



基于混沌理论和 Volterra 自适应滤波器的天然气负荷预测

乔伟彪^{①②③*}, 陈保东^②, 卢泓方^④

① 中国石油大学(华东)储运与建筑工程学院, 青岛 266555;

② 辽宁石油化工大学石油天然气工程学院, 抚顺 113001;

③ 华润(南京)市政设计有限公司, 南京 210000;

④ 西南石油大学石油与天然气工程学院, 成都 610500

* E-mail: stevenwbq@sina.com

收稿日期: 2014-08-12; 接受日期: 2014-10-24

中国石油集团公司重点研究项目(批准号: KY2011-13)和国家自然科学基金(批准号: 51474187)资助项目

摘要 天然气是一种重要的清洁能源, 如何预知未来时刻的天然气负荷, 对于燃气公司与上游气源及下游用户制定合适的商业计划具有重要的意义。针对天然气短期负荷预测问题, 提出基于混沌理论和 Volterra 自适应滤波器的预测模型。首先, 对天然气时负荷时间序列进行日相关性分析, 然后采用互信息法和伪最近邻域法确定延迟时间和最佳嵌入维数; 其次, 在相空间重构的基础上, 对天然气时负荷时间序列进行混沌特性分析; 再次, 针对现有预测模型多为主观预测模型的特点, 将 Volterra 自适应滤波器预测模型引入到天然气时负荷预测中, 降低了人为主观性; 最后, 给出预测算例, 探讨不同 Volterra 自适应滤波器阶数对预测效果的影响及对比了 Volterra 自适应滤波器预测模型与人工神经网络(*artificial neural network, ANN*)、傅里叶级数预测模型的预测精度, 验证二阶 Volterra 自适应滤波器预测模型在不同日时负荷的预测效果。结果表明: 二阶 Volterra 自适应滤波器预测模型较 ANN、傅里叶级数预测模型具有更高的预测精度, 为天然气短期负荷的在线工程应用提供了有益参考。

关键词

天然气负荷

混沌

Volterra 自适应滤波器

ANN

傅里叶级数

预测

1 引言

随着我国天然气工业的大力发展, 燃气企业的市场化运作, 作为燃气企业重要组成部分的短期负荷预测受到越来越多的重视。近些年来, 国家积极提倡节能减排, 这给天然气短期负荷预测的预测精度提出了更高的要求。天然气短期负荷是一个复杂的

非线性系统^[1], 系统本身具有伪周期性, 随机性和趋势性等特点。此外天然气短期负荷也与温度等多种气象自然因素息息相关, 同时还受重大事件及国家法定节假日的影响。因此, 这些都决定了很难建立起具体的天然气短期负荷预测数学模型。

现有的天然气短期负荷预测方法主要包括两种: 一种是物理方法^[2], 其利用数值天气预报的结果, 如

引用格式: 乔伟彪, 陈保东, 卢泓方. 基于混沌理论和 Volterra 自适应滤波器的天然气负荷预测. 中国科学: 技术科学, 2015, 45: 91–102

Qiao W B, Chen B D, Lu H F. Gas load forecasting based on chaotic theory and Volterra adaptive filter (in Chinese). Sci Sin Tech, 2015, 45: 91–102.
doi: 10.1360/N092014-00202

温度、湿度、风速、云量等物理信息建立刻画天然气短期负荷的模型进行预测; 另一种是统计方法, 即根据过往的历史样本数据建立起系统输入与输出的非线性映射关系来进行预测, 如小波分析综合模型^[3]、GM(1,1)灰色理论模型^[4]、人工神经网络(ANN)模型^[5]、时间序列模型^[6]等。物理方法不需要大量实测的历史样本数据, 但预测误差较大。统计方法共同的特点是需要预先建立时间序列的主观预测模型, 然后根据建立的主观预测模型进行计算及预测。随着混沌(chaos)科学的发展, 使得可不必预先建立主观预测模型, 而是直接根据数据样本时间序列本身所计算出的客观规律进行预测, 这样完全可以避免预测的人为主观性, 提高预测精度和可信度。我们知道, 对于某一混沌系统来说, 可采用混沌预测模型来进行预测, 并且能够获得较高的预测精度。因此, 我们设想采用混沌理论的方法来对天然气短期负荷进行建模和预测。基于此思想, 首要解决的问题是确定天然气短期负荷是否具有混沌特性。

近些年来, 混沌时间序列建模及预测已经成为混沌研究领域中的热点研究课题。文献[7]利用混沌理论对城市天然气日负荷进行了混沌特性识别, 采用最大 Lyapunov 指数法等方法对其进行了预测, 取得了很好的预测效果。但对天然气时负荷采用混沌理论进行预测研究还非常少见, 因此本文采用混沌理论对天然气时负荷进行预测研究。目前预测混沌时间序列的模型包括全局预测模型、局域预测模型、Volterra 自适应预测模型。全局预测模型是将全部轨迹中的点作为研究对象, 并进行拟合, 所以采用该模型进行预测, 预测精度相对较低。因此, 该预测模型在理论上具有一定的可行性^[8]。局域预测模型通常分为最大 Lyapunov 指数法^[9]、加权一阶局域法^[10]、一阶局域法^[11]、加权零阶局域法^[12]和零阶局域法^[13]预测模型等。加权一阶局域法和一阶局域法是在加权零阶局域法和零阶局域法发展起来的, 相比加权零阶局域法和零阶局域法, 加权一阶局域法和一阶局域法具有较高的预测精度。一阶局域法的思想是在预测过程中相互邻近的点对预测结果产生的影响是一样的, 但加权一阶局域法的思想是根据不同相互邻近的点距离远或近赋不一样的权值, 由于加权一阶局域法在预测过程中会增加计算工作且产生误差累积的缺点, 预测精度不高。Volterra 自适应滤波器预测模型是近些年发展起来的一种预测模型^[14], 分

为多步和单步预测。Volterra 自适应滤波器预测模型在预测过程中仅需很少的数据样本就可对混沌时间序列进行预测, 而且可自动追踪混沌运动轨迹, 该预测模型的预测精度较高。但文献[15]指出 Volterra 自适应滤波器预测模型的单步预测精度要高于多步预测精度, 因此本文试图利用单步 Volterra 自适应滤波器预测模型的优势, 用于天然气时负荷预测研究中, 以期获得天然气时负荷预测的新方法和新思路。

2 天然气时负荷数据序列的日相关性分析

2.1 样本预处理

本文中使用的样本数据来自江西省某市天然气门站实际采集的计量数据。在样本数据库中, 选取时间段为 2013 年 6 月 1 日早 8 点至 2013 年 6 月 30 日晚 8 点, 连续的 720 个数据样本点作为研究对象。首先采用小波软阈值去噪算法对样本数据进行降噪处理。为了计算方便, 将去噪后的样本数据进行归一化处理, 归一化后的样本数据如图 1 所示。下面对去噪后的样本数据进行日相关性分析。

2.2 日相关性分析

文献[16]利用小波理论-BP 神经网络对短期风电功率进行预测时, 要求短期风电功率具有明显的日相关性, 这样可提高预测精度。因此可以设想如果天然气时负荷时间序列具有日相关性, 在此基础上进行天然气时负荷的预测, 预测精度就会提高, 但其受温度等环境因素的影响, 很难保证具有日相关性, 因

