

输油管道压力时间序列混沌特性研究^{*}

刘金海^{1)†} 张化光¹⁾ 冯 健^{1)‡}

1) (东北大学信息科学与工程学院, 沈阳 110004)

2) (东北大学教育部暨辽宁省流程工业综合自动化重点实验室, 沈阳 110004)

(2008 年 3 月 21 日收到, 2008 年 5 月 12 日收到修改稿)

为了找出输油管道压力波动的内在动态特征, 采用非线性的分析方法探讨了实测输油管道压力波动中存在混沌动态特性的可能性. 以 6 个典型的输油管道压力实测数据集为研究对象, 重构相空间、求得了分形维数和 Lyapunov 指数谱、验证了数据的平稳性和非线性. 通过对所得结果的分析, 从理论上证明了输油管道压力信号具有严格混沌动态特性. 为基于输油管道压力时间序列的研究提供了混沌理论基础.

关键词: 压力时间序列, Lyapunov 指数, 分形维数, 混沌

PACC: 0545, 0555, 4752

1. 引 言

随着世界范围内石油资源的紧缺, 石油运输过程中的安全问题日益引起人们的重视. 由于管道老化、泥石流、打孔盗油等原因, 输油管道经常发生泄漏. 为了及时发现泄漏, 基于负压波、声波、光纤等原理^[1]的实时泄漏检测系统被设计和应用. 其中以负压力波为主要信号的输油管道泄漏故障诊断方法是目前我国输油管道泄漏检测所用的主要方法. 截至 2007 年底, 仅对中石化和中石油的初步统计就有超过一万公里的输油管道上安装了基于负压波的泄漏故障诊断系统. 该方法主要利用管段两端的瞬时压力数据, 通过压力的变化判断管道是否泄漏. 这种基于负压力波信号的分析方法对于泄漏量大于瞬时流量 3% 的泄漏可以实现较为准确的报警和定位, 但是很难检测较小的泄漏, 导致在检测小泄漏时经常漏报或者误报. 基于上述原因, 常用的相关函数法、小波方法、人工智能方法等^[2-4]对于大的信号波动有效, 但是在处理小泄漏的问题上效果不理想, 其根本原因是微弱的压力波动信号被噪声所掩盖而无法检测.

在实践中我们发现管道中的流体呈现类似湍流的运动, 湍流是一种典型的混沌现象. 由于混沌系统具有检测微弱信号能力^[5, 6], 如果能够在输油管道中发现混沌特征, 那么可以从混沌理论出发研究如何检测管道压力中的微弱信号, 为利用混沌相关理论^[7, 8]处理管道泄漏特别是小泄漏问题开辟一个新的研究途径. 所以, 本文从压力时间序列的内在特性着手, 研究输油管道压力时间序列存在混沌特性的可能性.

判断时间序列是否是混沌时间序列需要满足一系列的条件^[9, 10]. 这些条件涉及到时间序列的内在动态特性和吸引子的熵和几何特性. 根据输油管道压力时间序列的特点, 这些条件可以归纳为四个主要的条件:

- 1) 时间序列要足够长; 时间序列的状态应该是稳定的; 由于实际采集的时间序列一般都含有噪声, 在研究混沌特性之前需要对实际数据滤波.
- 2) 吸引子具有分形维数, 主要表现是关联维是分数值, 随着嵌入维的增长关联维会趋于一个常量.
- 3) 计算出的 Lyapunov 指数谱要满足: 至少一个 Lyapunov 指数是正值; 所有指数之和是负值; Lyapunov 维数和关联维数接近.

^{*} 国家自然科学基金(批准号: 60534010, 60572070, 60521003, 60774048, 60774093, 60728307), 长江学者和创新团队发展计划, 高校博士点基金(批准号: 20070145015), 辽宁省自然科学基金(批准号: 20062018), 中国博士后科学基金(批准号: 20060400962), 国家高技术研究发展专项(批准号: 2006AA04Z183), 教育部流程工业综合自动化重点实验室开放课题基金(批准号: PAL200503)资助的课题.

[†] E-mail: jh_lau@126.com

4) 时间序列应该非线性的, 这意味着理论上(实际当中一般很难找到)通过微分方程才可以描述它的时域行为。

本文首先对待测数据集进行说明, 并且根据数据集的特点对数据进行预处理, 然后, 对预处理后的数据集重构相空间; 第四部分求解和分析了数据集的分形维数; 第五部对目标数据集的 Lyapunov 指数进行了估计和分析; 第六部分评估了数据集的平稳性和非线性, 最后, 给出了本文的研究结论。

2. 数据采集与预处理

2.1. 数据集说明

本文中所用到的 6 个压力时间序列取自管道泄漏检测与定位系统实际采集到的平稳压力条件下的数据。该系统基于负压波原理研制, 由于对管道内在动态未知, 所以主要基于对压力信号的统计分析, 基本结构如图 1。信号采集器具有 12 位的采样精

度, 采样频率 1—50 Hz, 本文中数据集的采样频率为 1 Hz。由于采集到的数据可能受到温度、压变精度、管道长度等多因素影响, 为了所测得的数据具有普遍性, 所取的数据集在时间、地点、温度等参数都不尽相同, 具体参数见表 1。

