

基于回声状态网络的电力负荷采样概率分布估计

针对长期电力负荷预测中存在的预测准确度不高的问题,提出基于回声状态网络的电力负荷采样概率分布估计预测方法。首先,对电力负荷时间序列预测的回声状态网络模型进行研究,在构建网络训练集和测试集的基础上,利用主成分分析方法实现数据维度降低,提高预测效率;其次,为提高预测准确度,采用分布估计算法对模型参数进行优化,同时为提高分布估计算法性能,基于Gibbs采样方式构建Markov链特征的概率条件分布模型,实现联合概率分布逐渐逼近;最后,通过在标准测试函数和法国电力负荷实测数据上的实验对比,验证所提算法有效性。

肖利群 李云川 李 燕/四川工商学院计算机学院

0 引言

基于观测数值电力负荷时间序列预测(TSF)问题,准确的短期负荷预测(STLF)可降低运营成本^[1]。在TSF问题中,常使用人工神经网络(ANN),如文献[2-3]等。后来,基于回归神经网络(RNN)的预测系统逐渐流行,特别是对混沌时间序列(TS)问题预测^[4]。许多文献研究了RNN网络和自回归积分滑动平均模型(ARIMA)预测性能。如文献[5]使用RNN预测模型结合前馈反向算法获得良好性能。文献[6]提出新颖前馈神经网络,类似于傅里叶分解方法,通过在每层引入余弦函数来表达剩余非线性特征。

近期,回声状态网络(ESN)得到广泛研究^[7],例如入侵检测、自适应控制、金融信用评级、谐波失真测量及语音识别等。在电网应用中,ESN已被单独用于短期负荷预测^[8]。ESN可赋予该系

统模拟任何复杂动态行为的能力,且具备内连接存储功能,可简化训练过程。在ESN电力负荷预测研究中,文献[9]将ESN用于多变量TSF问题;文献[10]利用ESN对混沌TS进行预测;文献[11]提出ARIMA与ESN的混合预测模型等。

传统TSF数据降维方法有奇异值分解(SVD)。例如文献[12]利用SVD获得 $k < m$ 系数集。文献[13]使用自回归(AR)模型对日间电力负荷进行预测,然后利用SVD算法实现数据降维。但SVD算法是对原始数据进行分解,其形式与主成分分解(PCA)算法相似,意义不直观,且相对PCA算法需进行矩阵运算,而PCA算法是对原始数据协方差矩阵进行特征分解,便于评价指标的直接应用。对此,这里采用PCA算法进行数据降维,同时考虑使用ESN模型进行预测,此外为提高ESN模型预测准确度,考虑采用分布估计算法进行模型参数优化。

1 电力负荷TSF

在TSF问题中,预测步数为 m , $X(\tau)$ 表示TS问题 X 在所有时刻上的值: $x_0, x_1, \dots, x_{\tau-1}, x_{\tau}$,目标是对 $x_{\tau+m}$ 预测。对通用预测模型观测值进行序列表达,并实现 m 超前数据的训练预测。训练集然后由有序值集 $\{(x_0, x_m), (x_1, x_{1+m}), \dots, (x_T, x_{T+m})\}$ 组成,考虑使用 T 组采样数据,并将数据对 $(x_{\tau}, x_{\tau+m})$ 作为输入,然后在时刻 τ 预测输出。需注意 $x_{\tau+m}$ 的假设取值并非只取决于 x_{τ} ,与以往所有输入有关。



肖利群/讲师

关键词/Keywords

主成分分析·
回声状态网络·
电力负荷·
分布估计·
Gibbs采样·
Markov链·

然后, 构建电力负荷数据训练和测试集, 从原始数据 X 中产生2个长度相同TS, $X_{\text{obs}}(\tau)$ 和 $X_{\text{fut}}(\tau)$

$$\begin{cases} X_{\text{obs}}(\tau) = (x_0, \dots, x_\tau) \\ X_{\text{fut}}(\tau) = (x_m, \dots, x_{\tau+m}) \end{cases} \quad (1)$$

$X_{\text{obs}}(\tau)$ 和 $X_{\text{fut}}(\tau)$ 可在原始数据序列 X 中去掉首 m 个及尾 m 个数据构建, 其长度均为 $N-m$ 。 $X_{\text{obs}}(\tau)$ 和 $X_{\text{fut}}(\tau)$ 进行数据分裂, 可得训练数据 $(X_{\text{obs}}^{\text{tr}}, X_{\text{fut}}^{\text{tr}})$ 和测试数据 $(X_{\text{obs}}^{\text{ts}}, X_{\text{fut}}^{\text{ts}})$, 此过程如图1所示。

