

中国油气资源生产发展最优组合预测模型研建

陶冶, 薛惠锋

TAO Ye, XUE Hui-feng

西北工业大学 自动化学院, 西安 710072

College of Automation, Northwestern Polytechnical University, Xi'an 710072, China

E-mail: taoyemail@yahoo.com.cn

TAO Ye, XUE Hui-feng. Construction and application of combination forecasting model in Chinese energy output development tendency. *Computer Engineering and Applications*, 2009, 45(34): 208-211.

Abstract: In view of the complexity and non-linearity of energy system, this paper analyzes the historical data of energy production in China, the respective model of gray forecast and nonlinear regression model are established firstly, then the method of optimized combination is applied to allocate the weights for the accurate forecast model and forecast value. It is proved that the combination forecasting model is suitable and feasible for studying the production of energy system through analyzing characteristic of the model. The result shows that the model can be taken as an effective tool to predict Chinese future energy production, analyzing the development trend of energy production for future China, as well as providing reliable scientific basis for energy development strategy of China.

Key words: energy forecast; nonlinear regression; gray model; combination forecast

摘要: 为了得到更合理、准确的未来能源发展趋势预测结果, 利用非线性函数模型描述能源生产发展系统的非线性动态特征。分别采用非线性指数回归模型和灰色系统理论中的模型方法建立了中国油气资源生产发展预测模型。并利用组合预测理论及建模技术采用标准差法进行最优组合权重分配, 建立了中国油气资源生产发展的组合预测模型。经检验达到了较好的预测精度。结果表明, 这一模型适合于油气资源生产发展的趋势预测, 为能源预测提供了新的方法, 并在实例中进行了初步的预测应用, 取得了较好的效果, 为我国的能源发展战略提供可靠的科学依据。

关键词: 能源预测; 非线性回归; 灰色模型; 组合预测

DOI: 10.3778/j.issn.1002-8331.2009.34.065 **文章编号:** 1002-8331(2009)34-0208-04 **文献标识码:** A **中图分类号:** TP311

1 引言

油气资源是不可再生的重要矿产资源, 是关系到国际经济安全的重要战略物资。从 20 世纪 60 年代开始, 石油和天然气在世界一次能源消费结构中的比重达到 40% 以上, 并超过煤炭, 成为现代经济的主要动力。进入 21 世纪以来, 中国经济持续快速发展, 拉动能源特别是油气资源需求迅速增长, 供需矛盾越来越突出, 国家油气能源安全供应成为关注的焦点。

作为制定能源战略规划的重要组成部分, 能源生产发展趋势预测是通过对能源生产发展趋势的明确分析和较精确的预测, 为我国的能源发展战略提供可靠的科学依据, 据此确定我国未来一段时间甚至更长时期的能源发展对策, 制定最优的能源战略、安排能源建设及优化配置等工作, 确保经济社会发展目标的顺利实现, 具有重要的理论、经济和政治意义^[1]。

当前中国的能源生产, 伴随国家经济高速发展, 并同时受到多种因素及外部环境的影响和制约, 有其复杂性及非线性的历史发展演化趋势特征。有鉴于此, 利用非线性函数模型描述

能源生产发展系统能比较准确地符合其非线性动态特征。为了得到更合理、准确的未来能源发展趋势预测结果, 在对我国能源生产系统进行定性分析的基础上, 分别采用非线性指数回归模型和灰色系统理论中的 GM(1, 1) 模型方法建立了中国油气资源生产发展预测模型。并利用组合预测理论及建模技术克服单一模型在假设条件及适用范围的局限性, 有效地集结更多的有用信息, 采用标准差法进行最优组合权重分配, 建立了中国油气资源生产发展的组合预测模型, 从而获得更为精确的预测模型和预测值, 并对我国未来 20 年中国油气资源生产发展进行了预测。

2 中国油气资源开发现状和趋势

2.1 中国油气资源开发现状

国内油气资源总量丰富, 石油探明占总量的 1/3, 天然气探明占总量的 1/8, 石油探明稳定增长, 天然气探明快速增长, 石油产量小幅稳定增长, 天然气产量快速增长。“十五”期间, 油气

