

# 基于回声状态网络的 BOD 在线软测量模型

刘文波

(胜利油田电力管理总公司, 山东 东营 257200)

**摘要:** 针对污水处理过程的 BOD 建模问题, 提出了一种基于回声状态网络的 BOD 在线软测量方法; 基于梯度下降规则对回声状态网络的在线学习算法进行了研究; 为保证学习算法的收敛性, 基于 Lyapunov 理论对学习率范围进行了确定; 实验表明, 基于回声状态网络的在线 BOD 预测方法较常规神经网络预测精度提高约两个数量级, 模型的适应性也大幅提高。

**关键词:** 污水处理; BOD; 回声状态网络; 收敛性

## Online Biochemical Oxygen Demand Soft Measurement Based on Echo State Network

Liu Wenbo

(Power Management Corporation of Shengli Oil Field, Dongying 257200, China)

**Abstract:** In order to solve the modelling problem of biochemical oxygen demand (BOD) in wastewater treatment process, this paper proposes an online BOD predictive method based on echo state network (ESN). The gradient-based rule online algorithm is adopted to train the ESN model. To guarantee the convergence of the online learning algorithm, the range of the learning rate is determined based on Lyapunov theory. The experimental results demonstrate that the BOD prediction precision based on ESN is improved two orders of magnitude than conventional neural networks, and also the flexibility of the model is improved.

**Key words:** wastewater treatment; biochemical oxygen demand; echo state network; convergence

### 0 引言

城市污水处理过程的主要目标是对城市生产、生活污水进行处理, 使其满足国家出水水质要求。生化需氧量 BOD (Biochemical Oxygen Demand) 是污水处理过程中的关键质量指标之一<sup>[1-2]</sup>。由于仪器发展水平的限制, 目前 BOD 还无法实现在线的监测。由于 BOD 实验室检测存在滞后, 会导致水质超标不易被及时发现, 从而造成更大的后续污染。所以, 构建 BOD 的软测量模型, 实现 BOD 的实时监测, 对于提高污水处理过程的质量至关重要。

针对 BOD 的软测量问题, Scott<sup>[1]</sup>提出了一种基于机理模型的 BOD 软测量方法, 该方法通过构建污水处理过程的线性或非线性模型实现各出水指标的预测, 该模型存在大量需要在线辨识的参数, 应用极其不便。为降低建模的复杂性, 刘载文等<sup>[2]</sup>采用 RBF (Radial basis Function) 神经网络对出水 BOD 进行预测, 获得了较好的效果。神经网络 BOD 软测量模型实质上是采用一种高维非线性逼近的方式逼近易测量与 BOD 之间的输入输出关系。针对神经网络训练过程易陷入局部极小的问题, 田奕等<sup>[3]</sup>采用遗传算法对前馈神经网络的权值进行训练, 提高了 BOD 预测精度。研究表明, BOD 与易测辅助变量间存在复杂的时间动力学关系。而前馈神经网络实质上是一种静态映射关系描述, 其无法实现动力学关系的逼近。针对这一问题, 杨维维等<sup>[4]</sup>采用高阶递归神经网络实现了污水处理过程的建模, 其通过污水处理过程的易测量对出水 BOD、COD、

氨氮等进行了预测。由于递归神经网络能够逼近非线性系统的复杂动力学关系, 所以, 递归神经网络在 BOD 预测上获得了较前馈神经网络更高的精度。

目前递归神经网络建模仍存在两个问题: (1) 递归神经网络的结构复杂, 参数学习难度较大; (2) 目前神经网络建模多采取离线学习的方式, 这就需要神经网络的训练样本要有足够的代表性, 能够反映所有的工况信息。由于污水处理过程工况的复杂性, 获取反映所有工况的样本是比较困难的。所以, 实现神经网络的在线学习是解决这一问题的根本途径。为此, 提出了一种基于回声状态网络的 (Echo state network, ESN) 的 BOD 软测量模型, 并采用在线学习的方法实现网络权值的训练。