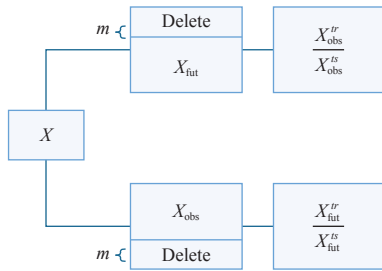


图1 训练集与测试集构建

基于ESN预测模型进行短期预测时, 可获得很高的预测准确度, 但随着 m 取值的增加, 预测准确度随之下降。常用的研究情形是对未来的24 h的电力负荷进行预测, 这里使用数据采集间隔为10 min, 则相应的预测间隔为 $m = 24 \times 6 = 144$ 。

2 PCA预测

2.1 PCA特性

TS问题 $M \in R^{n \times m}$, 协方差矩阵 $S_M \in R^{n \times m}$

$$S_M = \frac{1}{n-1} M^T M \quad (2)$$

式(2)为对称的 $m \times m$ 方形矩阵, 其中的对角项代表了列向量的方差 M_j , M_j 为 M 的列向量, $j = 1, \dots, m$ 。可通过正交矩阵 $C \in R^{m \times m}$ 识别来获得PCA, 实现给定协方差矩阵 S_M 的对角化。其列是特征向量和原始负荷矩阵 M 。因此 C 源自正交分解

$$A = CDC^T \quad (3)$$

式中, $A = M^T M \in R^{m \times m}$ 为半正定散射矩阵。 $D = R^{m \times m}$ 为对角矩阵, 对角元素为 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m$, 按递减排列, 为 $M^T M$ 的特征值。矩阵 C 是线性算子在新矩阵 $\beta = [\beta_1, \dots, \beta_m] \in R^{n \times m}$ 的投射 M , 则可得

$$M = \beta C^T \quad (4)$$

式中, C 为正交矩阵 ($C^T = C^{-1}$)。评价矩阵 β 可计算为

$$\beta = MC \quad (5)$$

自协方差矩阵的特征测量值用于表征 M 的“广义方差”, 该值越大表征数据主成分占据比例越大。当自协方差矩阵的特征测量值较小时, 表明存在冗余数据, 需要进行降维。定义累计贡献率 R_k 形式为

$$R_k = \frac{\sum_{j=1}^k \lambda_j}{\sum_{i=1}^m \lambda_i} \quad (6)$$

M 的近似值可构建如下

$$\bar{M}^{(k)} = \beta^{(k)} C^{(k)T}, \beta^{(k)} \in R^{n \times k}, C^{(k)} \in R^{m \times k} \quad (7)$$

式中, $\beta^{(k)} = [\beta_1, \dots, \beta_k]$ 和 $C^{(k)} = [c_1, \dots, c_k]$ 分别为矩阵 β 和 C 的前 k 列。

2.2 PCA降维

$\bar{M}^{(k)}$ 第 t 行元素为 $\bar{M}^{(k)}(t) = \sum_{j=1}^k \beta_j(t) c_j^T$, 其为原始TS数据长度 m 的第 t 个时间段逼近。对于第 t 个时段, 可表示成系数 $\beta_j(t)$, $j = 1, \dots, k$, 则TS问题的 m 步提前预测可用 k 组系数的一步提前预测求解, 例如 $t+1$ 时段值可近似为 $\sum_{j=1}^k \hat{\beta}_j(t+1) c_j^T$, 其中系数 $\hat{\beta}_j(t+1)$, $j = 1, \dots, k$ 是一步超前预测。由于主成分元素 $\beta_1, \dots, \beta_k$ 的正交性, 可认为TS问题是独立的, 可使用 k 个不同模型预测。

产生矩阵 M_{obs} 和 M_{fut} 分别为TS观测和预测值, 并分裂得到训练集 $(M_{\text{obs}}^{\text{tr}}, M_{\text{fut}}^{\text{tr}})$ 和测试集 $(M_{\text{obs}}^{\text{ts}}, M_{\text{fut}}^{\text{ts}})$ 。

为检索每个矩阵中的前 k 个主成分, 首先对 $M_{\text{obs}}^{\text{tr}}$ 和 $M_{\text{fut}}^{\text{tr}}$ 应用PCA。在测试阶段, 一旦有新测试元素输入, 可根据式(5)对其前 k 个主成分值近似。令 $\beta_{\text{obs}}^{tr(k)}$ 、 $C_{\text{obs}}^{tr(k)}$ 和 $\beta_{\text{fut}}^{tr(k)}$ 、 $C_{\text{fut}}^{tr(k)}$ 分别为 $M_{\text{obs}}^{\text{tr}}$ 和 $M_{\text{fut}}^{\text{tr}}$ 前 k 个主成分和负载。在给定测试集 $M_{\text{obs}}^{\text{ts}}$ 的第 t 行元素向量 $M_{\text{obs}}^{\text{ts}}$ 后, 第 k 个主成分的关联值可近似为

$$\bar{\beta}_{\text{obs}}^{tr(k)}(t) = M_{\text{obs}}^{\text{tr}}(t) C_{\text{obs}}^{tr(k)} \quad (8)$$