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.59808013)。

作者简介: 陶冶(1980-), 男, 博士生, 主要研究领域为资源环境系统工程、能源和气候问题; 薛惠锋(1964-), 男, 教授, 博士生导师, 主要研究领域为资源环境系统工程、管理系统工程。

收稿日期: 2008-12-11 **修回日期:** 2009-02-16

开发投入逐年增加,5年累计完成投资3470亿元,完钻开发井66700口,新建产能11250万吨。原油产量由2000年的1.63亿吨增加到2005年的1.815亿吨,年均增加370万吨,年均增长2.17%;其中2005年较2004年净增650万吨,增长3.7%。天然气产量由2000年的272亿立方米增加到2005年的500亿立方米,年均增加45亿立方米,年均增长13%。其中2005年较2004年净增92亿立方米,增长22.5%。石油天然气开发水平不断提高。

2.2 中国油气资源潜力

根据新一轮全国油气资源评价结果,我国油气资源总体探明程度不高,待探明油气资源潜力较大。

2.2.1 石油处于勘探中期,储量产量稳定增长

我国石油地质资源量765亿吨,可采资源量212亿吨。截至2005年底,累计探明石油地质储量258亿吨,待探明石油地质资源量507亿吨,平均探明程度34%,勘探已进入中期^[2]。预测到2020年,我国东部石油探明地质储量平缓下降,中西部和海域探明地质储量稳步上升,全国总体呈上升趋势。石油增储的主要领域有渤海湾、松辽、鄂尔多斯盆地的岩性、地层领域;渤海湾盆地的滩海和海域、塔里木盆地环满加尔凹陷深层坡折带和麦盖提斜坡、准噶尔盆地腹部、珠江口盆地深水等探区。渤海湾等7大含油盆地,除松辽外,石油储盆、产量均稳定增长。根据“十五”石油探明储量揭示,“十一五”期间,渤海海(含海域)盆地原油产量突破7000万吨。鄂尔多斯盆地原油产量接近2500万吨,新疆地区(塔里木、准噶尔和吐哈盆地)原油产量达到3500万吨。

2.2.2 天然气处于勘探早期,储量产量快速增长

天然气地质资源量35万亿立方米,可采资源量22万亿立方米,截至2005年底,累计探明天然气地质储量4.9万亿立方米,待探明天然气地质资源量30.1万亿立方米,平均探明程度13%,勘探处于早期天然气资源主要集中在塔里木、四川、鄂尔多斯三大盆地,占全国54%,奠定了“西气东输”的资源基础^[3]。预测到2020年,天然气探明地质储量年均4000~5000亿立方米,增长势头可持续到2030年之后;全国天然气产量呈现跨越式增长,年均增加90~110亿立方米。天然气增储主要领域为塔里木、准噶尔、柴达木等盆地前陆区;鄂尔多斯古生界、四川盆地中下部海相地层;松辽盆地断陷层系,海域中浅层。塔里木等7大含气盆地天然气储量、产量快速增长。

3 模型建立

3.1 非线性指数回归模型的建立

从2001年以来中国能源消费弹性系数始终保持高位,其中在2003年和2004年甚至一度达到了1.6。从1978年开始的以能源消耗的较低增长支撑快速发展的状况出现根本逆转,能源消费出现了过快增长的势头。对于具有增长趋势的能源生产问题,由于其发展趋势是逐年增长的非线性趋势,所以通常建立指数回归模型进行预测。

指数回归模型可表示为^[4]:

$$E_t = AB^t$$

式中: E_t 为因变量; t 为自变量; A, B 为计算参数。为将指数回归转化为线性回归,取对数得:

$$\lg E_t = \lg A + t \lg B \quad (1)$$

可得 $y = \lg E_t, a = \lg A, b = \lg B$

$$\text{则有 } y = a + bt \quad (2)$$

由最小二乘法可得公式(2)参数 a, b 为:

$$\hat{a} = \bar{y} - \hat{b} \bar{t} \quad (3)$$

$$\hat{b} = \frac{\sum_{i=1}^N (t_i - \bar{t})(y_i - \bar{y})}{\sum_{i=1}^N (t_i - \bar{t})^2} \quad (4)$$