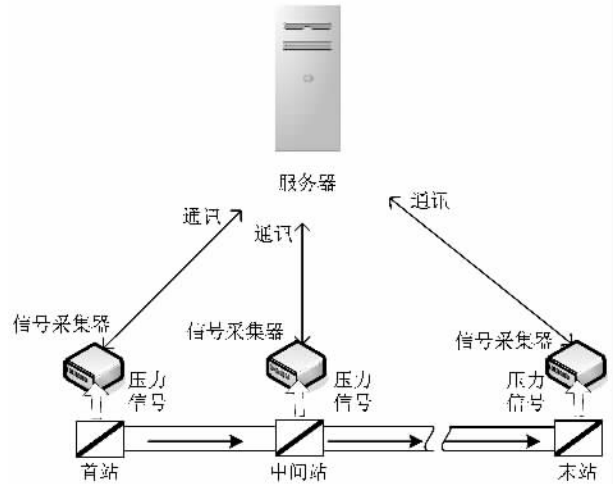


图 1 泄漏检测与定位系统结构

表 1 数据集参数表

数据集	A	B	C	D	E	F
采集地点	江西九江	湖南洪湖	安徽怀宁	江苏徐州	湖北襄樊	山东枣庄
时间	2007.10.25	2007.9.25	2007.9.12	2007.12.20	2007.5.4	2007.7.20
	4:30	1:30	14:00	22:00	13:00	8:30
管道温度/°C	22	21	28	38	41	26
压变精度/%	0.075	0.075	0.075	0.025	0.25	0.075
流量(m ³ /h)	330	480	2200	440	133	390
是否平稳	是	是	是	是	是	是
管段长/km	46.2	37.2	120.8	25.6	19.6	115
输送介质	原油	原油	原油	成品油	原油	成品油
数据个数	10021	8095	10800	9000	9301	9601

由于这些数据的取得条件(温度、流量、采集精度等)具有很大差异, 因此具有了广泛的代表性, 如果能够从这些数据集中发现普遍的现象, 则可以对输油管道压力时间序列的动态给出结论。

2.2. 数据滤波

由于实际采集的数据一般都含有噪声, 所以把经典理论应用于实际的数据具有较大的风险性。因为噪声的干扰有可能掩盖甚至改变原时间序列的特性, 所以如果对噪声不加以考虑和处理就很可能导致结论的误差甚至错误的结论。所以在对实测数据

进行系统分析前, 要考虑其噪声的大小, 并且在尽可能的情况下消除噪声对数据分析的影响。在不知道存在的噪声到底属于哪种噪声的情况下, 一般可以认为这种噪声为白噪声, 且其均值为 0。

本文研究的时间序列是压力时间序列, 属于低频范围, 所以可以对数据进行低通滤波。图 1 中所示的信号采集器设计了低通滤波电路, 通带范围是 1—100 Hz, 这就保证了高频信号在硬件部分被过滤掉。然后, 为了消除尖峰干扰, 采用滑动均值滤波方法对数据再进行滤波。通过均值滤波, 瞬间突变将被平均掉, 而慢变化则被保留。其滤波器的形式为

$$v_n = \sum_{i=1}^5 a_i \hat{v}_{n+i-4}, \quad (1)$$

式中取 $a_i = \frac{1}{5}$.

值得说明的是,在这里无论是硬件滤波还是滤波器滤波都不能消除噪声干扰,只是对特定的高频和尖峰噪声进行了处理,减小了这些噪声对数据的污染程度.

3. 重构相空间

Takens 定理指出一个时间序列包含系统的全部动态信息,并且可以通过重构还原系统状态,重构状态空间和原时间序列具有相同的动态行为. 根据 Takens 定理,我们可以利用时间序列 V 重新构造一个与之等价的动态系统,表示为 $d_i^T = (v_t, v_{t+\tau}, v_{t+2\tau}, \dots, v_{t+(d-1)\tau})$, $t = 1, 2, \dots, M_d$, 其中 $M_d = N - (d-1)\tau$, N 是数据的个数, τ 是时间延迟, d 为嵌入维. 在重构时, τ 和 d 选择直接影响重构的质量.

3.1. 嵌入延迟的确定

实测时间序列,因为存在噪声干扰和估计误差,如果 τ 太小,相空间轨迹会向同一位置挤压,新信息不容易显露;如果 τ 太大,会导致向量的分量变得相互独立,使得其后的计算无效. 所以有必要选择合适的 τ 以使前后两个时刻的值能较好地反应系统的特性. 计算 τ 的方法有相空间扩展法、自相关法、交互信息法^[11]等. 由于交互信息方法计算嵌入时间延迟效果较好,所以本文采用交互信息法计算时间延迟.

交互信息是从熵的角度推出的一种信息论性质. 它决定了被测信号 v_i 相对其他某个时刻的测量值 $v_{i+\tau}$ 具有多少信息. v_i 和 $v_{i+\tau}$ 间的交互信息定义为

$$K(\tau) = \sum_{v_i, v_{i+\tau}} P(v_i, v_{i+\tau}), \quad (2)$$

式中 $P(v_i)$ 和 $P(v_{i+\tau})$ 为 v_i 和 $v_{i+\tau}$ 观测值的归一化分布; $P(v_i, v_{i+\tau})$ 为其联合分布, $K(\tau)$ 第一次达到最小值时对应的 τ 作为嵌入延迟. 对本文 6 个数据集分别作图如图 2 所示,从图中可以看出数据集中 B 和 F 较为光滑,容易看出在 4 处达到第一个最小值;而其他数据集由于受到外界较大干扰,在下降的过程中有波动,这为寻找第一个达到最小值带来了一定困难. 从图中看到受噪声影响的曲线虽然第一次

达到最小值略有不同,但是可以看到这几条曲线都在 4 处达到了局部最小值. 这表明,这些被测的数据集可能产生于相同的动态系统,从而具有相同的动态特性.

