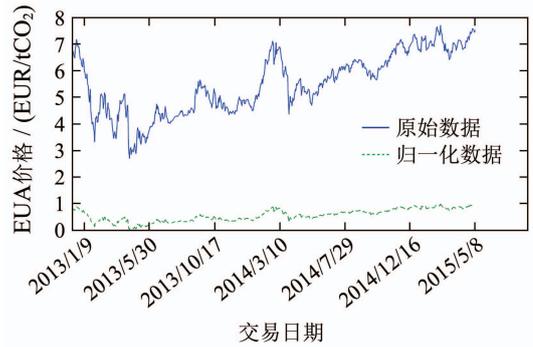


图 2 第3阶段EUA价格序列的归一化处理  
Fig. 2 Normalization processing of EUA price sequence in the third stage

##### 4.2 预测模型的构建(The prediction model construction)

###### 1) 最优滞后期的确定.

本文根据原始碳现货价格序列的SIC统计量来确定其最优滞后期. 原始碳价格序列在各滞后阶数下的SIC统计量如表3所示: 当  $k = 3$  时  $\ln SIC$  最小, 所以碳价格序列的最优滞后期为 3, 即  $m = 3$ . 选择  $P_{t-1}$ ,  $P_{t-2}$ ,  $P_{t-3}$  作为碳价格预测模型的自变量. 在此情况下, 欧盟碳市场第3阶段的有效数据为619个(因选择3阶连续滞后作为自变量从而丧失了3个自由度). 为使价格预测模型对碳价格波动的规律和特征进行充分的学习, 本文将前550个作为模型的训练集数据(样本内数据), 后69个作为模型的测试集数据(样本外数据)用以检验模型的泛化能力.

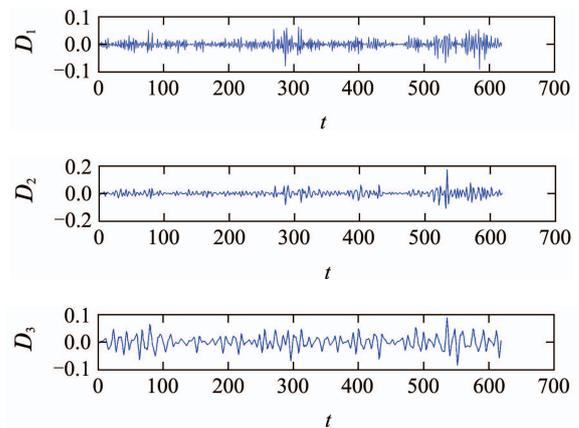
表 3 SIC统计量  
Table 3 SIC statistics

滞后期	1	2	3	4	5	6	7
SIC	-245.3954	-250.3487	-254.1127	-251.4343	-251.1232	-246.9998	-239.6131

###### 2) 不同尺度分量的获取.

小波变换中母小波的选择对准确描述原始信号的特征具有至关重要的意义. 本文采用多贝西小波(daubechies wavelet, DbN)来对碳价格序列进行小波变换和单支重构. 由于消失矩  $N_\psi = 3$  的 Daubechies 小波 (Db3) 最符合原始碳价格序列的波动特征, 所以选择基于Db3小波对碳价格序列进行3层小波分解和单支重构. 由此得到3个高频细节分量和一个低频结构分量(如图3所示). 由于通过小波变换和单支重构所得到的细节分量和结构分量分别表征出不同尺度上原始信号的特征, 所以本文将训练集中原始价格序列和滞后序列在相应尺度上的分量分别进行组合输入到神经网络模型中, 使该模型对碳价格序列不同尺度上的

特征进行充分学习.



<sup>1</sup> 由于第1阶段碳配额禁止储存(no banking)的规定使得第1阶段多余的碳配额在第2阶段失效, 所以第1阶段末碳配额价格趋于零, 考虑到数据的有效性, 本文剔除了2007年后半年的数据.