### 1 回声状态网络

#### 1.1 回声状态网络结构

回声状态网络是 Jaeger 于 2001 年提出的一种新型的递归神经网络<sup>[5]</sup>, 是池计算 (Reservoir computing, RC) 的典型代表<sup>[6]</sup>。ESN 的核心是一个动态神经元池 (Dynamic Neurons Reservoir, DNR)。神经元池内包含大量神经元, 这些神经元以完全随机和稀疏的方式连接<sup>[7]</sup>。Ozturk 指出, ESN 在结构上更接近于生物学网络<sup>[8]</sup>。ESN 只需确定隐层与输出层的连接权值, 这使得其学习非常快速, 同时缓解了传统人工神经网络学习过程中易陷入局部极小的问题。近十年来, ESN 在系统辨识、时间序列预测等领域表现出较高的精度和较强的应用前景<sup>[7-8]</sup>, 逐渐成为人工神经网络领域研究的热点之一。典型的 ESN 结构如图 1 所示, 其包括输入层、动态神经元池及输出层。

其中,  $u(k)$  为网络输入,  $x(k)$  为网络状态 (即隐层神

收稿日期: 2014-01-17; 修回日期: 2014-02-26。

作者简介: 刘文波 (1967-), 男, 山东东营人, 高级工程师, 硕士研究生, 主要从事神经计算与电力系统自动控制方向的研究。

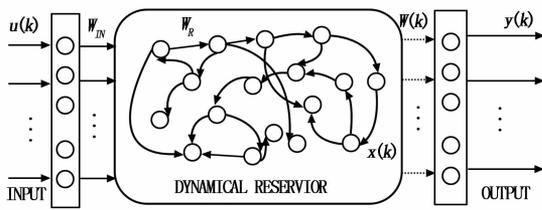


图 1 ESN 结构

经元输出),  $y(k)$  为网络输出, 其维数分别为  $K, N, L$ 。 $W_{IN}$ ,  $W_R$  分别为输入以及内部状态的连接权值矩阵, 其维数分别为  $N \times K$  和  $N \times N$ ,  $W$  为内部状态到输出的连接权值矩阵, 维数为  $N \times L$ 。ESN 的  $W_R$  以及  $W_{IN}$  在网络生成时确定, 并且在网络的学习和测试过程中保持不变。即 ESN 的学习只需确定  $W$  的值。

在不考虑输出层到输入层反馈的情况下, 内部神经元的输出为:

$$x(k) = f(W_{IN}u(k) + W_R x(k-1)) \quad (1)$$

$f$  为隐层神经元激活函数, 为 sigmoid 函数。

输出层神经元输出为:

$$y(k+1) = W^T x(k+1) \quad (2)$$

对于递归神经网络, 内部状态稳定是保证递归神经网络具有全局逼近能力的重要条件<sup>[5]</sup>。ESN 的内部稳定性可以通过选择合适的谱半径 ( $W_R$  的最大奇异值) 实现。

定理 1: 当谱半径  $\rho(W) < 1$  时, ESN 的内部状态是稳定的。

证明: 设  $x(k), x'(k)$  分别为 ESN 的两个不同的内部状态, 对于相同的输入  $u(k+1)$ , 根据式 (1) 有:

$$\|x(k) - x'(k)\| = \|f(W_{IN}u(k) + W_R x(k-1)) - f(W_{IN}u(k) + W_R x'(k-1))\| \quad (3)$$

由于  $f$  为 sigmoid 函数, 所以

$$\|x(k) - x'(k)\| = \|W_{IN}u(k) + W_R x(k-1) - W_{IN}u(k) + W_R x'(k-1)\| = \|W_R(x(k-1) - x'(k-1))\| \quad (4)$$

根据范数的相容性定理:

$$\|x(k) - x'(k)\| \leq \|W_R\| \|x(k-1) - x'(k-1)\| \quad (5)$$

显然, 当  $\rho(W) < 1$  时, (5) 式是收缩的, 即在输入不变的情况下, 当  $k \rightarrow \infty$  时,  $x(k), x'(k)$  将趋于一致, 即 ESN 的内部状态是稳定的。

### 1.2 回声状态网络的在线学习方法

从 (2) 式可以看出, ESN 的输出是各自独立的; 即:

$$y_i(k) = W_i^T(k)x(k) \quad (6)$$

$y_i(k)$  表示 ESN 的第  $i$  个输出。 $W_i(k)$  为与第  $i$  个输出相对应的权值列向量 (维数为  $N \times 1$ )。所以, 这里仅以  $W_i(k)$  的更新为例说明 ESN 的在线学习过程。