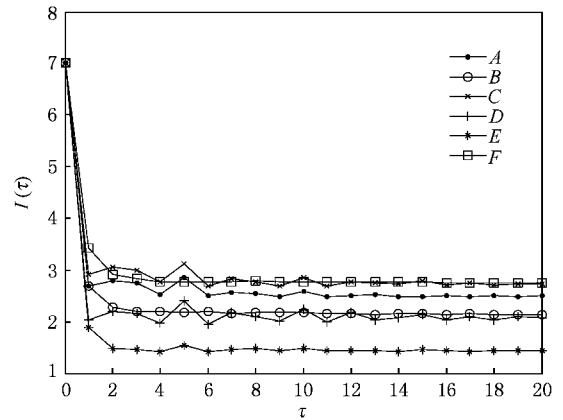


图 2 交互信息法计算嵌入时间延迟

3.2. 嵌入维的确定

作为重构相空间的另外一个重要参数,理论上嵌入维应该选取足够大,因为过小的嵌入维将使吸引子不能完全展开,背离了相空间重构的本来意义,但是大的嵌入维可能导致数据计算量的指数增加,且在含有噪声的系统中额外的嵌入维将引起很大的误差. 所以应该选择一种合适的算法去求取最佳嵌入维 d ,使之既能很好地展现吸引子的内部结构又能尽量地减小嵌入维数. 寻找 d 的有效方法有奇异系统分析法(singular system analysis, SSA)^[12]、伪最近邻方法(false nearest neighbors, FNN)^[13]. 由于 FNN 在计算最佳嵌入维 d 效果更好一些,所以本文采用 FNN 法计算最佳嵌入维. 令相点 $a_i^T = (v_t, v_{t+\tau}, v_{t+2\tau}, \dots, v_{t+(d-1)\tau})$ 最近距离内的相点为 $(a_i^{NN})^T = (v_t^{NN}, v_{t+\tau}^{NN}, v_{t+2\tau}^{NN}, \dots, v_{t+(d-1)\tau}^{NN})$, 其距离为

$$R_d = \| a_{t+d\tau}^T - (a_{t+d\tau}^{NN})^T \|, \quad (3)$$

当相空间的维数由 d 增加到 $d+1$ 时,这两个相点距离变为

$$R_{d+1} = \| a_{t+d\tau}^T - (a_{t+d\tau}^{NN})^T \|, \quad (4)$$

若 R_{d+1} 比 R_d 有较大变化,则可以认为 d 维中两个相邻的点是伪最近邻点. 即

$$\frac{R_{d+1}}{R_d} > R_{tol}, \quad (5)$$

则 $(a_i^{NN})^T$ 是 a_i^T 的伪最近邻点, 阈值 R_{tol} 一般在 10, 50 之间取值.

为了上述准则能够获得较好的结果,对于含有噪声实测数据,除了上面的第一条准则外,还需要增加第二条准则

$$\frac{R_{d+1}}{\sqrt{\frac{1}{N} \sum_{i=1}^N (v_i - \bar{v})^2}} > R_{10L2}, \quad (6)$$

其中 $\bar{v} = \frac{1}{N} \sum_{i=1}^N v_i$, R_{10L2} 一般取 2.

只要上面两个准则有一个成立,那么最近邻点就是伪的.对于实测数据, d 可以从 2 开始计算伪最近邻点的比例,然后增加 d 到伪最近邻点的比例小于 5% 或者伪最近邻点不再随 d 的增加而减少时,可以认为吸引子完全展开,此时的 d 即为最佳嵌入维.

对实测数据集 A—F 计算最佳嵌入维如图 3 所示,可以看出由于噪声的干扰,嵌入维从 4 到 5 有微小变化,但是以 5 为主,具体值见表 2.

这里需要说明,Takens 曾指出嵌入维应该满足 $d \geq 2D$ (D 为关联维),这是一个充分条件,Ding 等^[14]又证明只需要 $d \geq D$ 就是重构相空间的充分条件.关于关联维的计算将在下一部分说明.

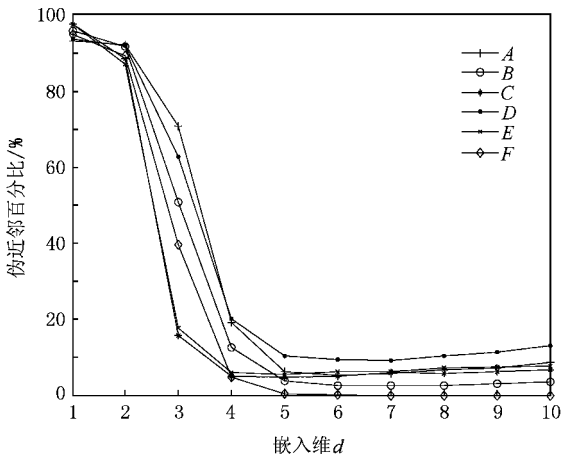


图 3 伪近邻法求最佳嵌入维

4. 分形维数

维数是空间和集合的重要集合特征量.分形维数是维数的推广.分形的维数有多种定义,如拓扑维、Hausdorff 维数、容量维数、相似维数和关联维数等.其中关联维 D 是被研究得最认真、最彻底的一个用于刻画动态系统的不变量.动态系统的这一维数直接反应了吸引子上的点在相空间或嵌入空间中

分布的方式.本文采用的关联维算法是 Grassberger 和 Procaccia 提出的称之为 G-P 算法.考虑关联积分

$$\alpha(r) = \frac{1}{M_d^2} \sum_{i=1}^{M_d} \sum_{j=1}^{M_d} \mathcal{A}[r - \|\beta_i - \beta_j\|], \quad (7)$$

其中, $\|\beta_i - \beta_j\|$ 为相空间中两点的距离, M_d 为相点个数, $\mathcal{A}(\cdot)$ 为 Heaviside 函数,即

$$\mathcal{A}(v) = \begin{cases} 0, & v < 0, \\ 1, & v \geq 0, \end{cases} \quad (8)$$

当 $r \rightarrow 0$ 时关联积分 $\alpha(r)$ 与 r 存在以下关系:

$$\lim_{r \rightarrow 0} \alpha(r) \propto r^D, \quad (9)$$

其中, D 是关联维,恰当地选取 r ,使得 D 能够描述奇异吸引子的自相似结构,由(9)式可得

$$D = \log \alpha(r) / \log r, \quad (10)$$

然后通过使嵌入维数从小到大的增加,对每个嵌入维取双对数关系 $\log \alpha(r) / \log r$ 中的直线段,用最小二乘拟合得出斜率叫做关联指数的直线,该斜率会随嵌入维的增大而增大,最后达到一个饱和值,这个饱和值就是关联维数 D .