为方便计算,取 N 为奇数,并把时间 $t=1, 2, \dots, N$ 的正中央 $t = \frac{1+N}{2}$ 的编号值取为零,前面部分取为负数(即: $-1, -2, \dots, -\frac{N-1}{2}$),后面部分取为正数(即: $1, 2, \dots, \frac{N-1}{2}$),可得 $\bar{t}=0$,参数 a, b 的估计式为:

$$\hat{a} = \bar{y}_t = \frac{1}{N} \sum_{i=1}^N y_i \quad (5)$$

$$\hat{b} = \frac{\sum_{i=1}^N t_i y_i}{\sum_{i=1}^N t_i^2} \quad (6)$$

对于指数回归模型可用标准离差进行检验,即,

$$s = \sqrt{\frac{\sum (y_i - \hat{y}_i)^2}{N-R-1}} \quad (7)$$

式中: \hat{y}_i 为预测值序列, N 为原始值序列长度, R 为自由度。若 $S/R < 10\%$,则模型预测精度良好。

3.2 等维递补灰色预测模型GM(1,1)的建立

灰色系统理论认为,尽管客观系统表象复杂,但其仍具有整体功能,因此必然蕴涵某种内在规律,一切灰色序列都能通过某种生成弱化其随机性,显其规律性^[5]。灰色理论的微分方程模型称为灰色模型(Grey Model, GM),单变量一阶微分方程GM(1,1)只涉及一个变量,原理如下^[6-7]:

设等间隔时间序列 $x^{(0)}$ 有 n 个观察值 $x^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$,作 m 次累加生成,即 $x^{(m)}(k) = \sum_{i=1}^k x^{(m-1)}(i)$,可得随机性被弱化 m 次后的序列 $x^{(m)} = \{x^{(m)}(1), x^{(m)}(2), \dots, x^{(m)}(n)\}$ ($m=1, 2, \dots$),令 $z^{(m)}$ 为 $x^{(m)}$ 的均值序列, $z^{(m)}(k) = \frac{1}{2}[x^{(m)}(k) + x^{(m)}(k-1)]$,则 $z^{(m)} = [z^{(m)}(2), z^{(m)}(3), \dots, z^{(m)}(n)]$ 。

随机非负序列经过多次累加后,一般可采用指数曲线逼近,以 $m=1$ 为例建立白化型灰色预测模型,白化型GM(1,1)微分方程为:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \quad (8)$$

式中 a 为发展系数; b 为灰作用量。

根据最小二乘原则,可得参数 a, b :

$$a = \frac{CD - (n-1)E}{(n-1)F - C^2} \quad (9)$$

$$b = \frac{DF - CE}{(n-1)F - C^2} \quad (10)$$

$$C = \sum_{k=2}^n z^{(1)}(k); D = \sum_{k=2}^n x^{(0)}(k) \quad (11)$$

$$E = \sum_{k=2}^n z^{(1)}(k)x^{(0)}(k); F = \sum_{k=2}^n z^{(1)}(k)^2$$

求解微分方程可得预测模型为:

$$\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a} \quad (k=0, 1, 2, \dots, n) \quad (12)$$

还原模型, 可得预测值:

$$\begin{cases} \hat{x}^{(0)}(1) = x^{(1)}(1) = x^{(0)}(0) \\ \hat{x}^{(0)}(k+1) = x^{(1)}(k+1) - x^{(1)}(k) \end{cases} \quad (k=0, 1, 2, \dots, n) \quad (13)$$

这样就得到 GM(1, 1) 的预测模型。

由于 GM(1, 1) 模型是由现在时刻(即时间原点)向未来时刻进行预测的, 未来时刻越远, 预测值灰区间就越大, 模型对问题的描述将因时间的逐渐外推而逐渐失真。因此, 用 GM(1, 1) 模型进行短期预测时, 预测结果较好, 当进行长期预测时效果较差, 仅反映一种趋势。原因是模型中灰参数估计值一旦确定, 则不再改变, 忽视了它的动态时变的性质。因此对模型修改如下: 先用已知数列建立的 GM(1, 1) 模型预测一个值, 并将这个预测值补充到已知数列的最后位置上, 同时去掉最老的一个数据, 使序列等维。接着再建立 GM(1, 1) 模型, 逐个滚动预测, 依次递补, 直到完成预测目标。由于考虑到与预测期接近的数据更有效, 因此该模型预测精度比全数据 GM(1, 1) 模型有显著提高^[8]。