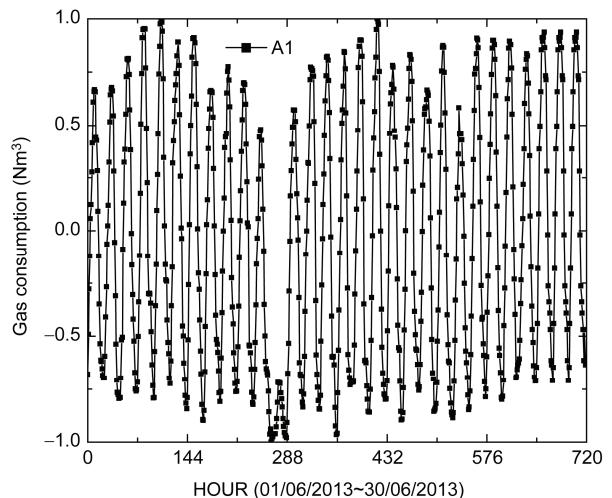


图 1 归一化后的时负荷

此首先运用相关性理论对天然气时负荷进行日相关性分析.

相关函数 r_k 是反应变量 $x(t, k)$ 及其下一周期 $x(t, k+T)$ 间线性相关的函数, 当 k 取 $T, 2T, \dots, kT$ 时, r_k 代表在不同周期的相关函数值. 判断天然气时负荷时间序列是否具有周期相关性, 主要是判断相关函数值波动是否平稳. 波动平稳, 可认为在一段时间内, 各周期数值具有较强的相关性; 波动剧烈, 可认为在一段时间内, 各周期数值不具有相关性.

定义:

$$\hat{C}_k = \frac{1}{T} \sum_{t=1}^T (x_{t,k} - \bar{x}_{t,k})(x_{t,k+T} - \bar{x}_{t,k+T}), \quad (1)$$

$$\hat{\sigma}_{t,k} = \left[\frac{1}{T} \sum_{t=1}^T (x_{t,k} - \bar{x}_{t,k})^2 \right]^{1/2}, \quad (2)$$

$$\hat{\sigma}_{t,k+T} = \left[\frac{1}{T} \sum_{t=1}^T (x_{t,k+T} - \bar{x}_{t,k+T})^2 \right]^{1/2}, \quad (3)$$

则

$$r_k = \frac{\hat{C}_k}{\hat{\sigma}_{t,k} \cdot \hat{\sigma}_{t,k+T}}, \quad (4)$$

类周期函数定义如下:

$$f(x) = 0.5(0.4 \sin \pi x / 5 + 0.6B(x) + 1), \quad (5)$$

式中, $B(x)$ 为 $0 \sim 1$ 均匀分布的随机量.

类周期序列的采样频率为 50, 共采集 20 个周期数据序列. 天然气时负荷时间序列数据及类周期数据在周期内的数据个数分别为 24 和 100, 类周期序列如图 2 所示. 图 3 为相邻每个周期的天然气时负荷时间序列和类周期序列的相关函数值. 从图 3 可以看出: 在 20 个周期内, 具有相似周期特性的类周期序列的相关函数值变化几乎不大, 而天然气时负荷时间序列在每个周期内的相关函数值变化较大, 主要在 $-0.75 \sim 0.97$ 变化. 但相邻的大部分日周期函数值或为正或为负, 因此具有日相似特性, 然而也有第 6 点与第 7 点, 第 7 点与第 8 点之间的相关函数值出现了正负或者是负正等较大的变化, 由此可知天然气时负荷不是每日都具有非常明显的日相似关系, 特殊情况偶尔也会出现. 考虑这些特殊情况, 可能带来的挑战, 所以本文试图采用混沌理论对天然气时负荷时间序列进行预测成为一种可能的选择.

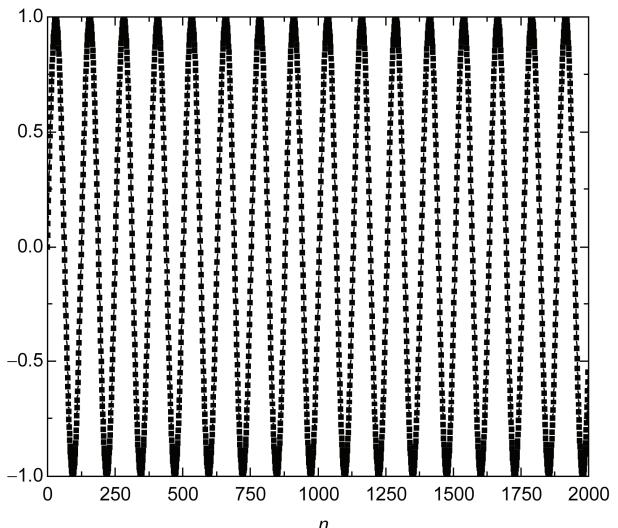


图 2 类周期序列

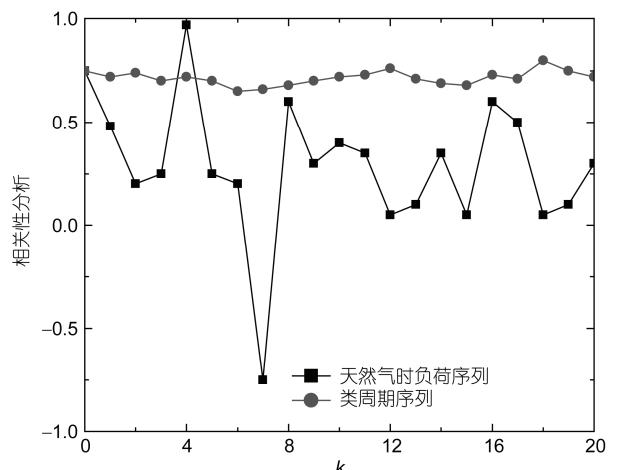


图 3 序列相邻周期的相关函数值

3 相空间重构及混沌特性分析

3.1 相空间重构

Takens 定理指出可以通过单一的混沌时间序列来构造一组表征原动力学系统特性的坐标分量, 使其在拓扑意义上等价的相空间, 从而准确把握混沌时间序列的规律和性质.

重构相空间^[17]的原理:

设时间序列为 $\{x(i), i=1, 2, \dots, N\}$, 通过延迟时间构成 m 维向量: $X(t) = [x(t), x(t+\tau), \dots, x(t+(m-1)\tau)]$, $t=1, 2, \dots, M$, 其中 m 为最佳嵌入维数, τ 为

延迟时间, $X(t)$ 为 m 维相空间中的相点, M 为相点个数, $M=N-(m-1)\tau$. 集合 $\{X(t), t=1, 2, \dots, M\}$ 描述了动力系统在相空间中的演化轨迹, 由此可在重构的 m 维相空间中研究动力学系统的混沌行为. 研究表明: 只要最佳嵌入维数 m 和延迟时间 τ 选择恰当, 重构的相空间则与原系统在拓扑意义上具有相同的性质.

延迟时间 τ 的确定方法有互信息法、平均位移法、自相关函数法等. 本文采用互信息法^[18]确定延迟时间 τ . 伪最近邻域法在最佳嵌入维数 m 的选取上具有很好的效果, 本文选用伪最近邻域法确定最佳嵌入维数^[19].

