

# AVC系统电压预测及控制研究

卢文浩<sup>1</sup>, 余天伦<sup>1</sup>, 杨立洪<sup>1</sup>, 魏勇军<sup>2</sup>, 刘有志<sup>2</sup>, 李东旭<sup>2</sup>, 胡 杨<sup>2</sup>

<sup>1</sup>华南理工大学数学学院, 广东 广州

<sup>2</sup>广州供电局电力调度控制中心, 广东 广州

Email: 197720327@qq.com

收稿日期: 2020年9月1日; 录用日期: 2020年9月18日; 发布日期: 2020年9月25日

---

## 摘 要

通过建立支持向量机回归模型, 对某一段母线电压进行预测。并通过几次模型参数的优化, 将误差逐渐减小, 取得令人满意的结果。利用模糊信息粒化方法, 获取电压曲线的特征。结合优化的支持向量机回归方法, 提出模式匹配方法对模型进行优化。实验结果表明, 该方法能取得进一步的改进结果。

## 关键词

支持向量机回归, 参数优化, 模糊信息粒化, 模式匹配

---

# Research on Prediction and Control of Voltage in AVC System

Wenhao Lu<sup>1</sup>, Tianlun Yu<sup>1</sup>, Lihong Yang<sup>1</sup>, Yongjun Wei<sup>2</sup>, Youzhi Liu<sup>2</sup>, Dongxu Li<sup>2</sup>, Yang Hu<sup>2</sup>

<sup>1</sup>School of Mathematics, South China University of Technology, Guangzhou Guangdong

<sup>2</sup>Power Dispatching Control Center of Guangzhou Power Supply Bureau, Guangzhou Guangdong

Email: 197720327@qq.com

Received: Sep. 1<sup>st</sup>, 2020; accepted: Sep. 18<sup>th</sup>, 2020; published: Sep. 25<sup>th</sup>, 2020

---

## Abstract

The voltage of a bus bar is predicted by the Support Vector Machine Regression Model. And through several times of parameters optimization, the error is reduced gradually and the result is satisfying. With the method of Fuzzy Information Granulation, we can get the feature of voltage curve. By combining this method with the previous one, we propose the pattern matching method in order to optimize our model. The result shows that we can achieve further improvement with the combined method.

文章引用: 卢文浩, 余天伦, 杨立洪, 魏勇军, 刘有志, 李东旭, 胡杨. AVC 系统电压预测及控制研究[J]. 应用数学进展, 2020, 9(9): 1604-1611. DOI: 10.12677/aam.2020.99188

## Keywords

The Support Vector Machine Regression, Parameter Optimization, Fuzzy Information Granulation, Pattern Matching

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

目前自动电压控制系统中用来控制电压的阈值线是固定的, 这种设定方法可能导致不必要的调压操作。本文希望提出一种能根据历史数据的变化, 智能地设定阈值的方法, 以提高电压调节的效率以及电压合格率。该方法的核心是对电压进行预测。本文利用历史数据中不同日期同一时刻的电压值作为输入变量, 预测未来某时刻的电压值, 然后根据电压值给出阈值设定方法。

## 2. 支持向量机回归

考虑样本如下

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_k, y_k)\}, x_i \in R^n, y_i \in R$$

考虑线性拟合的情况, 用下面的函数来拟合以上样本

$$f(x) = \omega \cdot x + b \quad (2.1)$$

假设函数(1)能够以精度  $\varepsilon$  零误差地拟合样本, 那么得到:

$$\begin{cases} y_i - \omega \cdot x_i - b \leq \varepsilon \\ \omega \cdot x_i + b - y_i \leq \varepsilon \end{cases}, i = 1, 2, \dots, k \quad (2.2)$$

考虑到可能允许出现的拟合误差的情况, 引入松弛变量  $\xi_i \geq 0$  和  $\xi_i^* \geq 0$ , (2.2)变成

$$\begin{cases} y_i - \omega \cdot x_i - b \leq \varepsilon + \xi_i \\ \omega \cdot x_i + b - y_i \leq \varepsilon + \xi_i^* \end{cases}, i = 1, 2, \dots, k \quad (2.3)$$

通过最小化  $(1/2)\|\omega\|^2$  来控制函数集的复杂度。再考虑到拟合误差的情况, 线性回归问题转化成二次优化问题

$$\min \frac{1}{2}\|\omega\|^2 + C \sum_{i=1}^k (\xi_i + \xi_i^*) \quad (2.4)$$

其中,  $C > 0$ , 为一常数, 作为超出拟合精度  $\varepsilon$  的样本的惩罚因子。前一部分代表了函数的推广能力, 后一部分代表了经验误差(即拟合函数对训练样本的误差), 两个合起来就是统计学理论中提到的结构风险最小化。采用优化方法得到其对偶问题, 在约束条件:

$$\begin{cases} \sum_{i=1}^k (\alpha_i - \alpha_i^*) = 0, \\ 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, 2, \dots, k \end{cases} \quad (2.5)$$