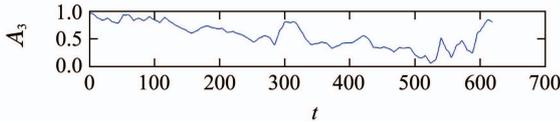


图3 碳现货价格序列的3层小波变换和单支重构分量  
Fig. 3 The 3-layer wavelet transform and single reconfiguration component of the carbon spot price sequence

### 3) 网络结构的确定.

表4显示了不同核函数数目下模型的训练误差和预测误差,从表4中可知:基于小波变换和单支重构的GA-RBF神经网络模型的训练误差随着核函数数目的增加持续下降,但预测误差在核函数数目为5时由下降转为上升,根据“最优”标准,选取模型最优隐含层核函数数目4,此时模型的泛化能力最强.

表4 不同核函数数目下模型的训练和预测误差  
Table 4 The model training and prediction error of different kernel function numbers

核函数数目	预测误差	训练误差
1	1.0208E-01	5.3711E-02
2	3.8177E-02	1.9413E-02
3	2.3803E-02	6.6660E-03
4	2.0885E-02	6.2562E-03
5	2.3135E-02	6.0579E-03
6	2.3464E-02	5.8882E-03
7	2.5186E-02	5.9301E-03
8	2.5256E-02	5.7966E-03
9	2.7644E-02	5.7099E-03
10	2.9010E-02	5.6740E-03

### 4) 网络参数的确定.

在确定了RBF神经网络的结构之后,则需要基于训练集数据采用遗传算法来确定模型最优的参数.在遗传算法中,初始种群的规模设为50,最大训练代数设为100,在进化过程中,为提高效率,将交叉概率设为0.9,变异概率设为0.01.当满足进化终止条件时,将最优参数保存到神经网络模型中,得到最优的RBF神经网络模型.

## 4.3 预测结果(The result of predicting)

运用训练好的RBF神经网络对测试集的碳现货价格进行预测结果见图4:该图显示了基于小波变换和单支重构的GA-RBF神经网络模型对测试集内碳现货价格的预测值与实际值的对比效果,从图中可以看出:预测值与实际值差异普遍较小.由此,本文所构建的价格预测模型对欧盟碳排放权市场的配额现货价格实现了不错的预测效果.

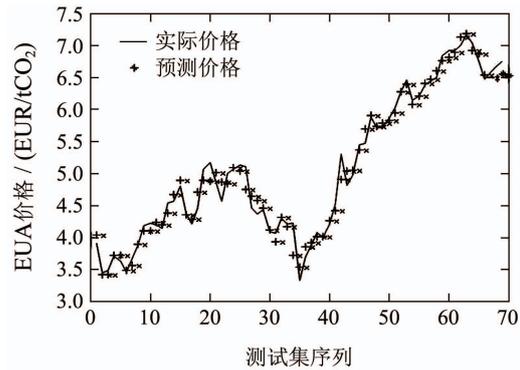


图4 Db3-GA-RBF模型的碳现货价格预测图

Fig. 4 The carbon spot price forecast of Db3-GA-RBF model

## 5 多种预测方法效果比较(The comparison of different prediction models)

为检验本文构建的价格预测模型的效果,将其与支持向量机模型(support vector machine, SVM)、遗传算法优化的支持向量机模型(genetic algorithm-support vector machine, GA-SVM)和遗传算法优化的径向基神经网络模型(genetic algorithm-radial basis function neural network, GA-RBF)对测试样本内碳价格的预测结果进行比较分析.比较从两个角度进行:一是预测值的精确度,采用均方误差(mean square error, MSE)和平均绝对误差(mean absolute error, MAE)两个指标来测度;二是预测方向的精确度,采用预测方向精度(sign correctness percentage, SCP)来衡量.3个指标的计算公式如下:

$$\text{MSE} = \frac{1}{N} \sum_{t=1}^N (\hat{y}_t - y_t)^2, \quad (12)$$

$$\text{MAE} = \frac{1}{N} \sum_{t=1}^N |\hat{y}_t - y_t|, \quad (13)$$

$$\begin{cases} \text{SCP} = \frac{100}{N} \sum_{t=1}^N d(t), \\ d(t) = \begin{cases} 1, & (\hat{y}_t - \hat{y}_{t-1})(y_t - y_{t-1}) > 0, \\ 0, & (\hat{y}_t - \hat{y}_{t-1})(y_t - y_{t-1}) \leq 0. \end{cases} \end{cases} \quad (14)$$

比较结果见表5.