这里有两点需要注意,首先即使在随机过程中,利用 G-P 算法所得到的关联维也可能为分数形式.再者,在实测数据中,由于存在模数转换噪声、原始噪声等原因,G-P 算法的精度也是一个问题.另外时间序列的长度也会影响关联维的估计.因此,需要求得最大关联维 D_{max} 对给定的时间序列关联维进行验证. D 不应该超过 D_{max} ,否则 D 的估计可能存在误差.

$$D_{max} = \frac{2 \log N}{\log(1/\rho)}, \quad (11)$$

其中, $\rho = r/\mu$,这里 μ 是重构吸引子的直径.

本文对所有数据集取 $r = 0.007$,计算结果如表 2 所示.结合表 2 和图 3,可以得到这样两个结果,首先从表中可以看出所有数据集的关联维具有分数的形式,且在一个范围内变化.再者,理论上关联维应该小于嵌入维,同时也应该小于最大嵌入维 D_{max} ,但是由于噪声和算法参数选择的影响,难以准确比较关联维和嵌入维,且在数据集 E 中 $D_{max} < D$,但是

表 2 嵌入延迟、嵌入维和关联维

数据集	A	B	C	D	E	F
嵌入维 d	5	5	4	5	4	5
延迟 τ	4	4	4	4	4	4
关联维 D	4.98	4.94	3.56	4.69	3.93	4.47
最大关联维 D_{max}	4.98	4.95	3.81	4.99	3.92	4.65

可以看出关联维与嵌入维基本在同一个范围内,总体上关联维 D 小于或者接近于 D_{\max} . 表 2 关联维 D 取两位有效值. 从上述两个结果可以得出产生时间序列的系统具有分形结构.

5. Lyapunov 指数谱

5.1. Lyapunov 指数

Lyapunov 指数^[15-17]作为混沌系统的一个重要特征量,是判断时间序列是否混沌的一个主要条件. 由于本文所研究的数据集含有噪声,所以以 Darbyshire 和 Broomhead^[18,19]所提出的实测数据 Lyapunov 指数谱理论为依据,计算数据集的 Lyapunov 指数谱.

考虑动态系统 (ϕ, M) , 流形 M 是系统的状态空间, ϕ 为 $M \rightarrow M$ 微分同构. Jacobian 矩阵 $T_x^{(N)}$ 为 M 的 tangent 空间中向量 x 到 $\phi^N x$ 的 tangent 映射. 根据系统的轨迹可以由一步 tangent 映射推导出 N 步 tangent 映射

$$T_{x_0}^{(N)} = T_{x_{n-1}}^{(1)} T_{x_{n-2}}^{(1)} \cdots T_{x_0}^{(1)}, \quad (12)$$

式中 $x_n = \phi^N x_0$. 则 Lyapunov 指数谱可由矩阵 $\Lambda_{x_0} = \lim_{N \rightarrow \infty} ((T_{x_0}^{(N)})^T T_{x_0}^{(N)})^{1/2N}$ 的特征值的对数集给出.

5.2. Lyapunov 指数谱与 Lyapunov 维数

我们利用最佳嵌入维 d 和 $N \times m$ 轨迹矩阵 A 计算维数为 $N \times d$ 子空间矩阵 X ,

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} a_1^T c_1 & a_1^T c_2 & \cdots & a_1^T c_d \\ a_2^T c_1 & a_2^T c_2 & \cdots & a_2^T c_d \\ \vdots & \vdots & \ddots & \vdots \\ a_N^T c_1 & a_N^T c_2 & \cdots & a_N^T c_d \end{bmatrix}. \quad (13)$$

根据(12)式, $T_x^{(n)}$ 可以通过计算 $T_{x_i}^{(1)}$ 求得. $T_x^{(1)}$ 可以通过为嵌入动态系统设计一个预估器来计算,即

$$\delta x_{i+1} \approx T_{x_i}^{(1)} \delta x_i, \quad (14)$$

式中 $\delta x_i = x_j - x_i$, x_j 是 x_i 临近的一个向量,且 $\{x_j \in R^d \mid j \in 1, \dots, N\}$, 同理有 $\delta x_{i+1} = x_{j+1} - x_{i+1}$. 计算出 δx_{i+1} 和 δx_i , 就可以通过(14)式计算出 $T_{x_i}^{(1)}$.

通过 $\|x_j - x_i\| \leq \epsilon$ (ϵ 是足够小的正数)确定 x_j 共有 N_b 个, 计算 $\{\delta x_i(j_k) = x_{j_k} - x_i \mid k \in 1, \dots, N_b\}$ 得到 N_b 个 d 维的向量, 利用这些向量组成矩阵 $B_{x_i} = (\delta x_i(j_1), \delta x_i(j_2), \dots, \delta x_i(j_{N_b}))^T$, 我们称之为

邻域矩阵. 再计算 $\{\delta x_{i+1}(j_k) = x_{j_{k+1}} - x_{i+1} \mid k \in 1, \dots, N_b\}$ 得到下一个邻域矩阵 $B_{x_{i+1}}^{(1)} = (\delta x_{i+1}(j_1), \delta x_{i+1}(j_2), \dots, \delta x_{i+1}(j_{N_b}))^T$. 当 ϵ 足够小时, 有下面的关系成立:

$$B_{x_{i+1}}^{(1)} = B_{x_i} (T_{x_i}^{(1)})^T. \quad (15)$$

设 $B_{x_i}^+$ 为 B_{x_i} 的逆矩阵, 因为 B_{x_i} 不一定存在逆矩阵, 所以称 $B_{x_i}^+$ 为 B_{x_i} 广义逆矩阵, 则由(15)式可以得到 $(T_{x_i}^{(1)})^T = B_{x_i}^+ B_{x_{i+1}}^{(1)}$, 又根据 $B_{x_i}^+ = (B_{x_i}^T B_{x_i})^{-1} B_{x_i}^T$, 有

$$T_{x_i}^{(1)} = ((B_{x_i}^T B_{x_i})^{-1} B_{x_i}^T B_{x_{i+1}}^{(1)})^T. \quad (16)$$