ESN 的第  $i$  个输出 ( $y_i(k)$ ) 与期望输出  $d_i(k)$  的偏差可表示为:

$$e_i(k) = y_i(k) - d_i(k) \quad (7)$$

则 ESN 的训练目标可设定为:

$$E(k) = \frac{1}{2} e_i^2(k) \quad (8)$$

根据梯度下降算法, ESN 的权值可通过下式更新:

$$\Delta W_i(k) = -\eta_i(k) \left[ \frac{\partial E(k)}{\partial W_i(k)} \right]^T \quad (9)$$

式中,  $\eta_i(k)$  为学习率。将 (7)、(8) 代入 (9), 可得:

$$\Delta W_i(k) = -\eta_i(k) e_i(k) \left[ \frac{\partial y_i(k)}{\partial W_i(k)} \right]^T \quad (10)$$

根据式 (2) 可知,

$$\left[ \frac{\partial y_i(k)}{\partial W_i(k)} \right]^T = x(k) \quad (11)$$

将 (11) 代入 (10) 可得:

$$\Delta W_i(k) = -\eta_i(k) e_i(k) x(k) \quad (12)$$

### 1.3 学习算法的收敛性分析

学习算法的收敛性是保证 ESN 在线学习性能的关键问题之一, 其直接影响到 ESN 模型的可用性。所以, 这里对 ESN 在线学习算法的收敛性进行分析。

定理 2: 若学习率  $\eta_i(k)$  满足

$$\eta_i(k) < \frac{2}{\|x(k)\|^2} \quad (13)$$

则 (12) 式所代表的 ESN 学习过程是收敛的。

证明: 定义离散的 Lyapunov 函数为:

$$L(k) = \frac{1}{2} e_i^2(k) \quad (14)$$

$e_i(k)$  的定义如式 (7) 所示。

ESN 学习过程中, Lyapunov 函数的增量为:

$$\Delta L(k) = \frac{1}{2} e_i^2(k+1) - \frac{1}{2} e_i^2(k) \quad (15)$$

显然,  $L(k) \geq 0$  成立, 并且仅当  $e_i(k) = 0$  时,  $L(k) = 0$ 。所以, 只要  $\Delta L(k) \leq 0$ , 则可以表明 ESN 的学习过程是收敛的。令:

$$\Delta e_i(k) = e_i(k+1) - e_i(k) \quad (16)$$

则 (15) 也可写为:

$$\Delta L(k) = \Delta e_i(k) \left[ e_i(k) + \frac{1}{2} \Delta e_i(k) \right] \quad (17)$$

根据全微分定理,

$$\Delta e_i(k) = \frac{\partial e_i(k)}{\partial W_i(k)} \Delta W_i(k) \quad (18)$$

将 (7)、(11) 代入 (18) 可得:

$$\Delta e_i(k) = x^T(k) \Delta W_i(k) \quad (19)$$

将 (12) 代入 (19), 可得:

$$\Delta e_i(k) = -\eta_i(k) e_i(k) \|x(k)\|^2 \quad (20)$$

将 (20) 代入 (17), 得:

$$\Delta L(k) = -\eta_i(k) e_i^2(k) \|x(k)\|^2 \left[ 1 - \frac{1}{2} \eta_i(k) \|x(k)\|^2 \right] \quad (21)$$

显然, 当

$$1 - \frac{1}{2} \eta_i(k) \|x(k)\|^2 \geq 0 \quad (22)$$

时,  $\Delta L(k) \leq 0$ , 并且仅当  $e_i(k) = 0$  时,  $\Delta L(k) = 0$ 。解不等式 (22) 即得到 (13) 式。

## 2 实验研究

### 2.1 BOD 预测实验

研究表明, 出水 BOD 主要与出水 COD, 溶解氧浓度, pH 值、以及固体悬浮物浓度 (Suspended solid, SS) 这 4 个量密切相关, 所以, 构建 4 输入 1 输出的 ESN 预测模型, 如图 2 所示。