式中, $\bar{\beta}_{\text{obs}}^{tr(k)}(t)$ 为 $\beta_{\text{obs}}^{tr(k)}(t)$ 的近似值。类似地, 可对 $\bar{\beta}_{\text{fut}}^{tr(k)}(t)$ 和 $C_{\text{fut}}^{tr(k)}$ 进行估算。

2.3 k 成分预测

由于正交特性, 元素 β_j , $j=1, \dots, k$ 均可视为不同TS问题, 其中第 n 个值 $\beta_j(n)$ 仅依赖于以前的值 $\beta_j(t)$, $t=1, \dots, n-1$, 其为同一TS问题的观测值。利用TS问题的 $\beta_{\text{obs},j}^r$ 作为输入, $\beta_{\text{fut},j}^r$ 作为理想输出对模型 j 训练, 其中 $\beta_{\text{obs},j}^r$ 为矩阵 β_{obs}^r 的第 j 列, $\beta_{\text{fut},j}^r$ 为矩阵 β_{fut}^r 的第 j 列。对关注的第 t 个测试成分预测, 根据上节内容对 $\bar{\beta}_{\text{obs}}^{(k)}(t)$ 的前 k 个主成分检索。然后对于每个成分 $\beta_{\text{obs},j}^s$ 可评价其 m 步提前预测 $\hat{\beta}_{\text{fut},j}^s(t)$ 。

对 $t+m$ 间隔的数据, $\mathbf{C}_{\text{fut}}^{tr(k)}$ 预测公式为

$$\hat{\mathbf{M}}_{\text{fut}}^{ts}(t) = \hat{\beta}_{\text{fut}}^{ts(k)}(t) \mathbf{C}_{\text{fut}}^{tr(k)} \quad (9)$$

随测试数据集行数增加, 近似误差也增加, 因为 k 个成分正交特性无法保证。对此提出负载矩阵 $\mathbf{C}_{\text{obs}}^{tr(k)}$ 和 $\mathbf{C}_{\text{fut}}^{tr(k)}$ 更新机制, 将预测模型内部状态改为最近观测值, 并利用新的 $\mathbf{M}_{\text{obs}}^{tr}$ 和 $\mathbf{M}_{\text{fut}}^{tr}$ 计算 $\mathbf{C}_{\text{obs}}^{tr(k)}$ 和 $\mathbf{C}_{\text{fut}}^{tr(k)}$ 。基于PCA降维的 k 成分预测过程如下。

1) 初始化。根据电力负荷采集数据, 构建初始训练集 $(\mathbf{M}_{\text{obs}}^{tr}, \mathbf{M}_{\text{fut}}^{tr})$ 和测试集 $(\mathbf{M}_{\text{obs}}^{ts}, \mathbf{M}_{\text{fut}}^{ts})$, 并根据式(8)计算初始 $\beta_{\text{obs}}^{tr(k)}$ 、 $\mathbf{C}_{\text{obs}}^{tr(k)}$ 、 $\beta_{\text{fut}}^{tr(k)}$ 和 $\mathbf{C}_{\text{fut}}^{tr(k)}$ 。

2) M 近似值计算。利用上步所得 $\beta_{\text{obs}}^{tr(k)}$ 、 $\mathbf{C}_{\text{obs}}^{tr(k)}$ 、 $\beta_{\text{fut}}^{tr(k)}$ 和 $\mathbf{C}_{\text{fut}}^{tr(k)}$, 结合式(2)~式(7)计算TS问题 $\mathbf{M} \in R^{n \times m}$ 的近似值, 判断是否达到终止条件, 是则终止算法, 否则跳转步骤3)。

3) k 成分预测。利用式(9)进行 $\beta_{\text{obs}}^{tr(k)}$ 、 $\mathbf{C}_{\text{obs}}^{tr(k)}$ 、 $\beta_{\text{fut}}^{tr(k)}$ 和 $\mathbf{C}_{\text{fut}}^{tr(k)}$ 预测, 并根据预测值, 跳转步骤2), 进行 M 近似值计算。

3 ESN采样概率分布预测

3.1 最小二乘ESN求解模型

在这项工作中, 考虑使用两种不同的预测方法进行电力负荷预测: ①用于基准测试的经典ARIMA模型; ②回声状态网络, 其结构和训练过程在本节中描述。电力负荷预测的ESN模型示意图如图2所示。