3.1.1 延迟时间的确定

假设时间序列为 $\{x(i), i=1, 2, \dots, N\}$, 根据 Takens 定理进行延时重构, $X(i)=[x(i), x(i+\tau), \dots, x(i+N\tau)]$.

$x(k)$ 在时间序列 $\{x(i), i=1, 2, \dots, N\}$ 中出现的概率定义为 $P(x(k))$, $X(k)$ 在时间序列 $\{X(i), i=1, 2, \dots, N\}$ 中出现的概率定义为 $P(X(k+\tau))$, $P(x(k))$ 和 $P(X(k+\tau))$ 在两个时间序列中出现的联合概率定义为 $P(x(k), X(k+\tau))$, 其中 $P(x(k))$ 和 $P(X(k+\tau))$ 可分别通过在各自时间序列中出现的概率来求解, 联合概率 $P(x(k), X(k+\tau))$ 可通过在平面 $(x(k), X(k+\tau))$ 上数格子得到, 则求互信息的互信息函数为

$$I(\tau) = \sum_{i=1}^N P(x(k), X(k+\tau)) \times \log_2 \frac{P(x(k), X(k+\tau))}{P(x(k)) \cdot P(X(k+\tau))}. \quad (6)$$

延迟时间 τ 的确定为当互信息函数 $I(\tau)$ 取第一个极小值点时的 τ , 因为此时产生的冗余最小, 并具有最大独立性. 互信息-延迟时间曲线如图 4 所示. 互信息与延迟时间的对应关系如表 1 所示. 由此可知: 当 τ 为 13 时, 互信息函数取得第一个局部极小值点, 因此天然气短期负荷序列延迟时间为 $\tau=13$ h.

3.1.2 最佳嵌入维数的确定

伪最近邻域法的基本原理: 当最佳嵌入维数比较小时, 混沌吸引子在相空间中不能完全展开, 混沌吸引子就会在其相空间中产生部分投影点, 相空间中的其他点与投影点形成最近邻域点, 而这两点在两空间中, 因此这两点并非真正的最邻近点, 称之为

表 1 天然气短期负荷时间序列互信息与延迟时间对应表

延迟时间	互信息	延迟时间	互信息
1	7.5136	26	2.7301
2	5.0125	27	2.7484
3	3.1256	28	2.7465
4	2.9312	29	2.7345
5	2.8458	30	2.7402
6	2.8212	31	2.7396
7	2.8121	32	2.7285
8	2.8112	33	2.7358
9	2.8100	34	2.7346
10	2.8085	35	2.7221
11	2.8065	36	2.7302
12	2.8045	37	2.7286
13	2.7801	38	2.7164
14	2.7900	39	2.7212
15	2.7794	40	2.7185
16	2.7886	41	2.7166
17	2.7721	42	2.7003
18	2.7710	43	2.7086
19	2.7700	44	2.6945
20	2.7554	45	2.7012
21	2.7621	46	2.7001
22	2.7612	47	2.6945
23	2.7400	48	2.6812
24	2.7584	49	2.6900
25	2.7525	50	2.6845

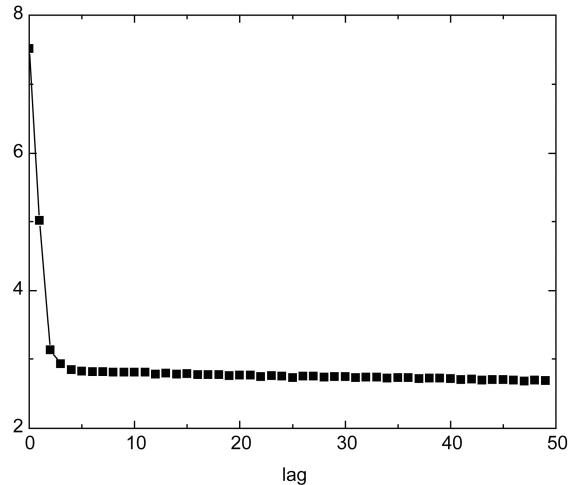


图 4 互信息-延迟时间曲线

伪最邻近点.

伪最近邻域法的计算过程如下:

- 1) 在嵌入维数为 m 的嵌入空间中, 找出所有点

的欧氏距离最邻近点, 任意 2 点 $X(i)$ 和 $X(j)$ 之间的欧式距离为

$$\|X(i), X(j)\| = \left[\sum_{l=0}^{m-1} (x(i+l\tau) - x(j+l\tau))^2 \right]^{1/2}, \quad (7)$$

2) 当任意一对最邻近点满足如下判据时, 则为伪最邻近点.

$$\begin{aligned} \left[\frac{R_{m+1}^2(i, j) - R_m^2(i, j)}{R_m^2(i, j)} \right]^{1/2} &= \frac{|x(i+m\tau) - x(j+m\tau)|}{R_m(i, j)} \\ &\geq R_{\text{tol}}, \end{aligned} \quad (8)$$

式中, $R_{m+1}^2(i, j)$, $R_m^2(i, j)$ 分别表示在最佳嵌入维数分别为 $m+1$ 和 m 时任意一对最邻近点之间的距离平方, R_{tol} 为门限值, 一般可以取 [10, 50], 本文取 $R_{\text{tol}}=16$.

3) 当 $m=1$ 时, 计算伪最邻近点与相点总数的比例, 然后逐渐增加 m , 直到伪最邻近点的比例小于 5% 时, 可认为混沌吸引子几何结构完全打开^[19], 此时的 m 为最佳嵌入维数.

取延迟时间 $\tau=13$ h, 根据文献[15]中的方法, 嵌入维数 m 从 1 取到 16. 计算结果如表 2 及图 5 所示. 当嵌入维数 $m=4$ 时, 伪最邻近点的比例小于 5%, 因此天然气短期负荷时间序列的最佳嵌入维数为 $m=4$.

3.2 混沌特性分析

3.2.1 小数据量法计算最大 Lyapunov 指数

判断动力系统是否具有混沌特性, 可以通过分析系统是否对初始值敏感, 而最大 Lyapunov 指数则是表征这一特性的特征量, 是识别系统混沌特性的重要依据之一.

1) 最大 Lyapunov 指数小于 0, 则说明系统具有稳定的不动点.

2) 最大 Lyapunov 指数等于 0, 则说明该系统具

表 2 天然气时负荷时间序列伪最邻近点百分比与最佳嵌入维数对应表

最佳嵌入维数	时负荷	最佳嵌入维数	时负荷
1	98.1242	9	3.1245
2	60.2514	10	3.2546
3	20.0145	11	3.4521
4	4.9654	12	3.2541
5	3.1245	13	3.3245
6	3.2451	14	3.1278
7	3.6524	15	3.3214
8	3.3214	16	3.1243

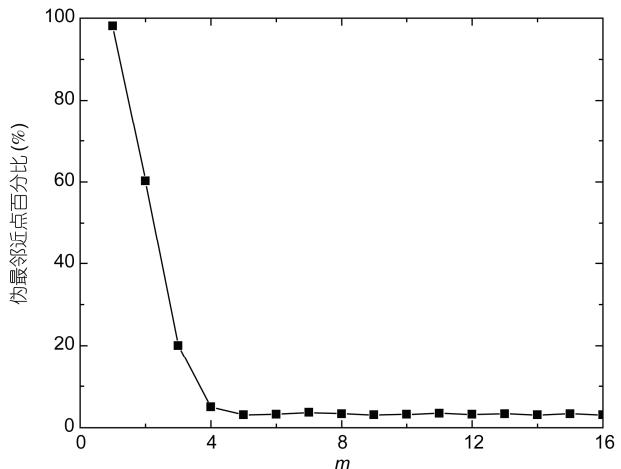


图 5 伪最邻近点-嵌入维数曲线

有周期解.