当求得了 $T_{x_i}^{(1)}$ 后, 通过 QR 分解来求 Lyapunov 指数谱. 对任意一个正交矩阵, 取其列作为 x_0 的正切空间的基, 在 $T_{x_0}^{(1)}$ 的作用下, 这个基通过 $T_{x_0}^{(1)} Q_0$ 的列映射到 x_1 点. 这些向量由于 $T_{x_0}^{(1)}$ 的影响不一定正交, 为了求 Lyapunov 指数谱, 要把 $T_{x_0}^{(1)} Q_0$ 的列正交化:

$$Q_1 R_0 = T_{x_0}^{(1)} Q_0, \quad (17)$$

Q_1 的列就是 x_1 的正交基; R_0 是有正对角元素的上三角矩阵. 这个过程一直循环下去, 即

$$Q_2 R_1 = T_{x_1}^{(1)} Q_1,$$

$$Q_N R_{N-1} = T_{x_{N-1}}^{(1)} Q_{N-1}, \quad (18)$$

则有

$$T_{x_0}^{(N)} Q_0 = Q_N R^{(N)} = Q_N R_{N-1} R_{N-2} \cdots R_1 R_0. \quad (19)$$

对于所有的 Q_0 有 $\lambda_i = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{j=0}^{N-1} \log (R_j)_{ii}$, 其中 $(R_j)_{ii}$ 是 R_j 的第 i 个对角元素. λ_i 就为时间序列的 Lyapunov 指数.

当计算出 Lyapunov 指数谱后, 就可以根据这些指数计算 Lyapunov 维数(也叫 Kaplan-Yorke 维数) D_{KY} :

$$D_{KY} = K + \frac{1}{|\lambda_{K+1}|} \sum_{a=1}^K \lambda_a, \quad (20)$$

其中 λ_a 是 Lyapunov 指数 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_L \geq 0$, 且

$$\sum_{a=1}^K \lambda_a \geq 0, \sum_{a=1}^{K+1} \lambda_a < 0.$$

计算数据集所得的 Lyapunov 指数如图 4 所示, 从图中可以看出根据数据量的不同得到的 Lyapunov 指数也不相同, 而数据量在 2000 个以上计算所得的 Lyapunov 指数基本平稳. 其中, 数据集 C 由于管道距离长 (120.8 km, 见表 1) 泵扬程大, 所以受到噪声

干扰大,导致计算出的 Lyapunov 指数有较大波动. 大于 500 个数据的所有数据集的 Lyapunov 指数谱的平均值、Lyapunov 指数和、Lyapunov 维数见表 3,从表中可以看出每个数据集的都有正的 Lyapunov 指数,且每个数据集的所有指数之和都是负数,这说明泄漏检测系统是个能量耗散系统. 理论上 Lyapunov 维数与关联维应该相等,但从表 3 中可以发现 Lyapunov 维与表 2 中关联维数接近但不相等,都在同一个范围内. 这是由于噪声干扰和算法参数选取的原因. 重要的是,关联维和 Lyapunov 维始终在同一个范围内.

表 3 Lyapunov 指数与 Lyapunov 维数表

数据集	A	B	C	D	E	F
指数 1	0.6461	0.6315	0.6048	0.7077	0.5675	0.5251
指数 2	0.3504	0.3436	0.2544	0.3505	0.2048	0.2608
指数 3	0.0914	0.0770	-0.1244	0.0614	-0.1574	0.0341
指数 4	-0.2530	-0.2563	-0.8541	-0.2842	-0.8219	-0.2964
指数 5	-0.9638	-0.9617		-1.0289		-0.9889
Lyapunov 指数和	-0.1289	-0.1659	-0.1193	-0.1935	-0.2069	-0.4652
Lyapunov 维数 D_{KY}	4.87	4.83	3.86	4.81	3.75	4.53