该模型分三部分, 即输入层、可擦写存储器及读取结构。 N_i 维的输入向量 $x[n] \in R^{N_i}$, 作为可擦写存储输入, 其内部状态 $h[n-1] \in R^N$ 可更新为

$$h[n] = f_{\text{res}}(\mathbf{W}_i^r x[n] + \mathbf{W}_r^r h[n-1] + \mathbf{W}_0^r y[n-1]) \quad (10)$$

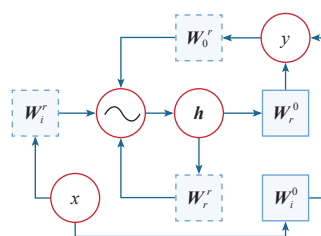


图2 电力负荷预测的ESN模型

式中, $\mathbf{W}_i^r \in R^{N_r \times N_i}$, $\mathbf{W}_r^r \in R^{N_r \times N_i}$ 以及 $\mathbf{W}_0^r \in R^{N_r}$ 初始为随机状态; $f_{\text{res}}(\cdot)$ 为非线性函数; $y[n-1] \in R$ 为网络前标量输出。在第二阶段中, ESN模型预测为

$$y[n] = (\mathbf{W}_i^0)^T x[n] + (\mathbf{W}_r^0)^T h[n] \quad (11)$$

式中, $\mathbf{W}_i^0 \in R^{N_i}$, $\mathbf{W}_r^0 \in R^{N_r}$ 为可训练权重。固定和自适应权重矩阵区别(见图2), 分别用实线和虚线表示。此外, 为提高整体稳定性, 在计算 $f_{\text{res}}(\cdot)$ 非线性变换前, 可插入小的均匀噪声项到状态更新方程中。

在 $\mathbf{W}_i^r$ 选择方面, 根据ESN理论, 存储器必须满足回声状态特性(ESP)。广泛使用的矩阵 $\mathbf{W}_i^r$ 重新调整经验规则是令 $\rho(\mathbf{W}_i^r) < 1$, 其中 $\rho(\cdot)$ 为操作算子谱半径。为确定权重矩阵, 需考虑 Q 输入-输出对序列

$$(x[1], d[1]), \dots, (x[Q], d[Q]) \quad (12)$$

输入为原序列 X 或第 i 个主成分 β_i , 输出为

$$d[t] = x[t+m] \quad (13)$$

式中, m 为预测期。输入根据式(10)设定为存储器输出, 产生内部状态序列: $h[1], \dots, h[Q]$ 。因此, ESN模型输出是无法定义的, 式(11)中的输出可用期望输出替代。状态被层叠在矩阵 $\mathbf{H} \in R^{Q \times N_i + N_r}$ 中, 期望输出表示为向量 $\mathbf{d} \in R^Q$

$$\mathbf{H} = \begin{pmatrix} x^T[1], h^T[1] \\ \vdots \\ x^T[Q], h^T[Q] \end{pmatrix}, \mathbf{d} = \begin{pmatrix} d[1] \\ \vdots \\ d[Q] \end{pmatrix} \quad (14)$$

此时所得训练问题是标准线性回归, 这里采用最小二乘回归(LSR)进行处理, 其正则化形式为

$$\mathbf{w}_{\text{ls}}^* = \arg \min_{\mathbf{w} \in R^{N_i + N_r}} \frac{1}{2} \|\mathbf{H}\mathbf{w} - \mathbf{d}\|_2^2 + \frac{\alpha}{2} \|\mathbf{w}\|_2^2 \quad (15)$$

式中, $\mathbf{w}_{\text{ls}} = [\mathbf{w}_i^0, \mathbf{w}_r^0]^T$, $\alpha \in R^+$ 为正标量, 式(15)可表达成如下闭合形式

$$\mathbf{w}_{ls}^* = (\mathbf{H}^T \mathbf{H} + \alpha \mathbf{I})^{-1} \mathbf{H}^T \mathbf{d} \quad (16)$$

对于任何 $N_r + N_i > Q$, 式 (16) 利用下式计算

$$\mathbf{w}_{ls}^* = \mathbf{H}^T (\mathbf{H}^T \mathbf{H} + \alpha \mathbf{I})^{-1} \mathbf{d} \quad (17)$$