下, 对拉格朗日因子  $\alpha_i, \alpha_i^*$ , 最大化目标函数

$$W(\alpha, \alpha^*) = -\varepsilon \sum_{i=1}^k (\alpha_i + \alpha_i^*) + \sum_{i=1}^k y_i (\alpha_i - \alpha_i^*) - \frac{1}{2} \sum_{i,j=1}^k (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)(x_i \cdot x_j) \quad (2.6)$$

可以求得回归函数

$$f(x) = (\omega \cdot x) + b = \sum_{i=1}^k (\alpha_i^* - \alpha_i)(x_i \cdot x) + b^* \quad (2.7)$$

当函数  $f(x)$  为非线性时, 通过一种事先规定好的映射  $\phi(x): R^n \rightarrow R^m (m \geq n)$ , 将样本投影到的一个特征空间(Hibert 空间), 进行线性回归。可以通过满足 Mercer 条件[1]的核函数  $k(x, x_i)$  代替在其高维特征空间内积运算。公式(2.7)用核函数替换, 得到特征空间的线性函数:

$$f(x) = \langle W, \phi(x) \rangle + b \quad (2.8)$$

本文利用 LIBSVM [2]工具箱实现支持向量机回归算法。其中最重要的参数是 SVM 方法、核函数及其参数  $g$ , 惩罚因子  $C$ , 以及在本问题中输入变量的维度  $N$ 。根据经验先采用  $\varepsilon$ -SVR 及高斯径向核函数,  $C=2, g=1, N=7$ , 检验模型是否有效。得到结果如下图 1、图 2 所示:

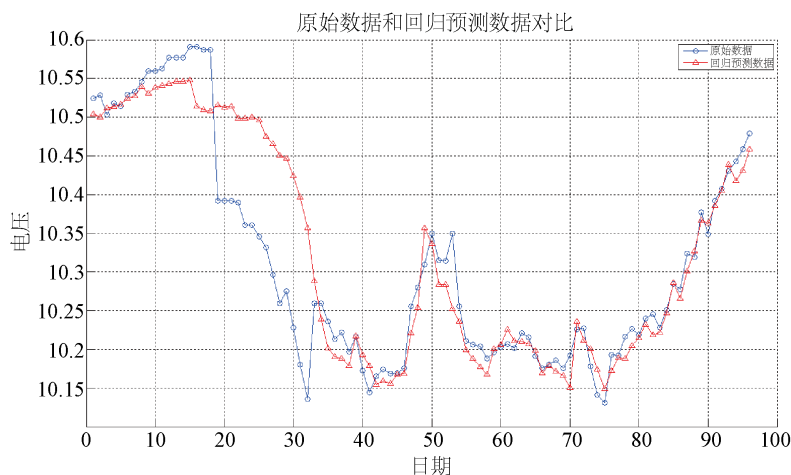


Figure 1. Comparison of forecast data and original data in the 8th day of 2012  
图 1. 2012 年第 8 天预测数据及原始数据对比

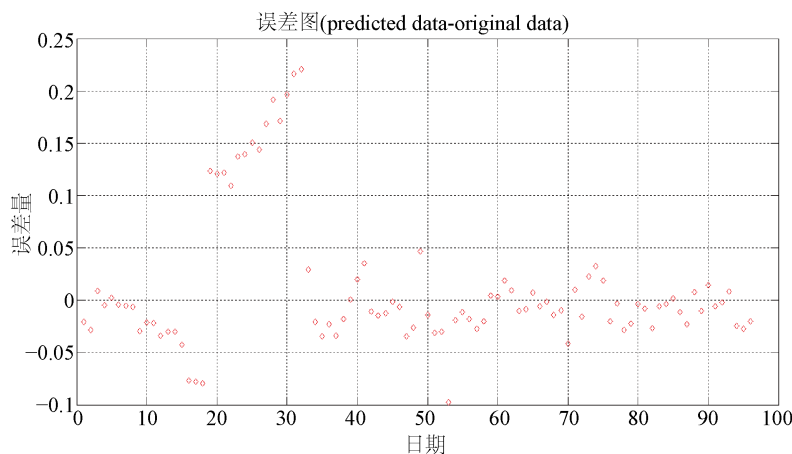


Figure 2. Absolute error of forecast data on the 8th day of 2012  
图 2. 2012 年第 8 天预测数据绝对误差