3) 最大 Lyapunov 指数大于 0, 则说明该系统具有混沌特性.

计算最大 Lyapunov 指数, 早些年主要是利用 Wolf 法^[20], 之后一种适用于小数据样本, 抗噪声能力较强, 计算速度较快的计算方法被 Rosenstein 提出, 称之为小数据量法^[21]. 小数据量法计算过程如下:

相空间重构后, 寻找特定的轨道上每个状态点 $X(j)$ 的最邻近点 $X(\hat{j})$, 即

$$d_j(0) = \min_{X(\hat{j})} \|X(j) - X(\hat{j})\|, |j - \hat{j}| \geq p, \quad (9)$$

式中, $j = N_0, N_0 + 1, \dots, N$, $N_0 = (m-1)\tau + 1$, p 表示时间序列的平均周期, 应用快速傅里叶变换来计算, $d_j(0)$ 表示在初始时刻一对最邻近点之间的距离, $X(j)$ 为重构相空间中的状态点. Sato 等人估计最大 Lyapunov 指数为

$$\lambda_1(i) = \frac{1}{i \cdot \Delta t} \cdot \frac{1}{N-i} \sum_{j=1}^{N-i} \ln \frac{d_j(i)}{d_j(0)}, \quad (10)$$

式中, $i = N_0, N_0 + 1, \dots, N$, Δt 表示样本数据周期, $d_j(i)$ 为基本轨道上的第 j 对最邻近点对经过 i 个离散时间步长后的距离. 之后 Sato 等人对其进行改进, 改进后的估计表达式为

$$\lambda_1(i, k) = \frac{1}{k \cdot \Delta t} \cdot \frac{1}{N-k} \sum_{j=1}^{N-k} \ln \frac{d_j(i+k)}{d_j(i)}, \quad (11)$$

式中, k 为一常数. 最大 Lyapunov 指数的几何意义量化初始闭轨道散发和估计系统的总体混沌水平的量, 结合 Sato 等人的估计式则有

$$d_j(i) \approx C_j e^{\lambda_j(\Delta t)}, \text{ 且 } C_j = d_j(0). \quad (12)$$

将上式两边取倒数得

$$\ln d_j(i) \approx \ln C_j + \lambda_j(i \cdot \Delta t). \quad (13)$$

最大 Lyapunov 指数可以近似地看成公式所代表的一组直线斜率, 可以利用最小二乘法逼近这组直线来得到, 即

$$y(i) = \frac{1}{w\Delta t} \sum_{j=1}^w \ln [d_j(i)], \quad (14)$$

式中, w 为非零 $d_j(i)$ 的数目, $y(i)$ 为距离 $d_j(i)$ 对 w 累积和的平均值.

采用小数据量法对天然气时负荷时间序列求最大 Lyapunov 指数, 演化步数 i 取 80, 距离对数平均值 $y(i)$ 与步数 i 关系如图 6 所示. 求得的最大 Lyapunov 指数为 0.1455, 大于 0, 所以验证了天然气时负荷时间序列具有明显的混沌特性, 即为典型的混沌时间序列. 因此可以利用混沌预测模型对天然气时负荷时间序列进行预测.

4 预测模型

4.1 Volterra 自适应滤波器基本理论

文献[14]建立了混沌时间序列的非线性自适应混沌预测技术的初步框架, 目前已经得到了广泛的应用^[22,23], 本文基于天然气时负荷时间序列具有混沌特性的这一特点, 将 Volterra 自适应滤波器预测模型应用到天然气负荷预测领域, 对天然气短期负荷预测进行应用研究, 以期得到天然气短期负荷预测的新思路和新方法. 下面对 Volterra 自适应滤波器

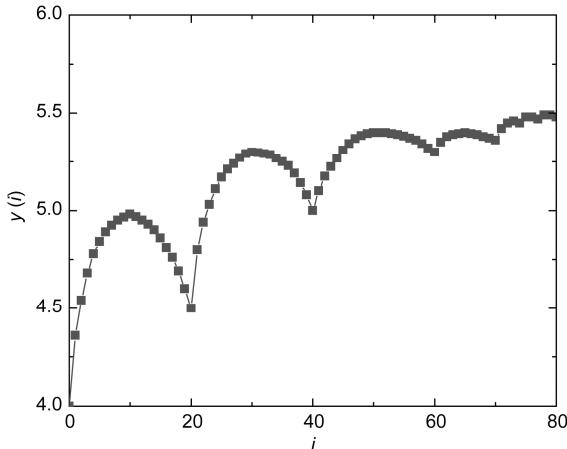


图 6 $y(i)$ 随 i 变化的关系曲线图

基本理论进行介绍.

设非线性动力离散系统的输入为 $x(t)$ ($t=1, 2, \dots, M$), 输出为 $Y(t) = \hat{x}(t+1)$ ($t=1, 2, \dots, M$), 则该非线性系统的 Volterra 级数展开式为

$$\begin{aligned} \hat{x}(t+1) &= F[x(t)] \\ &= h_0 + \sum_{m=0}^{+\infty} h_1(m)x(t-m) \\ &+ \sum_{m_1=0}^{+\infty} \sum_{m_2=0}^{+\infty} h_2(m_1, m_2)x(t-m_1)x(t-m_2) + \dots \\ &+ \sum_{m_1=0}^{+\infty} \sum_{m_2=0}^{+\infty} \dots \sum_{m_p=0}^{+\infty} \left[h_p(m_1, m_2, \dots, m_p)x(t-m_1) \right. \\ &\quad \left. \times x(t-m_2) \dots x(t-m_p + \dots) \right], \end{aligned} \quad (15)$$

式中, $h_p(m_1, m_2, \dots, m_p)$ 为 p 阶 Volterra 核, m 为滤波器的输入维数. Volterra 无穷级数展开式在实际中难以实现, 必须采用有限截断和有限求和的形式. 以二阶截断求和为例, 其形式为

$$\begin{aligned} \hat{x}(t+1) &= F[x(t)] \\ &= h_0 + \sum_{m=0}^{N_1} h_1(m)x(t-m) \\ &+ \sum_{m_1=0}^{N_2-1} \sum_{m_2=0}^{N_2-1} h_2(m_1, m_2) \times x(t-m_1)x(t-m_2), \end{aligned} \quad (16)$$

式中表示的非线性 Volterra 自适应滤波器结构如图 7 所示. 非线性扩展后的信号为

$$U(t) = [u(t), u(t-1), \dots, u(t-L-1)]^T, \quad (17)$$

式中, $U(t)$ 为滤波器的输入矢量, 并且 $L \geq N_1, N_2$.