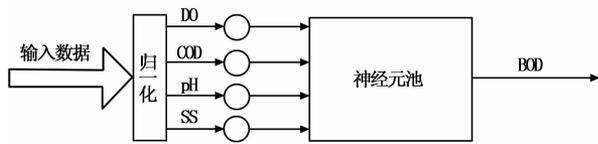


图 2 ESN 预测模型结构

由于输入数据是不同尺度的，ESN 隐层神经元为 Sigmoid 类型，当输入数据值过大时，神经元易陷入饱和区，这会大大降低 ESN 的预测性能。所以，对不同尺度的数据首先进行归一化处理。

$$s_1 = \frac{s - s_{\min}}{s_{\max} - s_{\min}} \quad (23)$$

$s_1$  表示归一化数据， $s$  代表原始数据信息。

实验数据来源于某小型污水处理厂 2008 年报表，该报表中包含 200 组数据。ESN 的隐层神经元个数为 50，连接度为 10%（即有 10% 的神经元与其它神经元存在连接，这样构造的神经元池是稀疏连接的，即  $W_R$  为一稀疏矩阵）。ESN 的谱半径设置为 0.9。BOD 的预测结果如图 3 所示（图 3 (a) 为 BOD 预测结果，图 3 (b) 为预测误差曲线）。

从图 3 (b) 可见，随着学习的进行，ESN 软测量模型的预测误差总趋势呈减小状态。

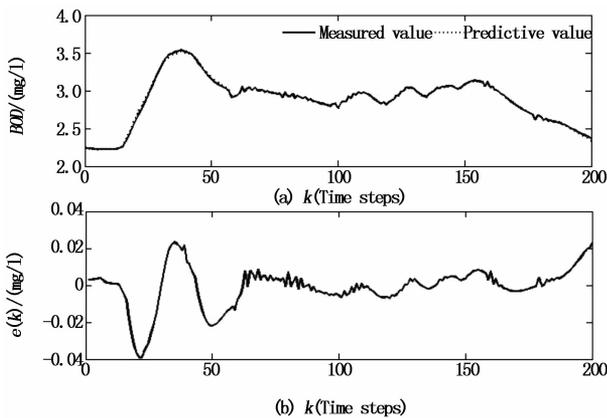


图 3 BOD 预测结果

表 1 不同网络 BOD 预测精度对比

神经网络类型	均方误差
前馈神经网络 (BP)	0.046
高阶递归神经网络 (HONN)	0.003 1
回声状态网络	$7.11 \times 10^{-5}$

表 1 列出了不同神经网络对 BOD 预测的精度对比，其中前馈神经网络及高阶递归神经网络的预测结果分别取自文献 [3] 和 [4]。从表 1 可见，ESN 对 BOD 的预测精度较 BP 网络及 HONN 网络提高了两个数量级，这说明 ESN 模型可以获得较高的预测精度。

### 2.2 BOD 模型收敛性实验

对于 ESN 的训练来说，除了训练误差趋近于 0 外，还要保证 ESN 的权值在学习过程中是收敛的。根据定理 2 的证明可知，若学习率满足 (13) 式，则当  $k \rightarrow \infty$  时，误差  $e_i(k) \rightarrow 0$ 。根据 (12) 式，当  $e_i(k) \rightarrow 0$  时， $\Delta W_i(k) \rightarrow 0$ ，即当学习率的选择满足定理 2 时，ESN 的权值也是收敛的。

图 4、5 为选择不同的学习率时 ESN 权值的收敛情况。图 4 (a)、(b) 的学习率值分别为  $\eta_i(k) = 2.05/\delta(k)$  及  $\eta_i(k) = 2/\delta(k)$  ( $\delta(k) = 1/\|x(k)\|^2$ )。从图 4 (a) 可见，当 ESN 的学习率  $\eta_i(k) > 2/\delta(k)$  时，ESN 的权值在学习过程中呈发散状态，此时 ESN 的学习过程是不稳定的。当  $\eta_i(k) = 2/\delta(k)$  时（图 4 (b)），ESN 的权值在学习过程中呈等幅震荡状态，此时 ESN 的学习过程处于临界稳定状态。

图 5 (a)、(b) 中的学习率分别为  $\eta_i(k) = 0.99/\delta(k)$  及  $\eta_i(k) = 1/\delta(k)$ 。在这两种情况下，学习率均满足定理 2，所以，ESN 的权值在学习过程中都是收敛的。但是当  $\eta_i(k) = 0.99/\delta(k)$ （图 5 (b)）时，ESN 学习的收敛速度明显加快。令

$$\varphi(k) = 1 - \frac{1}{2} \eta_i(k) \|x(k)\|^2 \quad (24)$$

根据式 (21) 可知，当  $\varphi(k)$  取得最大值时， $\Delta L(k)$  取得最小值，此时学习误差可以获得最快的下降速度（或最大梯度）。对 (24) 式求导，可得：