3.2 参数优化

最小二乘ESN模型参数优化过程如图3所示。

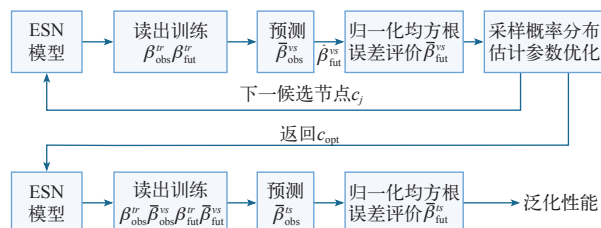


图3 参数优化算法框架

图3采用分布估计对参数优化, 所涉及参数搜索空间及最优搜索分辨率如下: ① 储存器神经元数目[100, 1 000], 分辨率为100; ② 权重 $\mathbf{W}_r^r$ 在 $[-1, 1]$ 中均匀提取, 设定比例初值 $p=0$, 对矩阵重新调整, 以获得所需谱半径 ρ^* , ρ^* 取值范围 $[0.5, 0.99]$, 存储层连通比例 p 范围 $[0.1, 0.4]$ 。③ 式 (10) 中的附加噪声项具有零均值, 方差为 β^2 , 搜索范围 $[0.000\ 01, 0.001]$ 。④ 所提供输入信号、期望响应及反馈信号通过三固定因子进行缩放, 分别为 γ_{in} 、 γ_{out} 和 γ_{feed} , 前两参数优化区间为 $[0.1, 1]$, γ_{feed} 优化区间为 $[0, 1]$, $\gamma_{feed}=0$ 表示ESN无反馈信号。

4 采样概率分布估计

4.1 算法框架

这里提出Gibbs采样概率分布估计算法 (SPEDAs)。若 X 为随机高维向量, 其联合概率密度为 $f_X(x)$, 因其具备不规则随机性, 导致无法对其直接进行建模。对此构建概率分布 $f_{X_i|X^{(i)}}(x_i|x^{(i)})$ 模型, 具体流程如图4所示。图4框架采用Gibbs

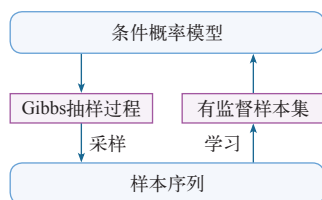


图4 采样概率模型

采样建立Markov链特征的联合概率分布模型 $f_{X_i|X^{(i)}}(x_i|x^{(i)})$, 实现逐渐逼近联合概率分布模型 $f_X(x)$ 的效果。

算法的Gibbs采样是对于每个变量, 都按照其概率条件数据及当前种群个体的分布特点, 通过随机形式构建新的种群个体, 利用此方式可构建Markov链分布特征的样本采样序列, 其数学过程描述为

$$\lim_{T \rightarrow \infty} \sum_{t=1}^T \phi[X(t)] / T = E\{\phi(X)\} \quad (18)$$

式中, 可对生成的Markov链序列进行表达, 该序列同联合分布模型采样所得样本性质相同。

4.2 样本序列训练

按照有监督训练方式, 执行自学习调整

$$Q_i = \{(x^{(i)}, x_i) | x \in P\} \quad (19)$$

式中, 对模型 $f_{X_i|X^{(i)}}(x_i|x^{(i)})$ 参数进行估计, $(x^{(i)}, x_i)$ 表示样本序列 $x^{(i)}$ 均值期望为 x_i , 为简化算法, 设定 X_i 取值区间为 $X_i \in \{-1, 1\}$ 。则概率条件模型 $f_{X_i|X^{(i)}}(x_i|x^{(i)})$ 的估计步骤可表述为: ① 结合样本序列 Q_i 及相应的分类算法对样本分类器进行构建; ② 利用所构建样本分类器实现模型 $f_{X_i|X^{(i)}}(x_i|x^{(i)})$ 参数分布估计。分类界面可定义如下

$$g = \arg \min_{g \in G} \left\{ \sum_Q L[y, g(x)] \right\} \quad (20)$$

式中, $Q = \{(x, y)\}$ 为样本的有监督序列; $L[y, g(x)]$ 为样本损耗指标; G 为函数候选集, 作用是屏蔽分类输出不理想数据; 所构建的分类界面可表示成 $g(x)=0$ 。则概率条件估计模型可表达成

$$\hat{P}(Y|X) \exp\{-L[Y, g(X)]\} \quad (21)$$

式中, 基于前述损耗函数并结合分类过程最大似然类型函数, 可获得模型式 (21) 等价形式

$$g = \arg \min_{g \in G} \left\{ \prod_{(x,y) \in Q} \hat{P}(y|x) \right\} \quad (22)$$

4.3 算法计算流程

假定 $(x_{opt}, fit_{opt}) = \text{IEDA}(eval)$ 是EDAs的模型表达形式, 模型中 $eval$ 表示评价指标, 那么改进的EDAs计算过程如伪代码2采样概率分布估计所述。

(1) 算法初始化

$P = \{x(i)|i=1, \dots, N\}$, $x(i) \sim \text{Uniform}(S)$, $P_{\text{opt}} = \emptyset$;

(2) 执行迭代过程

//寻优

for $ep=1$; $epoch$ do

$x^* = \arg \max_{x \in P} \{eval(x)\}$; $P_{\text{opt}} = P_{\text{opt}} \cup \{x^*\}$; $P = P \setminus \{x^*\}$;

endfor

//学习

for $i=1 : D$ do

$P_i = \{(x^{(i)}, x_i)|x \in P_{\text{opt}}\}$; $g_i = \arg \min_{g \in G} \left\{ \sum_{\mathcal{Q}_i} L[y, g(x)] \right\}$;