在应用设计中, N_1 和 N_2 应为有限长, 根据 Takens 的嵌入定理, 可将 N_1 和 N_2 分别取为 $N_1=N_2=m \geq 2D_2+1$, 这里 D_2 表示关联维数. 因此, 混沌时间序列预测的滤波器为

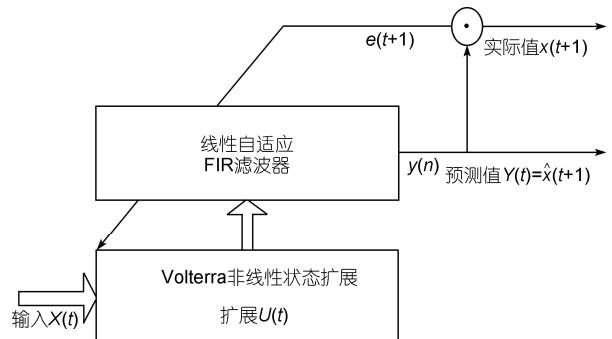


图 7 非线性 Volterra 自适应滤波器结构

$$\begin{aligned}\hat{x}(t+1) &= h_0 + \sum_{i=0}^{m-1} h_1(i)x(t-i) \\ &+ \sum_{i=0}^{m-1} \sum_{j=i}^{m-1} h_2(i,j)x(t-i) \times x(t-j),\end{aligned}\quad (18)$$

其状态扩展后的系数总个数为

$$L = 1 + m + m(m+1)/2.$$

线性自适应有限脉冲响应滤波器的输入矢量 $\mathbf{U}(t)$ 定义如下:

$$\begin{aligned}\mathbf{U}(t) &= [1, x(t), x(t-1), \dots, x(t-m+1), \\ &x^2(t), x(t)x(t-1), \dots, x^2(t-m+1)]^T,\end{aligned}\quad (19)$$

系数向量为

$$\begin{aligned}\mathbf{H}(t) &= [h_0, h(0), h(1), \dots, h(m-1), h_2(0,0), \\ &h_2(0,1), \dots, h_2(m-1, m-1)]^T.\end{aligned}\quad (20)$$

则(18)式可表示为

$$\hat{x}(t+1) = \mathbf{H}^T(n)\mathbf{U}(n). \quad (21)$$

(18)式表示的 Volterra 自适应滤波器系数可以直接利用线性有限脉冲响应自适应滤波器的自适应算法来计算得到。对于(21)式这种二阶 Volterra 自适应滤波器, 可采用的自适应算法为时间正交自适应算法, 对于系数矢量 $\mathbf{H}(t)$, 输入矢量 $\mathbf{U}(t)$, 时间正交自适应算法可以表示为

$$\begin{aligned}x(t) &= d(t), \\ \hat{x} &= \hat{d}(t), \\ \hat{d}(t) &= \mathbf{H}^T(t-1)\mathbf{U}(t-1), \\ \mathbf{H}(t) &= \mathbf{H}(t-1) + c \times \frac{e(t-1)}{\mathbf{U}^T(t)\mathbf{U}(t)} \mathbf{U}(t-1), \\ e(t) &= d(t) - \hat{d}(t),\end{aligned}\quad (22)$$

式中, c 为控制收敛性的参数。

4.2 天然气时负荷时间序列的 Volterra 预测模型

给定天然气时负荷时间序列的样本数据 $D = \{x(i), i = 1, 2, \dots, N\}$, 取上文计算出的最佳嵌入维数 $m=4$, 延迟时间 $\tau=13$, 按照重构相空间的方法, 就可以得到相空间域中的样本数据 $\bar{D} = \{X(t), Y(t)\}, t = 1, 2, \dots, M$, 其中 $X(t) = [x(t), x(t+\tau), \dots, x(t+(m-1)\tau)]$, $Y(t) = x(t+1+(m-1)\tau), t = 1, 2, \dots, M$. 表示成矩阵的

形式即

$$\begin{aligned}\mathbf{X} &= \begin{pmatrix} x(1) & x(1+\tau) & \cdots & x(1+(m-1)\tau) \\ x(2) & x(2+\tau) & \cdots & x(2+(m-1)\tau) \\ \cdots & \cdots & \cdots & \cdots \\ x(M) & x(M+\tau) & \cdots & x(M+(m-1)\tau) \end{pmatrix}, \\ \mathbf{Y} &= \begin{pmatrix} x(2+(m-1)\tau) \\ x(3+(m-1)\tau) \\ \cdots \\ x(M+1+(m-1)\tau) \end{pmatrix}.\end{aligned}$$

根据相空间中已知相点 $X(t)$ 来预测未来时刻相点 $Y(t)$ 的映射函数 F , 可由上文中的 Volterra 自适应滤波器来逼近, 使得

$$Y(t) = F[X(t)]. \quad (23)$$

5 算例分析

上文在重构相空间的基础上, 验证了天然气时负荷时间序列具有明显的混沌特性, 可采用 Volterra 自适应滤波器预测模型对天然气时负荷进行预测。一般的预测过程都需要较多的数据样本, 但在对天然气时负荷时间序列预测过程的研究中发现并不需要太多的样本, 否则会增加仿真的复杂程度和运算时间。根据多次实验, 可采用前 29 d 的数据来预测第 30 d 的数据。故选取前 29 d 的 657 个相点作为训练样本, 后 24 个相点作为测试样本。

为了加快 Volterra 自适应滤波器预测模型的收敛速度, 首先将数据样本转化为标么值, 并进行 $[-1, 1]$ 的归一化处理(如图 1 所示)。由于每次预测的负荷值均与其前 1 h 的负荷值相关, 只能进行单步预测, 而且文献[9]指出 Volterra 自适应滤波器预测模型的单步预测较多步预测具有较高的预测精度, 因此本文尝试只执行单步预测, 即将每次预测出的一个负荷值 $Y(i)$ 后, 将该负荷值填入到下一个输入 $X(i+1)$ 的最后一维中。需要执行迭代次数为 24 次, 才能够完成未来 24 h 负荷值的预测。

Volterra 自适应滤波器的阶数 p 和输入维数 m 对模型的预测精度有很大的影响, 但是目前仍没有统一的选择标准。文献[23]指出: 输入维数的选取可依据前文所述的最佳嵌入维数来确定, 因此 Volterra 自适应滤波器的输入维数确定为 4, 为了分析 Volterra 自适应滤波器模型不同阶数对天然气时负荷预测精

度的影响, 分别选取 2 阶, 3 阶, 4 阶, 5 阶 Volterra 自适应滤波器预测模型进行预测。定量评价不同阶数 Volterra 自适应滤波器预测模型的预测精度, 采用常用的 MAPE 值指标进行评价, 如公式(24)所示。

利用 Matlab7.7 分别编制 2 阶, 3 阶, 4 阶, 5 阶 Volterra 自适应滤波器预测模型、ANN 预测模型及傅里叶级数预测模型的程序, 运行环境为 core(TM) 2Quad CPU Q8400, 主频 2.66 GHz, 内存 2 G 的微机。

$$\text{MAPE} = 100 \frac{\sum_{i=1}^n |(Y(i) - \hat{Y}(i)) / Y(i)|}{n}, n = 24, \quad (24)$$