$$\frac{d\varphi(k)}{d\eta_i(k)} = 1 - \eta_i(k) \|x(k)\|^2 \quad (25)$$

当  $\eta_i(k) = 1/\delta(k)$  时， $\varphi(k)$  取得最大值时，此时，ESN 的学习可以获得最快的下降速度。

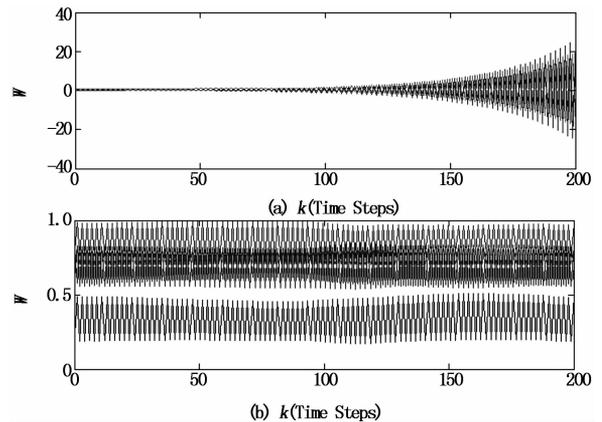


图 4  $\eta_i(k) \geq 2/\delta(k)$  时学习过程的权值变化

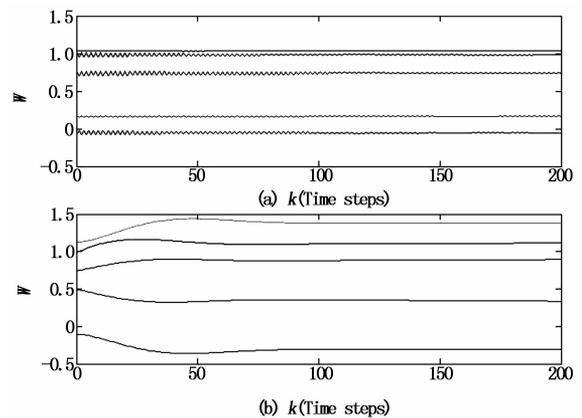


图 5  $\eta_i(k) < 2/\delta(k)$  的权值变化

### 2.3 BOD 模型适应性实验

为测试模型的适应性，添加该污水处理厂 2010 年的 200 组数据，采用在线和离线两种方式建立 BOD 预测模型。在离线建

模过程中,采用 2008 年的 200 组数据对 ESN 网络进行学习,2010 年的 200 组数据用于测试。图 6 显示了两种方式下 ESN 的建模效果(图 6 (a) 为在线学习,图 6 (b) 为离线学习)。

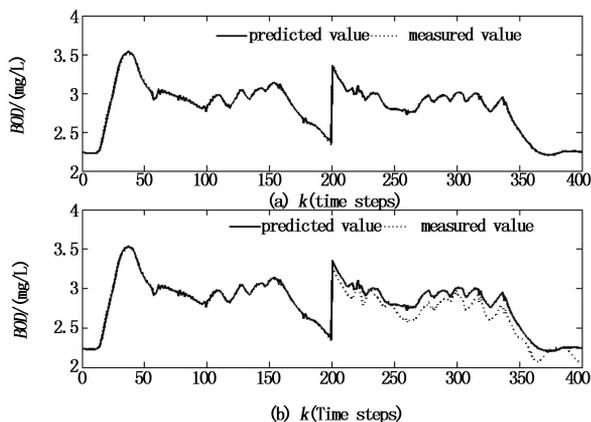


图 6 不同工况下的模型适应性对比

由图 6 (a) 可见,在线学习的情况下,ESN 预测模型能够在不同的工况下获得较好的预测结果。而离线学习(图 6 (b))的 ESN 模型虽然在学习时获得了较高的精度,但是,当工况发生变化时,模型预测的精度明显降低。这说明在线模型较离线模型有更好的适应性,其能够适应变化的情况。

### 3 结论

BOD 是污水处理过程的关键水质参数之一,对其进行实时的预测对于实现水质参数的闭环优化控制具有重要意义<sup>[9]</sup>。

BOD 的在线 ESN 模型在学习算法收敛的情况下,能够获得较高的精度,与离线的模型相比,在线 ESN 模型具有较好的工况适应性。

#### 参考文献:

[1] Dellana S A, West D. Predictive modeling for wastewater applications: Linear and nonlinear approaches [J]. Environmental Modelling & Software, 2009, 24 (1): 96-106.