$P(X_i | X^{(i)}) \propto \exp\{-L[X_i, g(X^{(i)})]\}$;

endfor

//采样

$P_{\text{new}} = \text{sample}(N - N_0, \{\hat{P}(X_i | X^{(i)})\})$;

//更新

$P = P_{\text{new}} \cup P_{\text{opt}}$; $P_{\text{out}} = \emptyset$;

(3) 算法输出

$x_{\text{opt}} = \arg \max_{x \in P} \{eval(x)\}$; $fit_{\text{opt}} = eval(x_{\text{opt}})$.

5 实验分析

5.1 分布估计算法性能测试

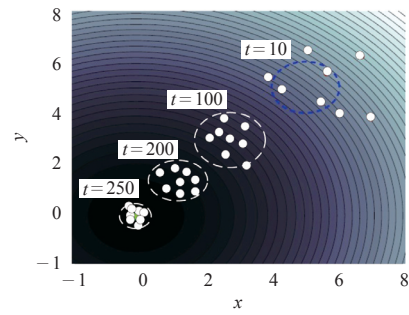
测试函数选取标准函数，仿真平台：系统为Windows7，软件为Matlab2013a，CPU为AMD X4 641，RAM为6G ddr3-1600。对比算法为标准EDAs及文献[14]环境识别分布估计算法(EIEDAs)。测试函数为

$$f_1(X) = \sum_{i=1}^n [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2] \quad (23)$$

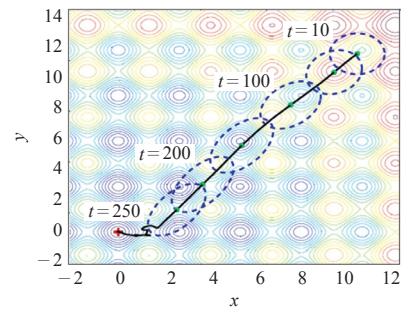
$$f_2(X) = \left[\sum_{i=1}^n x_i^2 - \prod_{i=1}^n \cos(x_i / \sqrt{i}) + 1 \right] / 4000 \quad (24)$$

参数设置：终止代数300，终止阈值 $\lg(f) = 10^{-7}$ ，种群规模1024，淘汰率为0.5，算法覆盖空间为 $100 \times 1024 \approx 10^5$ 。首先利用式(13)分别在二维和三维情况下对算法寻优过程进行仿真，如图5和图6所示。

图5a种群规模为8时，对于给定初始种群为

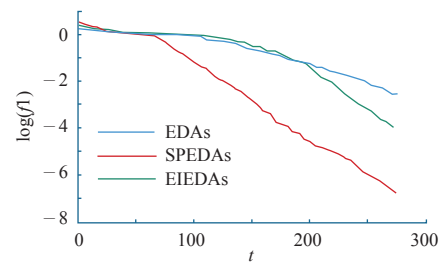


(a) 二维函数进化过程

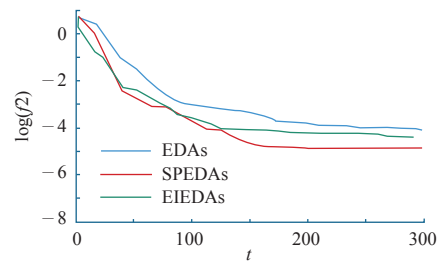


(b) 三维函数进化过程

图5 二维和三维寻优过程



(a) 测试函数式(23)对比结果



(b) 测试函数式(24)对比结果

图6 算法对比结果

$x \in [4, 8]$, $y \in [4, 8]$, 全局最优值点在 $x=y=0$ 处，可看出种群从开始的分散到逐渐聚敛，从最初的离最优点位置较远到后来逐渐靠近全局最优值附近，体现了算法的智能寻优过程。图5b为抽样概率分布估计算法每隔一定代数的最优值进化过程，从三维图像进化过程看，算法能够有效躲避局部极值点的诱惑，最终到达全局最优点。

如图6所示, 算法初始空间为 $x \in [0, 1]$, $y \in [0, 1]$, $n = 10$ 。从图6a可看出, SPEDAs与EDAs和EIEDAs相比, 性能优势较明显, 在运行到283步时, 即达到最小终止阈值 $\lg(f) = 10^{-7}$, 而EIEDAs此时收敛准确度为 $\lg(f) \approx 10^{-4}$, EDAs为 $\lg(f) \approx 10^{-2}$ 。从图6b可看出, 三种算法在某准确度上均出现早熟收敛, 均未达到最小终止阈值 $\lg(f) = 10^{-7}$, 而是在最大终止迭代数为300时算法终止。但在收敛准确度上SPEDAs收敛准确度 $\lg(f) \approx 10^{-5}$, EDAs和EIEDAs的收敛准确度均 $\lg(f) \approx 10^{-4}$, 其中EIEDAs要略优于EDAs。上述过程验证了SPEDAs的有效性。