式中, $Y(i)$ 和 $\hat{Y}(i)$ 分别为天然气时负荷的实际值和预测值。

不同阶数 Volterra 自适应滤波器预测模型的预测结果如图 8 所示。由图 8 可知: Volterra 自适应滤波器预测模型能够很好的预测天然气时负荷时间序列的变化规律, 但高阶 Volterra 自适应滤波器预测模型预测结果波动较大, 出现了很多恶劣解, 因此高阶

Volterra 自适应滤波器预测模型不适合天然气时负荷预测。

不同阶数 Volterra 自适应滤波器预测模型的预测误差如表 3 所示。MAPE 值表明: 2 阶 Volterra 自适应滤波器预测模型具有较高的预测精度, 能够满足实际预测的需要。

图 9 给出了不同预测模型的预测结果。为了比较本文预测模型的预测精度, 将本文预测模型与目前天然气时负荷预测领域常用的 ANN 及傅里叶级数预测模型进行对比分析。ANN 预测模型的预测过程见参考文献[5]。文献[1]指出天然气时负荷时间序列可用傅里叶级数数学模型来拟合, 具体做法为采用逐步回归提取法确定可能隐含的周期, 在给定可能隐含周期的基础上, 利用最小二乘法确定傅里叶级数中的系数 A 和 B , 这样就可以确定傅里叶级数的具体形式, 进而根据前 29 d 的时负荷数据里预测第 30 d 的各时负荷数据, 预测过程详见参考文献[1]。不同预测模型与实际值的定性对比见图 9。由于天然气时

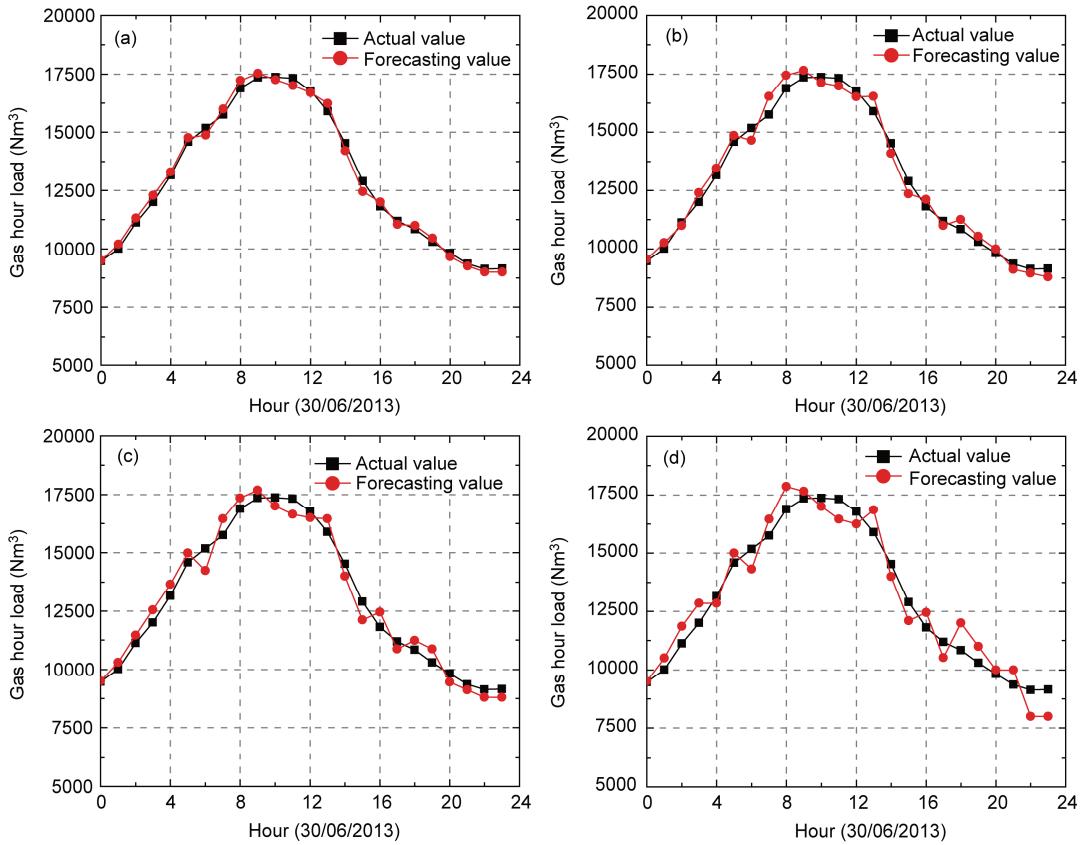
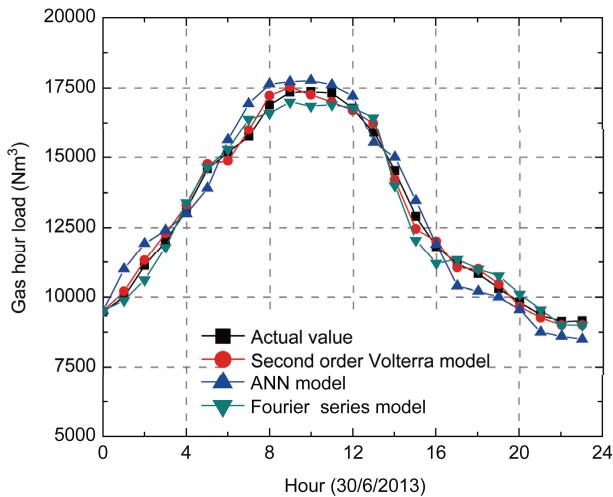


图 8 (网络版彩图)不同阶数 Volterra 自适应滤波器预测对比图

(a) 2 阶 Volterra 预测模型; (b) 3 阶 Volterra 预测模型; (c) 4 阶 Volterra 预测模型; (d) 5 阶 Volterra 预测模型

表 3 不同阶数 Volterra 自适应滤波器预测误差

指标名称	2 阶	3 阶	4 阶	5 阶
MAPE 值(%)	0.5671	1.5746	1.8163	2.1078

**图 9 (网络版彩图)不同预测模型对比图**

负荷的预测误差大小与某日各时负荷的最大值有直接关系,为了定量的评价各方法的预测结果,以国际上普遍采用的归一化均方根误差、归一化绝对平均误差和归一化均方根误差、最大绝对误差以及反映预测值对实际值平均偏离程度的相对均方误差为依据反映各种模型的预测性能,计算公式可参考文献[15],计算结果见表 4。

1) 相对均方误差 E_{re}

$$E_{re} = \frac{\sum_{i=1}^{24} (Y(i) - \hat{Y}(i))^2}{\sum_{i=1}^{24} Y(i)^2}, \quad (25)$$

2) 归一化均方误差 E_{rmse}

$$E_{rmse} = \frac{1}{G_{peak}} \sqrt{\frac{1}{23} \sum_{i=1}^{24} (Y(i) - \hat{Y}(i))^2}, \quad (26)$$

3) 归一化绝对平均误差 E_{nmae}

$$E_{nmae} = \frac{1}{G_{peak}} \frac{1}{24} \sum_{i=1}^{24} |Y(i) - \hat{Y}(i)|, \quad (27)$$

4) 归一化均方根误差 E_{nrmse}

$$E_{nrmse} = \frac{1}{G_{peak}} \sqrt{\frac{1}{24} \sum_{i=1}^{24} (Y(i) - \hat{Y}(i))^2}, \quad (28)$$