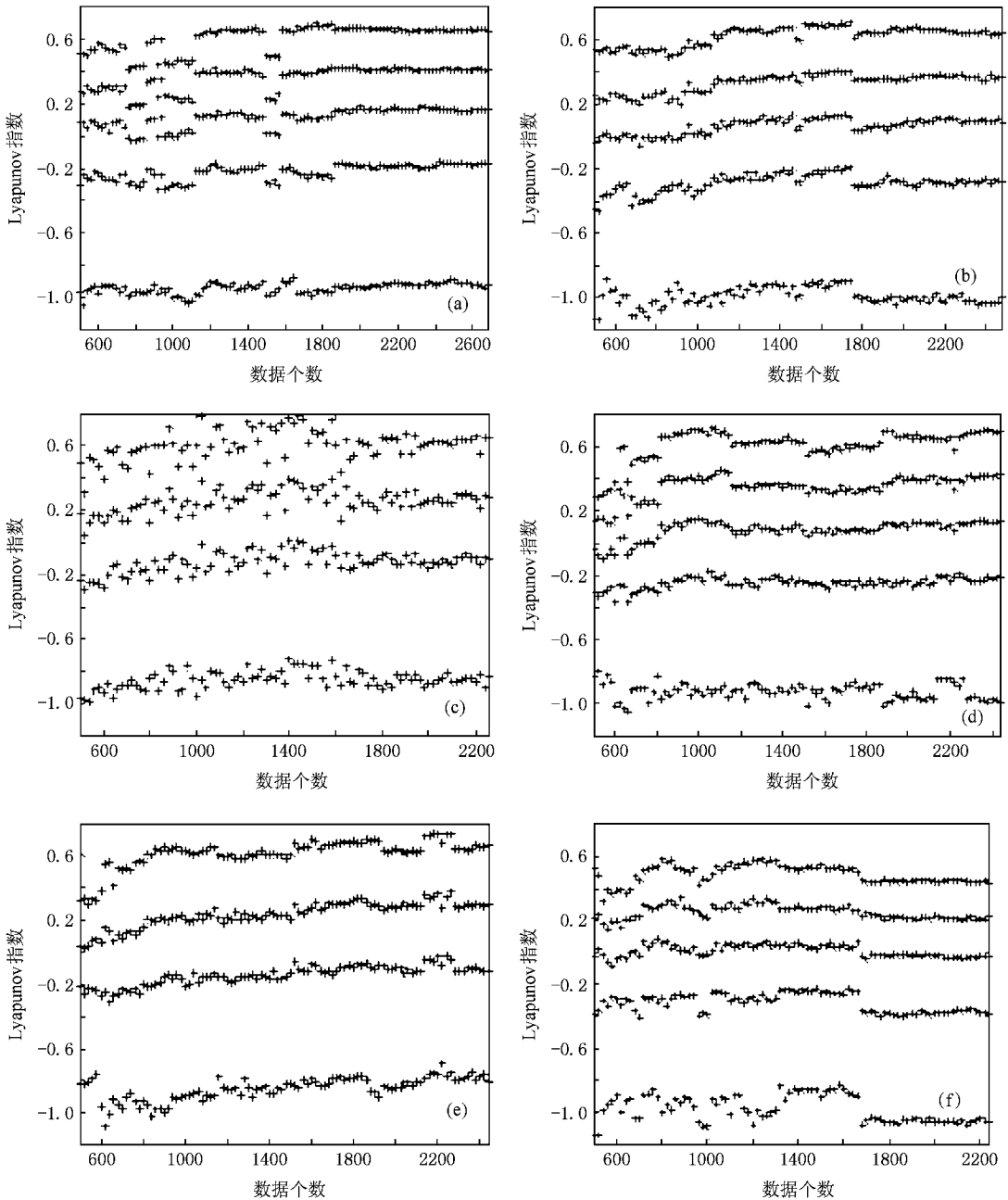


图 4 输油管道压力时间序列 Lyapunov 指数图 (a)~(f)分别是数据集 A~F 的 Lyapunov 指数

6. 系统的非线性证明

6.1. 平稳性分析

理论研究中假设被研究的时间序列是平稳的,但是从实际采集来的数据是否平稳是未知的,需要通过特定的方法来进行验证.这里采用递归图作为分析工具来验证数据是否平稳. a_i^T 为 d 维空间中描述动态系统轨道上的第 i 个点.只要轨道中另外一个点 a_j^T 与 a_i^T 接近,就在以 i 和 j 组成的坐标系的

(i, j) 处画出一.具体的步骤如下:

- 1) 利用嵌入维 d 和嵌入延迟 τ 构建 d 维轨道 $a_i^T = (v_i, v_{i+\tau}, v_{i+2\tau}, \dots, v_{i+(d-1)\tau})$;
- 2) 以 a_i^T 为中心,选择合适的 r_i 为搜索半径的球包含适当数目的其他点 a_j^T ;
- 3) 在以 i 和 j 组成的坐标系的 (i, j) 处画出一;
- 4) 重复上述过程.

对本文的数据集所做的递归图如图 5 所示,从图中可以看到所有的数据集的递归图分布都较均匀,证明了所研究的数据集的平稳性.