5.2 电力负荷预测实验

实验设置: Intel i7-5440 4-core CPU @2.80 GHz, 8 GB RAM。实验电力负荷数据来自文献[15]。时间自2005年2月到2006年12月共100周的电力负荷采集数据。采用绝对误差均值 (MAPE) 评价指标为

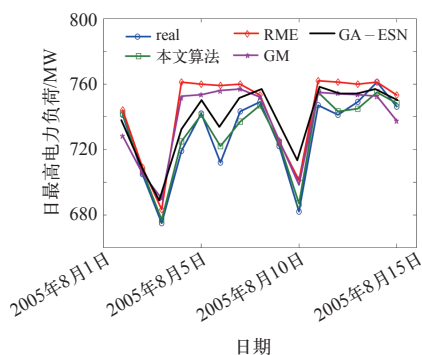
$$\text{MAPE} = 100 \left[\frac{\sum_{i=1}^n |(L_i - \hat{L}_i)|}{(n \cdot L_i)} \right] \quad (25)$$

式中, L_i 及 $\hat{L}_i$ 分别为数据集中第*i*天真实电力负荷及预测数值; n 为采样预测点数。MAPE越小, 代表算法的负荷预测准确度越准确。

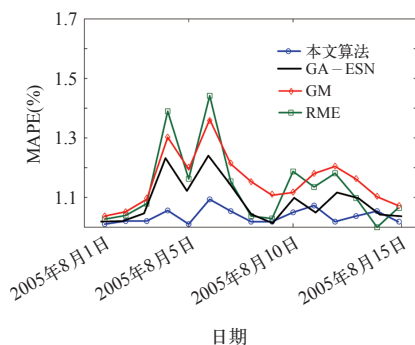
负荷预测准确度: 选取灰度值电力负荷预测 (GM) 算法、比率中位数估计 (RME) 算法及遗传算法的ESN电力负荷预测 (GA-ESN) 进行对比。图7为算法预测值同真实值的对比曲线 (该曲线由每天真实负荷最大值与算法预测负荷最大值构成)。

图7为上述算法在2005年8月1~15日, 日最高预测负荷值同真实值及其MAPE指标。由图7a可看出, 在预测准确度上, 本文预测值要比RME、GM及GA-ESN算法预测值更接近真实值, 说明其预测准确度最高。同时GA-ESN预测准确度要优于GM预测准确度, RME预测准确度最差。这在图7b所示的预测值MAPE指标上得到印证。

接下来考虑固定时间负荷预测准确度对比, 固定时间设为9:00与21:00, 考虑10天为周期的



(a) 每天负荷预测最高值



(b) MAPE指标对比

图7 负荷预测准确度

电力负荷MAPE指标, 具体预测数据如图8所示, 对比算法仍选取RME、GM及GA-ESN算法。

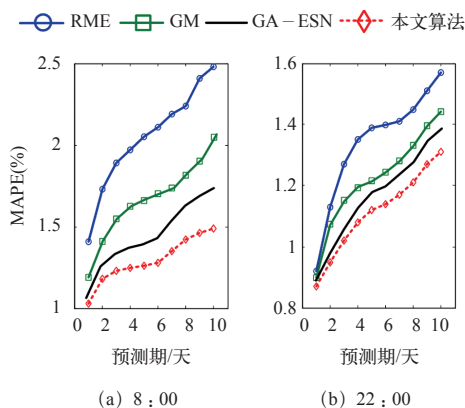


图8 负荷MAPE预测指标

根据图8指标可知, 上述算法均可有效地进行负荷值预测, 但随着预测期延长, 其预测值MAPE指标均呈现增长态势, 预测准确度下降, 这符合真实逻辑。同时可知, 本文算法在周期内的MAPE预测指标要优于RME、GM以及GA-ESN算法的MAPE预测指标, 这表明本文算法在固定时刻的负

荷预测值仍然是相对最优的。

负荷预测时间：与预测准确度指标相比，预测算法运行时间要求相对要低，但对于算法评价角度仍给出对比算法在图7每天负荷预测最高值上的运行时间对比情况，如下表所示。

表 预测时间对比

算法	运行时间/s
RME	15.3
GM	11.4
GA-ESN	35.9
本文算法	28.7

根据上表对比数据可知，在算法运行效率上，GA-ESN算法所用时间最长，达到35.9 s，其次为本文算法，这两种算法的运行时间较长，主要是因为采用群智能优化算法所致。RME、GM两种算法的运行时间相对较短，算法实现较为简单，这体现了群智能优化算法在计算效率上的相对劣势，但是考虑到电力负荷预测过程中，对于计算时间的要求相对较宽松，所以上述时差并不能造成实质性的影响。