5) 最大绝对误差 E_{mae}

表 4 不同预测模型预测误差对比

指标名称	二阶 Volterra 模型	ANN 模型	傅里叶级数模型
相对均方误差	0.000287	0.001847	0.000827
归一化均方误差	0.002772	0.007027	0.004701
归一化绝对平方误差	0.011699	0.029154	0.018282
归一化均方根误差	0.013013	0.032990	0.022072
最大绝对误差	0.025976	0.067734	0.049015

$$E_{mae} = \frac{1}{G_{peak}} \max(|Y(i) - \hat{Y}(i)|), i = 1, 2, \dots, 24, \quad (29)$$

式中, G_{peak} 为某日各时用气负荷值的最大值。

从图 9 和表 4 中可以看出三种预测模型均能很好地反映星期日 0 时到 24 时的变化规律。反映预测值与实际值偏离程度的相对均方误差,本文建立模型较 ANN 及傅里叶级数预测模型要小,且比 ANN 预测模型小了一个数量级。在高峰用气段 8 点到 12 点, ANN 预测模型和傅里叶级数预测模型与实际值点偏离较多,误差较大,而本文预测模型则具有很好的拟合效果。这主要是因为在用气高峰段,用气日不均匀所致,主观模型显然不具备这种拟合能力。本文建立的预测模型较 ANN 及傅里叶级数预测模型在 4 种定量误差评价指标上都要小。可见,二阶单步 Volterra 预测模型能够很好地反映被预测的变化规律,且具有更高的预测精度。

为了进一步验证二阶 Volterra 自适应滤波器预测模型的有效性,本文选取一周之内不同类型的日子(包括一般日和休息日),即 6 月 24 日(星期一)、6 月 25 日(星期二)、6 月 26 日(星期三)、6 月 27 日(星期四)、6 月 28 日(星期五)和 6 月 29 日(星期六)分别进行预测,预测结果如图 10 和表 5 所示。从图 10 和表 5 可以看出,二阶 Volterra 自适应滤波器预测模型能够很好的预测星期一至星期六天然气时负荷的变化规律,同时星期三到星期六的预测结果令人满意,星期一和星期二的预测误差稍大,星期五个别预测点出现了过预测和欠预测,实际值在 8 点到 12 点出现波动,这主要是因为星期五用气高峰段用气不均匀系数较大所致,而二阶 Volterra 自适应滤波器预测模型仍能很好的反映星期一和星期二负荷的变化规律。

6 总结

针对天然气短期负荷的预测问题,笔者建立了

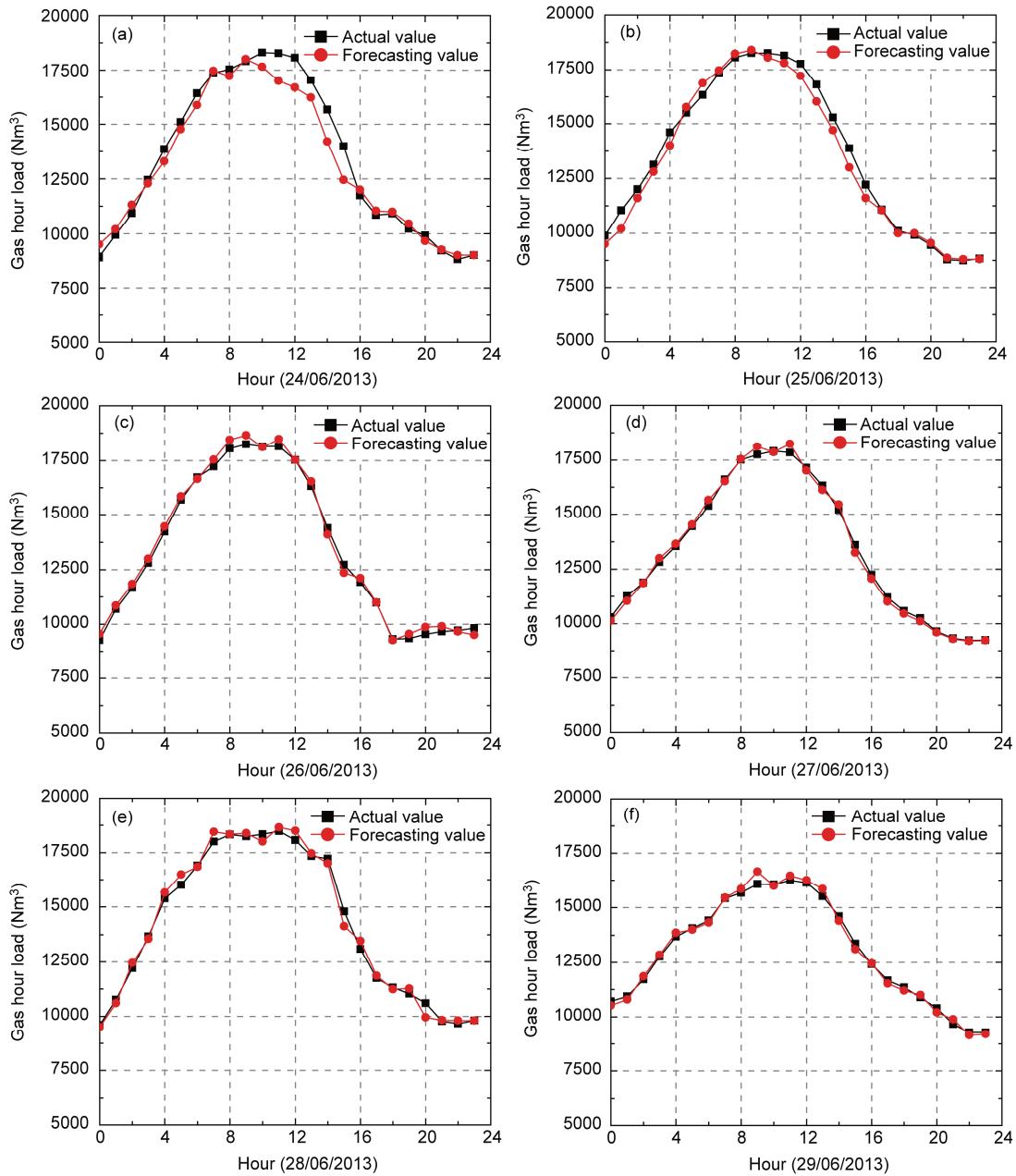


图 10 (网络版彩图)不同日二阶阶数 Volterra 自适应滤波器预测结果

(a) 星期一预测结果; (b) 星期二预测结果; (c) 星期三预测结果; (d) 星期四预测结果; (e) 星期五预测结果; (f) 星期六预测结果

基于混沌理论和 Volterra 自适应滤波器的预测模型。首先, 为揭示天然气短期负荷时间序列内蕴含的动态特性, 利用小数据量法求得了最大 Lyapunov 指数, 验证了天然气短期负荷时间序列具有混沌特性, 为典型的混沌时间序列, 将 Volterra 自适应滤波器预测模型引入到天然气短期负荷预测领域中, 其为天然

气短期负荷预测提供了一种新的思路, 现将本文的工作总结如下。

1) 对天然气时负荷时间序列进行日相关性分析, 分别采用互信息法和伪最近邻域法计算了天然气时负荷时间序列的延迟时间和最佳嵌入维数, 利用小数据量法进行最大 Lyapunov 指数的求解, 结果表明

表 5 不同日二阶阶数 Volterra 自适应滤波器预测误差

指标名称	MAPE 值(%)
星期一(2013 年 6 月 24 日)	0.6872
星期二(2013 年 6 月 25 日)	0.5070
星期三(2013 年 6 月 26 日)	0.3268
星期四(2013 年 6 月 27 日)	0.3276
星期五(2013 年 6 月 28 日)	0.3280
星期六(2013 年 6 月 29 日)	0.3296

该指数大于零, 验证了天然气时负荷时间序列具有明显的混沌特性.