表5 不同预测模型的预测精度比较

Table 5 The comparison of the prediction accuracy of different prediction models

模型	MSE	MAE	SCP
SVM	0.0801	0.2230	61.7647
GA-SVM	0.0797	0.2208	60.2941
GA-RBF	0.0765	0.2127	60.2941
Db3-GA-RBF(AC)	0.0411	0.1524	74.6269
Db3-GA-RBF(SIC)	0.0209	0.1057	80.8824

表5显示:

1) Db3-GA-RBF(包括 AC 和 SIC)模型与常用的

其他价格预测模型预测结果比较:Db3-GA-RBF模型的预测误差(无论是均方误差或平均绝对误差)均小于SVM,GA-SVM和GA-RBF模型,其预测精度分别提高了73.9%,73.7%和72.7%,而且预测方向的正确率也有大幅提升,说明Db3-GA-RBF类模型对碳现货价格实现了更为准确的预测;

2) Db3-GA-RBF(SIC)与Db3-GA-RBF(AC)的预测结果比较:Db3-GA-RBF(SIC)的预测误差明显小于Db3-GA-RBF(AC),预测方向精度明显高于Db3-GA-RBF(AC),AC准则下的碳价格预测模型并没有显著提高对测试集碳价格预测的准确性,相较之下,SIC准则更为合理和有效;

3) Db3-GA-RBF(包括AC和SIC)模型与没有经过小波变换和单支重构的GA-RBF模型比较:Db3-GA-RBF模型的预测误差小于GA-RBF,预测精度远大于GA-RBF模型的预测精度,说明通过小波变换单支重构所得的各尺度下的分量能够有效地刻画具有多重分形特征的碳现货价格的时、频两域局部特征,从而较大程度提升了模型的预测精度。

## 6 结论(Conclusions)

本文的研究结论如下:

1) 欧盟碳排放权市场配额(european union allowance, EUA)现货价格波动均具有局部尺度多样性特征,且第3阶段碳价格波动最不均匀,多重分形特征最强.表明经典的有效市场假说无法解释碳排放权市场的价格行为特征,本质上,碳排放权市场是一个多重分形和混沌市场;

2) 与原始价格系列相比,经过小波变换单支重构后的价格序列已经有效地平滑掉原始信号的边缘和噪声,得到了能够准确反映信号本身特征的细节信息.说明小波变换单支重构可以有效提高数据的准确性,继而使模型的预测精度更强;

3) 与其他预测模型效果相比,本文所构建的Db3-GA-RBF(SIC)模型能较大幅度地避免预测误差,提高预测精度,预测精度提高约70%.这一研究为碳排放权价格及其他金融资产的价格预测提供了一种新的思路和方法。