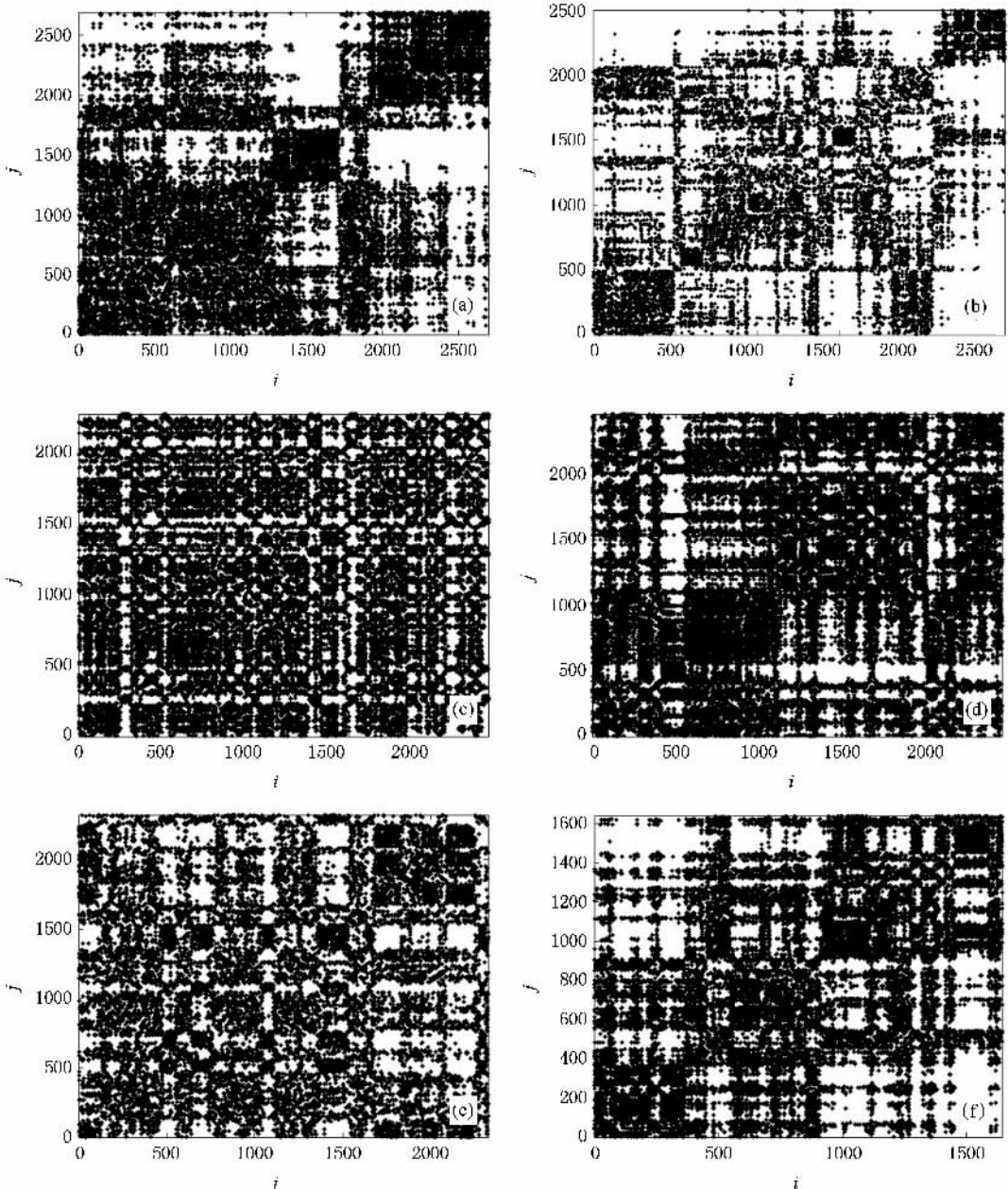


图 5 输油管道压力时间序列递归图 (a)~(f)分别是数据集 A—F 的递归图

6.2. 非线性证明

为了证明时间序列是非线性的,我们必须推翻时间序列是随机过程的假设.这里我们用替代数据分析方法来验证解决^[20-23].其基本思想是提出零假设,即假设原时间序列是线性相关的,再以某种方式生成线性相关的数据集,称为替代数据集(surrogate data set).然后利用统计检验方法检验其统计差异的显著性,并根据显著性水平来判断是否拒绝零假设.如果零假设成立,即原始数据和替代数据服从相同的分布,具有相同的统计特性,两者的特征量相近;反之如果零假设不成立,即原始时间序列为非线性的,那么原始数据要服从确定性的非线性规律,而被均匀随机化后的替代数据则不具有这一性质,二者的统计特性会有很大的差异.所以,要检验原始数据是否含有非线性成分,只需要比较生成的代替数据集和原始数据集的关联维的统计量来判断原始数据是否具有非线性.替代数据集利用相位随机算法产生,该算法保留了原始数据的功率谱幅值,但随机产生替代数据的相位值,过程如下:

- 1) 将原始数据输入数组 $y[t], t=1, 2, \dots, N$;
- 2) 对 $y[t]$ 计算傅里叶变换 $\alpha[t]=\text{DFI}(y[t])$;
- 3) 取 $\alpha[t]$ 的幅度 $|h_2(\omega_i)|$ 和相位 $\phi_i, i=1, 2, \dots, N$;
- 4) 在 $[-\pi, \pi]$ 之间产生一个任意相位数据数组 $\phi_i^*, i=2, 3, \dots, N/2$, 如果 N 是偶数,则令 $\phi_1^* = \phi_{N/2+1}^* = 0, \phi_{N+2-i}^* = -\phi_i^* (i=2, 3, \dots, N/2)$, 如果 N 为奇数,则令 $\phi_1^* = 0, \phi_{N+2-i}^* = -\phi_i^* (i=2, 3, \dots, \text{cei}(N/2))$, 其中 $\text{cei}(N/2)$ 是取 $N/2$ 的上限整数;
- 5) 利用 $|h_2(\omega_i)|$ 和 ϕ_i^* 做傅里叶逆变换得到 $y^*[t]$;

6) 重复上述的步骤得到不同的替代数据集.

确定替代数据集后,需要找到一个统计量来判断原始数据和替代数据的特征量的异同.这个统计量可以利用均方差、关联维、Liyapunov 指数、熵等.本文中采用关联维^[21]作为判定统计量.

$$S = \frac{|D_{\text{test}} - D_{\text{surr}}|}{\varrho_{\text{surr}}}, \quad (21)$$

式中 D_{test} 原始数据的关联维, D_{surr} 替代数据的关联维, ϱ_{surr} 替代数据的标准差.取显著水平 $\alpha = 0.05$, 则当 $S \geq 1.96$ 时,表明原始数据与替代数据有明显的差别,即原始数据以 95% 的置信水平为非线性.文献^[21]仿真验证了此方法的有效性.各数据集的

计算结果如表 4 所示,所有的判据 S 都明显大于 1.96,说明数据集具有强烈的非线性.

表 4 时间序列统计量的比较

数据集	时间序列	均值 \bar{y}/MPa	标准差 ϱ	关联维 D	判据 S	
A	原始序列	0.38	5.68×10^{-4}	4.9826	24.36	
	替代序列 1	0.38374	5.6762×10^{-4}	4.9749		
	替代序列 2	0.38374	5.6762×10^{-4}	4.9651		
B	原始序列	1.1535	1.0365×10^{-3}	4.9382	11.627	
	替代序列 1	1.1535	1.0365×10^{-3}	4.9501		
	替代序列 2	1.1535	1.0365×10^{-3}	4.9516		
C	原始序列	0.36859	2.2812×10^{-3}	3.557	6.3	
	替代序列 1	0.36859	2.2812×10^{-3}	3.5686		
	替代序列 2	0.36859	2.2812×10^{-3}	3.5688		
D	原始序列	3.4291	7.9081×10^{-3}	4.6936	3.445	
	替代序列 1	3.429	7.9081×10^{-3}	4.738		
	替代序列 2	3.429	7.9081×10^{-3}	4.5834		
E	原始序列	1.9141	2.6329×10^{-3}	3.9327	5.57	
	替代序列 1	1.9141	2.6329×10^{-3}	3.953		
	替代序列 2	1.9141	2.6329×10^{-3}	3.9245		
F	原始序列	0.16725	1.6725×10^{-3}	4.4645	12.157	
	替代序列 1	0.16725	1.6725×10^{-3}	4.4287		
	替代序列 2	0.16725	1.6725×10^{-3}	4.4312		
		替代序列 3	0.16725	1.6725×10^{-3}	4.4726	

7. 结 论

总结本文的研究,得到这样的结果:

1. 输油管道压力时间序列是一个非线性的物理过程;
2. 输油管道压力时间序列关联维具有分形特征且与压力值、温度等参数无关,是不依赖于环境的常量;
3. 输油管道压力时间序列的嵌入维约为 5 维;
4. 所有数据集的 Lyapunov 指数谱都有正的指数,所有指数之和为负数,且 Lyapunov 维数与关联维在同一个范围内,这些参数基本上与数据集的参数无关.