6 结束语

提出基于PCA回声状态网络的电力负荷采样概率分布估计预测方法，用来提高长期电力负荷预测过程中存在的预测准确度不高的问题。利用Gibbs采样方式的分布估计改进算法可实现PCA回声状态网络状态参数的高准确度优化，通过实验验证了算法的预测准确度优势。但是，因为所采用群智能优化算法相对实现较为复杂，导致算法的运行时间相对较长，如何提高算法的收敛效率是一项有意义的工作。

参考文献

- [1] Ruiz P A, Gross G. Short-term resource adequacy in electricity market design [J]. IEEE Transactions on Power System, 2008, 23 (3) : 916-926.
- [2] Xiao Zhiqiang, Liang Shunlin, Wang Jindi. Use of general regression neural networks for generating the GLASS leaf area index product from time-series

MODIS surface reflectance [J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52 (1) : 209-223.

- [3] Rocio Cogollo M, Velasquez J D. Are neural networks able to forecast nonlinear time series with moving average components [J]. IEEE (Revista IEEE America Latina) Latin America Transactions, 2015, 13 (7) : 2292-2300.
- [4] Hayashi H, Shibasaki T, Shima K. A recurrent probabilistic neural network with dimensionality reduction based on time-series discriminant component analysis [J]. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26 (12) : 3021-3033.
- [5] Chandra R. Competition and collaboration in cooperative coevolution of elman recurrent neural networks for time-series prediction [J]. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26 (12) : 3123-3136.
- [6] Erol Egrioglu, Ufuk Yolcu, Cagdas Hakan Aladag, et al. Recurrent multiplicative neuron model artificial neural network for non-linear time series forecasting [J]. Neural Processing Letters, 2015, 41 (2) : 249-258.
- [7] Ding Haiyan, Pei Wenjiang, He Zhenya. A multiple objective optimization based echo state network tree and application to intrusion detection [C]. Proceedings of 2005 IEEE International Workshop on VLSI Design and Video Technology, IEEE, 2005: 443-446.
- [8] Rabin M J A, Hossain M S, Ahsan M S. Sensitivity learning oriented nonmonotonic multi reservoir echo state network for short-term load forecasting [C]. 2013 International Conference on Informatics, Electronics & Vision (ICIEV), Dhaka, IEEE, 2013: 1-6.
- [9] Jayawardene I, Venayagamoorthy G K. Comparison of echo state network and extreme learning machine for PV power prediction [C]. 2014

- IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG), Orlando, FL, IEEE, 2014: 1-8.
- [10] Fink O, Zio E, Weidmann U. Fuzzy classification with restricted boltzman machines and echo-state networks for predicting potential railway door system failures [J]. IEEE Transactions on Reliability, 2015, 64 (3) : 861-868.
- [11] Niu D, Ji L, Xing M, et al. Multi-variable echo state network optimized by bayesian regulation for daily peak load forecasting [J]. Journal of Networks, 2012, 7 (11) : 1790-1795.
- [12] Scardapane S, Wang D, Panella M. A decentralized training algorithm for echo state networks in distributed big data applications [J]. Neural Network, 2015, 32 (8) : 236-243.
- [13] Peng Y, Lei M, Li J B, et al. A novel hybridization of echo state networks and multiplicative seasonal ARIMA model for mobile communication traffic series forecasting [J]. Neural Computing and Applications, 2014, 24 (4) : 3-8.
- [14] Peng Xingguang, Gao Xiaoguang, Yang Shengxiang. Environment identification-based memory scheme for estimation of distribution algorithms in dynamic environments [J]. Soft Computing, 2011, 15 (3) : 311-326.
- [15] Yacine Chakhchoukh, Patrick Panciatici, Lamine Mili. Electric load forecasting based on statistical robust methods [J]. IEEE Transactions on Power Systems, 2011, 26 (3) : 982-992.

(收稿日期: 2016-01-18) **EA****Weidmüller**

魏德米勒Energy Meter电力参数测量仪 —— 高效透明的生产电能监测

魏德米勒Energy Meter电力参数测量仪不仅能监测电能的消耗, 它还能监测电能质量的基本参数和分析来自不同导体的电流的细微差别, 使电力行业用户能够快速地了解生产设备中电能参数的变化情况, 保证其稳定和高效地使用。具有综合功能的Energy Meter电力参数测量仪使您可以将生产点的能源网络分块测量, 尽可能精确地进行监测, 配备有一体的液晶显示屏, 重要的监测数据如电压、电流、功率和能耗都能一目了然。

扫一扫, 关注魏德米勒官方微信



Let's connect.