灰色模型通常用后验方法检验^[9], 设序列:

$$\hat{x}^{(0)} = \{x^{(0)}(2), \dots, x^{(0)}(n)\}$$

计算序列残差得:

$$e(k) = \left| x^{(0)}(k) - \hat{x}^{(0)}(k) \right|, k=1, 2, \dots, n$$

得残差向量:

$$e = e(1), e(2), \dots, e(n) \quad (14)$$

设定原始列 $x^{(0)}$ 及残差数列 e 的方差分别为 s_1^2, s_2^2 , 则有:

$$s_1^2 = \frac{1}{n} \sum_{k=1}^n (x^{(0)}(k) - \bar{x}^{(0)})^2 \quad (15)$$

$$s_2^2 = \frac{1}{n} \sum_{k=1}^n (e(k) - \bar{e})^2 \quad (16)$$

其中: $\bar{x}^{(0)} = \frac{1}{n} \sum_{k=1}^n x^{(0)}(k), \bar{e} = \frac{1}{n} \sum_{k=1}^n e(k)$

计算后验差比值 $C = \frac{s_2}{s_1} \quad (17)$

计算小误差概率 $P = P\{|e(k) - \bar{e}| < 0.674 5s_1\} \quad (18)$

3.3 非线性优化组合预测模型的建立

将指数回归模型与灰色预测模型进行加权组合, 权系数是以方差最小作为优化标准而求得。通过非线性优化组合, 综合了两类趋势预测模型的优点, 使得预测模型信息更加全面, 预测结果更加可靠。设指数回归模型得到的预测值为 y_1 , 预测误差为 e_1 ; 灰色预测模型得到的预测值为 y_2 , 预测误差为 e_2 。

设 $y = w_1 y_1 + w_2 y_2, (w_1 + w_2 = 1)$, 则预测误差 $e_c = w_1 e_1 + w_2 e_2$ 。

$$Var(e_c) = w_1^2 Var(e_1) + w_2^2 Var(e_2) + 2w_1 w_2 cov(e_1, e_2) \quad (19)$$

关于 w_1 , 对 $Var(e_c)$ 求极小值得:

$$w_1 = \frac{var(e_2) - cov(e_1, e_2)}{Var(e_1) + Var(e_2) - 2cov(e_1, e_2)} \quad (20)$$

$$\text{设 } Var(e_1) = E(e_1^2) - E(e_1)^2 = \sigma_{11}$$

$$Var(e_2) = E(e_2^2) - E(e_2)^2 = \sigma_{22}$$

$$cov(e_1, e_2) = E(e_1 \cdot e_2) - E(e_1) \cdot E(e_2) = \sigma_{12}$$

则 $w_1 = \frac{\sigma_{22} - \sigma_{12}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} \quad (21)$

$$w_2 = \frac{\sigma_{11} - \sigma_{12}}{\sigma_{11} + \sigma_{22} - 2\sigma_{12}} \quad (22)$$

计算可得: $\lim_{\sigma_{11} \rightarrow \infty} w_1 = 0, \lim_{\sigma_{22} \rightarrow \infty} w_1 = 1, \lim_{\sigma_{12} \rightarrow \infty} w_1 = 1, \lim_{\sigma_{12} \rightarrow \infty} w_1 = 0$ 。

可以看出, 预测值 y_1 越可靠, y_1 的权数值也会相应增大。

同时可以证明: $Var(e_c)_{\min} \leq \sigma_{11}, Var(e_c)_{\min} \leq \sigma_{22}, Var(e_c)_{\min} \leq \min(\sigma_{11}, \sigma_{22})$, 这表明优化组合方法优于单一的方法。