上述 4 个结果符合了混沌时间序列的条件,揭

示了输油管道压力具有严格混沌动态特征.

本文的研究结论说明输油管道压力时间序列内在动态是混沌现象而不是随机过程,压力的正常波动也不是白噪声,为从本质上研究输油管道压力时间序列提供了理论依据.从而为泄漏检测系统的分

析开辟了一条新的研究途径.即如果证明泄漏检测系统是混沌的,则可以流体运动的本质上对系统进行研究,可以通过短期的时间预测来预测系统的正常趋势,这样可以对微弱的泄漏信号进行判断,从而达到诊断泄漏故障的目的.

- [1] Shen G T , Li G H , Jing W K , Zuo Y T 2006 *Nondestructive Testing* **28** 267 (in Chinese) [沈功田、李光海、景为科、左延田 2006 无损检测 **28** 267]
- [2] Zhang D L , Wang S Q , Zhang M 2007 *Acta Petrolei Sin.* **28** 131 (in Chinese) [张东领、王树青、张敏 2007 石油学报 **28** 131]
- [3] Wang L N , Li J , Jin S J 2000 *Acta Petrolei Sin.* **21** 93 (in Chinese) [王立宁、李健、靳世久 2000 石油学报 **21** 93]
- [4] Feng J , Zhang H G 2005 *Acta Auto. Sin.* **31** 484
- [5] Wang Y S , Jiang W Z , Zhao J J , Fan H D 2008 *Acta Phys. Sin.* **57** 2053 (in Chinese) [王永生、姜文志、赵建军、范洪达 2008 物理学报 **57** 2053]
- [6] Chen L , Wang D S , 2007 *Acta Phys. Sin.* **56** 5098 (in Chinese) [谯龙、王德石 2007 物理学报 **56** 5098]
- [7] Yang D S , Zhang H G , Li A P , Meng Z Y 2007 *Acta Phys. Sin.* **56** 3121 [杨东升、张化光、李爱平、孟子怡 2007 物理学报 **56** 3121]
- [8] Wang Z S , Zhang H G , Wang Z L 2006 *Acta Phys. Sin.* **55** 2687 [王占山、张化光、王智良 2006 物理学报 **55** 2687]
- [9] Haykin S , Puthusserypady S 1997 *Chaos* **7** 777
- [10] Varadan V , Leung H , Bossé É 2006 *IEEE Trans. Instrumentation and Measurement* **55** 327
- [11] AntzHolger 1994 *Phys. Lett. A* **185** 77
- [12] Broomhead D S , King G E 1986 *Phys. D* **20** 217
- [13] Kennel M B , Brown R , Abarbanel H D I 1992 *Phys. Rev. A : Gen. Phys.* **45** 3403
- [14] Ding M Z , Grebogi G , Ott E , Sauer T , Yorke J A 1993 *Phys. Rev. Lett.* **70** 3872
- [15] Eckmann J P , Kamphorst S O , Ruelle D , Ciliberto S 1986 *Phys. Rev. A* **34** 4971
- [16] Ramazanetal G 1992 *Phys. D* **599** 142
- [17] Jan F 1984 *Phys. Rev. A* **29** 2928
- [18] Darbyshire A G , Broomhead D S 1996 *Phys. D* **89** 287
- [19] Banbrook M , Ushaw G , McLaughlin S 1997 *IEEE Signal.* **45** 1378
- [20] Theiler J , Eubank S , Longtin A , Galdrikian B , Farmer J D 1992 *Phys. D* **58** 77
- [21] Song A J , Han L 2007 *Acta Phys. Sin.* **56** 3821 [宋爱军、韩雷 2007 物理学报 **56** 3821]
- [22] Barnett A G , Wolff R C 2005 *IEEE signal.* **53** 26
- [23] Jiang K Y , Cai Z M , Lu Z B 2008 *Acta Phys. Sin.* **57** 1471 (in Chinese) [姜可宇、蔡志明、陆振波 2008 物理学报 **57** 1471]

Investigation of chaotic behavior for press time series of oil pipeline *

Liu Jin-Hai^{1)†} Zhang Hua-Guang¹⁾ Feng Jian¹⁾²⁾

1) (*School of Information Science and Engineering , Northeastern University , Shenyang 110004 , China*)

2) (*Key Laboratory of Process Industry Automation of Ministry of Education , Northeastern University , Shenyang 110004 , China*)

(Received 21 March 2008 ; revised manuscript received 12 May 2008)

Abstract

Usually , pressure fluctuation in oil pipeline is assumed to be stochastic or as white noise , which hinders the improvement of fault diagnosis technology based on pressure fluctuation . To find the internal dynamics of pressure fluctuation in oil pipeline , the possibility of existing chaotic behavior is validated using the method of nonlinear analysis . Six experimental data sets are studied on which phase spaces are reconstructed , the fractal dimensions and Lyapunov exponents are computed and the stationary feature and nonlinearity are validated . By the analysis of the results , the rigorous chaotic character of pressure fluctuation in oil pipeline is found , which is a theoretical basis for the correlative investigation based on pressure time series of oil pipeline .

Keywords : press time series , Lyapunov exponents , fractal dimensions , chaos

PACC : 0545 , 0555 , 4752

* Project supported by the National Natural Science Foundation of China (Grant Nos. 60534010 , 60572070 , 60521003 , 60774048 , 60774093 , and 60728307) , the Program for Changjiang Scholars and Innovative Research Groups of China , the Research Fund for the Doctoral Program of China Higher Education (Grant No. 20070145015) , the Natural Science Foundation of Liaoning Province , China (Grant No. 20062018) , the China Postdoctoral Science Foundation (Grant No. 20060400962) , and the National High Technology Research and Development Program of China (Grant No. 2006AA04Z183) . This Work is also supported in Part by Open Project Foundation of Key Laboratory of Process Industry Automation , Ministry of Education China (Grant No. PAL200503) .

† E-mail : jh_lau@126.com