2) 基于天然气时负荷时间序列具有混沌特性这一特点, 将 Volterra 自适应滤波器预测模型引入到了天然气短期负荷预测领域中, 分析了不同 Volterra 自适应滤波器预测模型的阶数对预测精度的影响及对比了二阶 Volterra 自适应滤波器预测模型与 ANN 及

傅里叶级数预测模型的预测精度, 结果表明二阶 Volterra 自适应滤波器预测模型较 ANN 及傅里叶级数预测模型具有更高的预测精度.

3) 为了进一步验证二阶 Volterra 自适应滤波器预测模型的有效性, 本文选取星期一、星期二、星期三、星期四、星期五及星期六分别进行预测, 结果表明除星期一和星期二的预测误差稍大外, 其余日的预测误差均很小, 预测结果能够充分反映负荷变化的规律, 且均能够满足实际预测需要, 模型有效.

此外, 将混沌理论及混沌预测方法引入到天然气短期负荷预测中, 可不必考虑天气、日期类型等外界因素的影响, 仅利用一维负荷数据样本就可以得到较好的预测效果, 简单有效, 可为以后天然气短期负荷预测提供一种新的思路和方法.

参考文献

- 焦文玲, 廉乐明, 崔建华, 等. 城市燃气短期周期负荷预测的时序模型. 天然气工业, 2002, 22: 92–94
- 何春蕾, 段言志, 邬宗婧, 等. 基于气温的城市燃气短期日负荷预测模型—以四川省成都地区为例. 天然气工业, 2013, 4: 131–134
- 李持佳, 焦文玲, 赵林波. 燃气短期负荷预测的小波分析综合模型. 天然气工业, 2007, 8: 103–105
- 焦文玲, 展长虹, 廉乐明, 等. 城市燃气短期负荷预测研究. 煤气与热力, 2001, 6: 483–486
- 杨昭, 刘燕, 苗志彬, 刘振能. 人工神经网络在天然气负荷预测中的应用. 煤气与热力, 2003, 6: 331–333
- 焦文玲, 赵林波, 秦裕琨. 城市燃气负荷非平稳时序预测模型的研究. 煤气与热力, 2003, 8: 451–454
- 刘洋. 基于混沌理论的城市天然气日负荷预测研究. 硕士学位论文. 成都: 西南石油大学, 2011
- 韩敏. 混沌时间序列预测理论与方法. 北京: 中国水力水电出版社, 2007. 167–169
- 张勇, 关伟. 基于最大 Lyapunov 指数的多变量混沌时间序列预测. 物理学报, 2009, 58: 756–763
- 孟庆芳, 张强, 龙文英. 混沌时间序列多步自适应预测方法. 物理学报, 2006, 55: 1666–1671
- 吕涛, 唐巍, 所丽. 基于混沌相空间重构理论的风电场短期风速预测. 电力系统保护与控制, 2010, 38: 113–117
- 罗海洋, 刘天琪, 李兴源. 风电场短期风速的混沌预测方法. 电网技术, 2009, 33: 67–71
- 李恒超, 张家树. 混沌时间序列局域零阶预测法性能比较. 西南交通大学学报, 2004, 39: 328–331
- 张家树, 肖先赐. 混沌时间序列的 Volterra 自适应预测. 物理学报, 2000, 3: 403–408
- 张学清, 梁军. 风电功率时间序列混沌特性分析及预测模型研究. 物理学报, 2012, 19: 190507
- 师洪涛, 杨静玲, 丁茂生, 等. 基于小波-BP 神经网络的短期风电功率预测方法. 电力系统自动化, 2011, 35: 44–48
- 张华强, 张晓燕. 基于混沌理论和 LSSVM 的蒸汽负荷预测. 系统工程理论与实践, 2013, 33: 1058–1066
- Henry D, Abarbanel N M, Rabinovich M I. Distribution of mutual information. Phys Lett A, 2001, 281: 368–373
- Kennel M B, Brown R, Abarbanel H D I. Determining embedding dimension for phase space reconstruction using a geometrical reconstruction. Phys Rev A, 1992, 45: 3403–3411
- Wolf A, Swift J B, Swinney H L, et al. Determining Lyapunov exponents from a time series. Physica D, 1985, 16: 285–317
- Rosenstein M T, Collins J J, De Luca C J. A practical method for calculating largest Lyapunov exponents from small data sets. Physica D, 1993, 65: 117
- 张玉梅, 吴晓军, 白树林. 交通流量序列混沌特性分析及 DFPSOVF. 物理学报, 2013, 62: 190509
- 张永明, 齐维贵. 供热负荷时间序列混沌特性分析及预报模型研究. 物理学报, 2011, 60: 100508

Gas load forecasting based on chaotic theory and Volterra adaptive filter

QIAO WeiBiao^{1,2,3}, CHEN BaoDong² & LU HongFang⁴

¹ College of Pipeline and Civil Engineering, China Petroleum University, Qingdao 266555, China;

² College of Oil and Gas Engineering, Liaoning Shihua University, Fushun 113001, China;

³ China Resource (Nanjing) Municipal Design Co Ltd, Nanjing 210000, China;

⁴ School of Petroleum Engineering, Southwest Petroleum University, Chengdu 610500, China

Gas is an important source of clean energy, so in the future, gas load forecasting will have a very important meaning for the gas company to develop an appropriate business plan between ‘upstream’ gas supplier and ‘downstream’ users. Aiming at the gas short-term load forecasting problem, a prediction model has been put forward based on chaotic theory and Volterra adaptive filter. Firstly, conduct day-to-day correlation analysis for collected gas hour load data, and then calculate delay times and find out the optimal embedding dimension by the approach of mutual information and pseudo-nearest-neighbor. Secondly, on the basis of phase-space reconstruction, carry out chaotic characteristic analysis for collected gas hour load data. Once again, aiming at the current situation of most existing prediction models being subjective, reduce the subjectivity in the process of gas load forecasting by introduction of Volterra adaptive filter prediction model. Finally, the influence of the different order of Volterra adaptive filter on prediction results, and the comparison of accuracy among the Volterra adaptive filter prediction model, ANN (artificial neural network, ANN) prediction model and Fourier series prediction model are discussed through gas load forecasting example. Additionally, the forecasting results of the second order Volterra adaptive filter prediction model showed that: compared with ANN prediction model and fourier series prediction model, the second order Volterra adaptive filter prediction model has higher accuracy, and may provide a useful reference for practical engineering applications of short-term gas load forecasting.

gas load, chaos, Volterra adaptive filter, ANN, fourier series, forecast

doi: 10.1360/N092014-00202