4 中国油气资源生产发展预测

中国石油、天然气资源生产及新增探明存量的历史数据见表 1。

表 1 中国石油、天然气生产历史数据统计表

| 时间 | 石油年产量/万吨 | 石油探明储量增长率/亿吨 | 天然气年产量/亿立方米 | 天然气探明储量增长率/亿吨 |
|------|----------|--------------|-------------|---------------|
| 1960 | 420 | 2.85 | 10.39 | 58.16 |
| 1961 | 482 | 20.62 | 14.72 | 58.16 |
| 1962 | 527 | 0.05 | 12.05 | 58.16 |
| 1963 | 595 | 0.00 | 10.42 | 143.2 |
| 1964 | 794 | 0.66 | 10.57 | 143.2 |
| 1965 | 1 076 | 4.78 | 11.13 | 143.2 |
| ... | ... | ... | ... | ... |
| 2000 | 16 100 | 7.04 | 223 | 2 478.19 |
| 2001 | 16 000 | 7.94 | 245 | 1 149.41 |
| 2002 | 16 300 | 7.25 | 272 | 4 951.69 |
| 2003 | 16 500 | 4.69 | 303 | 4 421.71 |
| 2004 | 16 700 | 8.00 | 326.3 | 3 687.63 |
| 2005 | 16 988 | 8.64 | 341.28 | 4 968.55 |

利用公式(1)至公式(6)建立指数预测回归模型, 利用公式(7)计算预测结果的标准离差。

石油年产量: $\hat{E}_t = 11 482.05(1.036 4)^t$, 标准离差: $\frac{S}{y} = 12.93\%$ 。

石油探明储量增长率: $\hat{E}_t = 4.607 3(1.042 3)^t$, 标准离差:

$$\frac{S}{y} = 6.81\%$$

天然气年产量: $\hat{E}_t = 131.345 7(1.051 7)^t$, 标准离差: $\frac{S}{y} =$

13.76%。

天然气探明储量增长率: $\hat{E}_t = 444.732 2(1.139 2)^t$, 标准离差: $\frac{S}{y} = 14.86\%$ 。

利用公式(8)至公式(13)建立等维递补灰色预测模型 GM(1, 1), 利用公式(14)至公式(18)计算预测模型的后验差比值和小误差概率。

石油年产量: $X^{(1)}(k+1) = 890 841.9e^{0.015 78k} - 877 010.9$, 后验差比值 $C = 0.023 3$, 小误差概率 $P = 1$ 。

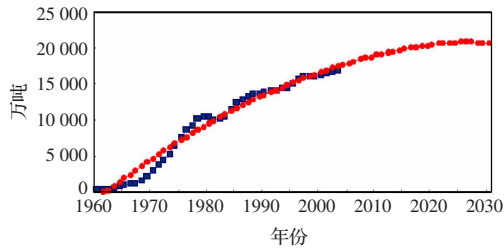


图1 中国石油年产量增长趋势

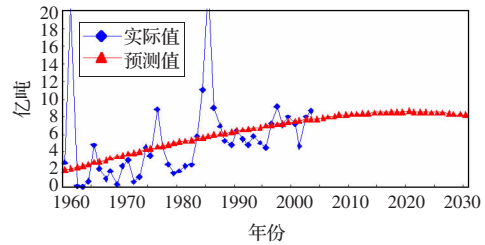


图2 中国石油探明储量增长趋势

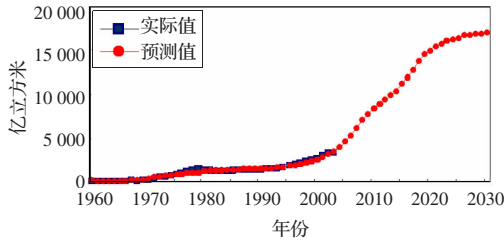


图3 中国天然气产量增长趋势

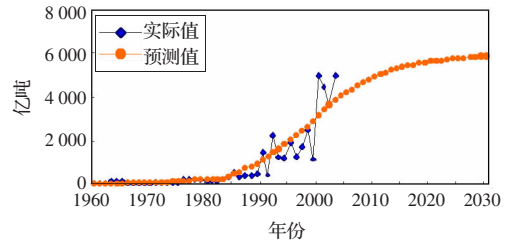


图4 中国天然气探明储量增长趋势

石油探明储量增长量: $X^{(1)}(k+1)=146.474e^{0.035406k}-140.074$, 后验差比值 $C=0.2567$, 小误差概率 $P=1$ 。

天然气年产量: $X^{(1)}(k+1)=1750.993e^{0.07576k}-1608.993$, 后验差比值 $C=0.2125$, 小误差概率 $P=1$ 。