[2] 刘载文,崔莉凤,祁国强,等. SBR 出水 BOD 值的 RBF 软测量法[J]. 中国给水排水, 2004, 20 (5): 18-20.

[3] 田奕,乔俊飞. 基于遗传算法的 BOD 神经网络软测量[J]. 计算机技术与发展, 2009, 19 (3): 127-133.

[4] 杨维维,乔俊飞. 基于高阶递归神经网络的污水处理系统建模[J]. 信息与控制, 2011, 40 (5): 710-714, 720.

[5] Jaeger H. The echo state approach to analysing and training recurrent neural networks [R]. GMD Report German National Research Center for Information Technology, 2001, 12 (8): 1-43.

[6] Lukochecksevicius, Jaeger H. Reservoir computing approaches to recurrent neural network training [J]. Computer Science Review, 2009, 3 (3): 127-149.

[7] Jaeger H. Harnessing Nonlinearity: Predicting Chaotic Systems and Saving Energy in Wireless Communication [J]. Science, 2004, 304 (5667): 78-80.

[8] Ozturk M C, Xu D G, Principe J C. Analysis and Design of Echo State Networks [J]. Neural Computation, 2007, 19: 111-138.

[9] 孙红,吴钱忠,王晓婉,等. BAF 小区生活污水处理智能控制系统应用[J]. 计算机测量与控制, 2013, 21 (5): 1233-1235.



(上接第 1332 页)

第二个例子使用了阿伦偏差来分析一个 4 GHz LO 信号。测试信号为调制 0.5 Hz 三角波 2 Hz 频偏的调频信号。53230A 设置无间隙测量速率为 10 ms,进行 1 000 次测量,产生 10 s 数据。阿伦偏差计算时  $\tau$  值范围从 10 ms 至 5 s,执行“all tau analysis”数据分析<sup>[4]</sup>。通过“all Tau analysis”可以很容易辨认低频周期干扰<sup>[4]</sup>。“all tau analysis”绘图使用 MATLAB 产生如图 6。

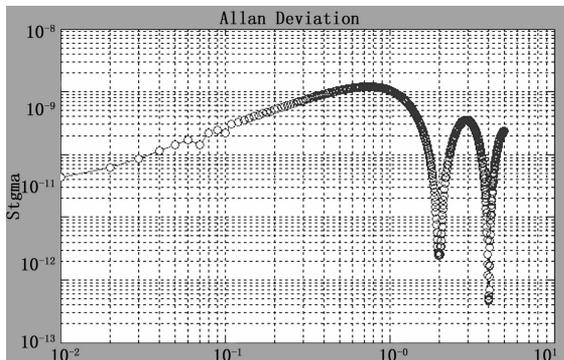


图 6 执行“all tau analysis”数据分析可以很容易辨认低频周期干扰

文献 [4] 揭示了“all tau analysis”数据分析中阿伦方差的极点以周期性干扰信号的半周期或全周期出现,在图 6 中,

阿伦方差极点对应的  $\tau$  值与附加的调制信号周期完全吻合,在阿伦偏差测量中对应某些  $\tau$  值的下降及趋势变化,揭示出影响信号寄生噪声频率范围。这种噪声分析尤其有助于查明接近被测信号的噪声来源,可以用来减少或消除噪声。

### 3 结束语

相位噪声参数是评估连续波信号频率短期稳定度的重要指标<sup>[5]</sup>,尤其是近载频相位噪声性能对雷达系统的作用距离,目标分辨率,数字通信系统的误码率等都有直接影响。本文论述的无间隙采样技术不仅可以方便地捕捉近载频相位噪声,还可以查找高性能振荡器噪声源的线索,从而减少甚至消除高性能振荡器中的噪声。

#### 参考文献:

[1] Agilent, Agilent Radar Measurements [Z]. 5989-7575EN, 70-76.

[2] Agilent, Agilent 53, 220A/53, 230A user guide [Z]. 53220-90417, p71.

[3] Riley W J. Handbook of Frequency Stability Analysis [M]. NIST Special Publication 1065, p59.

[4] Riley W J. Methodologies for Time - Domain Frequency Stability Measurement and Analysis [DB/OL]. <http://www.wiley.com/M-ETHODOLOGIES.pdf>, p8-9.

[5] 战永红,等. 一种改进的时钟定时抖动估计方法[J]. 计算机测量与控制, 2008, (3): 411-414.