## 参考文献(References):

- [1] YANG Xing, LIANG Jingli. Analysis and test of fractal and chaotic behavior characteristics of the international carbon emissions market [J]. *Systems Engineering—Theory & Practice*, 2017, 37(6): 1420 – 1431.  
(杨星, 梁敬丽. 国际碳排放权市场分形与混沌行为特征分析与检验——以欧盟碳排放交易体系为例 [J]. *系统工程理论与实践*, 2017, 37(6): 1420 – 1431.)
- [2] CAROLINA G M, JULIO R, MARIA J S. Modelling and forecasting fossil fuels, CO<sub>2</sub> and electricity prices and their volatilities [J]. *Applied Energy*, 2013, 101(C): 363 – 375.
- [3] NIBLOCK S. A diachronic informational efficiency investigation of European carbon markets [D]. Lismore NSW: Southern Cross University, 2011.
- [4] MARIUS C F, DOMINIQUE G, ANTONIN L. *Statistical Evidence of Tax Fraud on the Carbon Allowances Market* [R]. <https://halshs.archives-ouvertes.fr/halshs-00523458>, 2010.
- [5] BYUN S J, CHO H. Forecasting carbon futures volatility using GARCH models with energy volatilities [J]. *Energy Economics*, 2013, 40(2): 207 – 221.
- [6] CHEVALLIER J. Nonparametric modeling of carbon prices [J]. *Energy Economics*, 2011, 33(6): 1267 – 1282.
- [7] BENZ E, TRUCK S. Modelling the price dynamics of CO<sub>2</sub> emission allowances [J]. *Energy Economics*, 2009, 31(1): 4 – 15.
- [8] FAN X H, LI S S, TIAN L X. Chaotic characteristic identification for carbon price and an multi-layer perception network prediction model [J]. *Expert Systems with Applications*, 2015, 42(8): 3945 – 3952.
- [9] ZHU B Z, WEI Y M. Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology [J]. *Omega*, 2013, 41(3): 517 – 524.
- [10] GAO Yang, LI Jian. International carbon finance market price prediction based on EMD-PSO-SVM error correction model [J]. *China Population, Resources and Environment*, 2014, 24(6): 163 – 170.  
(高杨, 李健. 基于EMD-PSO-SVM误差校正模型的国际碳金融市场价格预测 [J]. *中国人口·资源与环境*, 2014, 24(6): 163 – 170.)
- [11] ZHU Bangzhu, WEI Yiming. Carbon price prediction based on integration of GMDH, particle swarm optimization and least squares support vector machines [J]. *Systems Engineering—Theory & Practice*, 2011, 31(12): 2265 – 2271.  
(朱帮助, 魏一鸣. 基于GMDH-PSO-LSSVM的国际碳市场价格预测 [J]. *系统工程理论与实践*, 2011, 31(12): 2265 – 2271.)
- [12] HU Chaofang, XIE Qianqian. Sum of squares-robust model predictive controller for nonlinear system with input saturation [J]. *Control Theory & Applications*, 2016, 33(3): 321 – 327.  
(胡超芳, 解倩倩. 非线性系统有输入饱和时基于平方和的鲁棒模型预测控制器 [J]. *控制理论与应用*, 2016, 33(3): 321 – 327.)
- [13] CHHABRA A, MENEVEAU C, JENSEN R V, et al. Direct determination of the  $f(\alpha)$  singularity spectrum and its application to fully developed turbulence [J]. *Physical Review A General Physics*, 1989, 40(9): 5284 – 5294.
- [14] HUANG W, NAKAMORI Y, WANG S. A general approach based on autocorrelation to determine input variables of neural networks for times series forecasting [J]. *Journal of Systems Science and Complexity*, 2004, 17(3): 297 – 305.
- [15] GRAHAM M, NIKKINEN J. Co-movement of the Finnish and international stock markets: a wavelet analysis [J]. *The European Journal of Finance*, 2011, 17(5/6): 409 – 425.
- [16] HOSSEIN M, TAMER I. An RBF collocation method for solving optimal control problems [J]. *Robotics and Autonomous Systems*, 2017, 87: 219 – 225.
- [17] JIN Min, LU Huaxiang. A multi-subgroup hierarchical hybrid of genetic algorithm and particle swarm optimization [J]. *Control Theory & Applications*, 2013, 30(10): 1231 – 1238.  
(金敏, 鲁华祥. 一种遗传算法与粒子群优化的多子群分层混合算法 [J]. *控制理论与应用*, 2013, 30(10): 1231 – 1238.)

## 作者简介:

杨星 (1965–), 女, 教授, 博士生导师, 主要研究方向为碳金融、金融创新, E-mail: tyxing@jnu.edu.cn;

梁敬丽 (1990–), 女, 硕士, 主要研究方向为碳金融, E-mail: 18319567825@163.com;

蒋金良 (1953–), 男, 副教授, 主要研究方向为管理科学与工程, E-mail: jiangjl@gcu.edu.cn;

米君龙 (1987–), 男, 硕士, 讲师, 主要研究方向为金融风险, E-mail: mijl@gcu.edu.cn.