天然气探明储量增长量: $X^{(1)}(k+1)=8192.0274e^{0.1209k}-6743.897$, 后验差比值 $C=0.45$, 小误差概率 $P=0.75$ 。

$$x^{(0)}(k+1)=x^{(1)}(k+1)-x^{(1)}(k)$$

利用式(19)~式(22)建立非线性优化组合预测模型。

$$\text{石油年产量: } f=0.3514f_1+0.6486f_2$$

$$\text{石油探明储量增长量: } f=0.6327f_1+0.3673f_2$$

$$\text{天然气年产量: } f=0.5879f_1+0.4121f_2$$

$$\text{天然气探明储量增长量: } f=0.7732f_1+0.2278f_2$$

式中: f 为组合预测值; f_1 为指数回归预测; f_2 为灰色预测。

预测结果如图1至图4。

5 结论

所提出的非线性指数回归与灰色预测优化组合预测模型, 通过对能源生产历史数据不同变化规律的分析, 充分利用了灰色预测建模所需信息少、方法简单的优点和非线性指数回归具有较强的非线性映射能力的特性, 用最优组合权系数法进行组合, 得到了预测精度较高且方便使用的组合预测模型。对中国

未来油气能源的生产发展状况进行了预测分析, 实例表明, 该模型可以作为能源生产发展预测的有效工具。

参考文献:

- [1] 刑棉. 能源发展趋势的非线性优化组合预测模型研究[J]. 华北电力大学学报, 2002, 29(3): 64-67.
- [2] 朱康福. 对中国2020年能源消费及结构中石油比例的再思考[J]. 当代石油石化, 2003, 11(6): 15-17.
- [3] 周总瑛, 唐跃刚. 中国天然气资源特点及其在能源结构中的地位分析[J]. 中国矿业, 2004, 13(6): 17-28.
- [4] 孙永富, 王良曦, 庞磊. GM(1,1)模型与指数回归模型的研究[J]. 装甲兵工程学院学报, 2000, 14(1): 76-79.
- [5] Liu Si-feng, Forrest J. The role and position of grey system theory in science development[J]. The Journal of Grey System(UK), 1997, 9(4): 351-356.
- [6] Dang Yao-guo, Liu Si-feng, Chen Ke-jia. The GM model that $x(n)$ be taken as initial value[J]. Journal of Xiamen University: Natural Science, 2002, 41(Sup.): 276-277.
- [7] Liu Si-feng, Deng Jun-long. The range suitable for GM(1,1)[J]. The Journal of Grey System(UK), 1999, 11(1): 131-138.
- [8] 刘凤朝, 潘雄锋, 王元地. 基于灰色系统理论的中国专利分析与预测[J]. 情报杂志, 2004(12): 53-55.
- [9] 赵金环, 陈建强, 郭立智. Excel在动态GM(1,1)模型预测及检验中的应用[J]. 青岛理工大学学报, 2007, 28(4): 69-73.
- [5] Liang H L, Bressler S L. Empirical mode decomposition: A method for analyzing neural data[J]. Neurocomputing, 2005: 65-66, 801-807.
- [6] 熊学军, 郭焰火, 胡筱敏, 等. EMD方法和Hilbert谱分析法的应用与探讨[J]. 海洋科学进展, 2002, 20(2): 12-21.
- [7] 毛炜, 金荣洪, 耿军平, 等. 一种基于改进Hilbert-Huang变换的非平稳信号时频分析法及其应用[J]. 上海交通大学学报, 2006, 40(5): 724-729.
- [8] 张晓玲. Hilbert-Huang变换在气候反演模型中的应用[D]. 石家庄: 河北师范大学, 2007.
- [9] 张瑞, 汪亚平, 潘少明. 长江大通水文站径流量的时间序列分析[J]. 南京大学学报: 自然科学版, 2006, 42(6): 423-434.

(上接 207 页)

参考文献:

- [1] Huang N E, Shen Z. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis[J]. Proc R Socland A, 1998, 454(1): 903-995.
- [2] 赵雪花. 河川径流演变规律的挖掘与识别技术[D]. 西安: 西安理工大学, 2005.
- [3] 吴湘淇. 信号、系统与信号处理[M]. 北京: 电子工业出版社, 1996.
- [4] 谭善文. 多分辨希尔伯特-黄(Hilbert-Huang)变换方法的研究[D]. 重庆: 重庆大学, 2001.