从图 1、图 2 中结果看来，对于 2012 年第 8 天的电压预测，模型是有效的。我们定义一个绝对误差平均值来衡量对全年预测的效果。绝对平均误差：

$$abs\_error = \frac{1}{96 \times 359} \sum_{i=1}^{96} \sum_{j=8}^{366} |error_{ij}| \tag{2.9}$$

其中， $abs\_error$  为绝对误差平均值，是预测误差取绝对值后的平均值。 $error_{ij}$  是每个预测点的绝对误差。

### 3. 对参数进行选择优化

对比不同 SVM 方法下，不同核函数的绝对误差平均值，我们得到表 1、表 2：

**Table 1.** Comparison of fitting effect of different kernel functions of  $\epsilon$ -SVR method

**表 1.**  $\epsilon$ -SVR 方法不同核函数的拟合效果对比

核函数类型	abs_error	运行时间(秒)	svmtrain 参数选项
线性核函数	0.0734	1.750270	'-c 2 -g 1 -t 0'
多项式核函数	112.2287	206.220078	'-c 2 -g 1 -t 1'
高斯径向核函数	0.0721	1.771225	'-c 2 -g 1 -t 2'
Sigmoid 核函数	0.1120	1.801441	'-c 2 -g 1 -t 3'

**Table 2.** Comparison of fitting effect of different kernel functions of v-SVM method

**表 2.** v-SVM 方法不同核函数的拟合效果对比

核函数类型	abs_error	运行时间(秒)	svmtrain 参数选项
线性核函数	0.0724	2.994160	'-c 2 -g 1 -t 0'
多项式核函数	68	700.043180	'-c 2 -g 1 -t 1'
高斯径向核函数	0.0716	2.373454	'-c 2 -g 1 -t 2'
Sigmoid 核函数	0.1126	1.117464	'-c 2 -g 1 -t 3'

可以看出在 v-SVM 下，利用高斯径向核函数能取得最小的绝对误差平均值，因此模型 SVM 方法跟核函数就取为 v-SVM 及高斯径向核函数。

通过基于正交验证的网格寻优方法，将参数  $C$  和参数  $g$  在某区间内以某步长离散的取值，然后根据不同参数组合下的均方误差，取其最小值的组合作为我们模型的参数输入。先在大区间，利用大步长进行粗略的搜索，如图 3 所示。

其中， $CVmse$  的定义如下：

$$CVmse = \sqrt{\frac{1}{n \times m} \sum_{i=1}^n \sum_{j=1}^m (x_{ij} - \hat{x}_{ij})^2} \tag{3.1}$$

其中， $\hat{x}_{ij}$  为电压  $x$  矩阵中，电压  $x_{ij}$  的预测值。 $m$ 、 $n$  为要预测电压的行数和列数。

我们的目标是要找出  $CVmse$  最小的区域，然后进行精细搜索。观测  $CVmse$  等高线图，我们缩小搜索范围得到精细结果，如图 4 所示。

于是，得到目前最优的参数组合  $C = 0.125$ ， $g = 3.7321$ 。

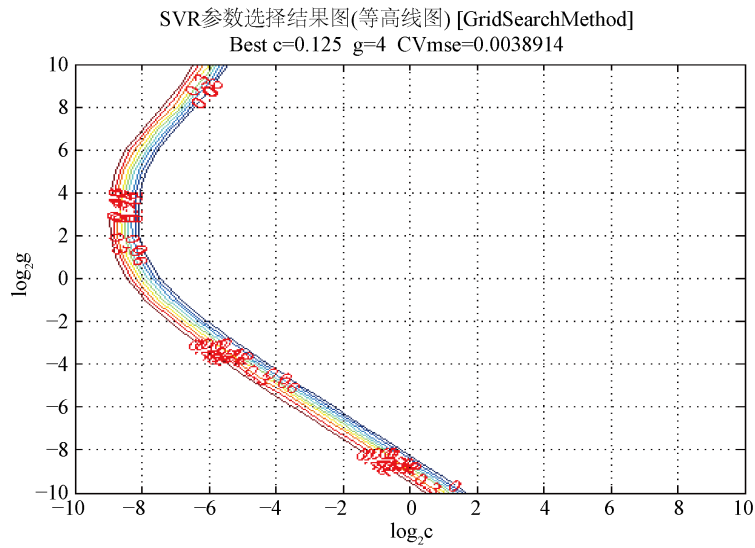


Figure 3. Rough selection of CVmse contour map with the lowest mean square error

图 3. 粗略选择结果最低均方误差 CVmse 等高线图

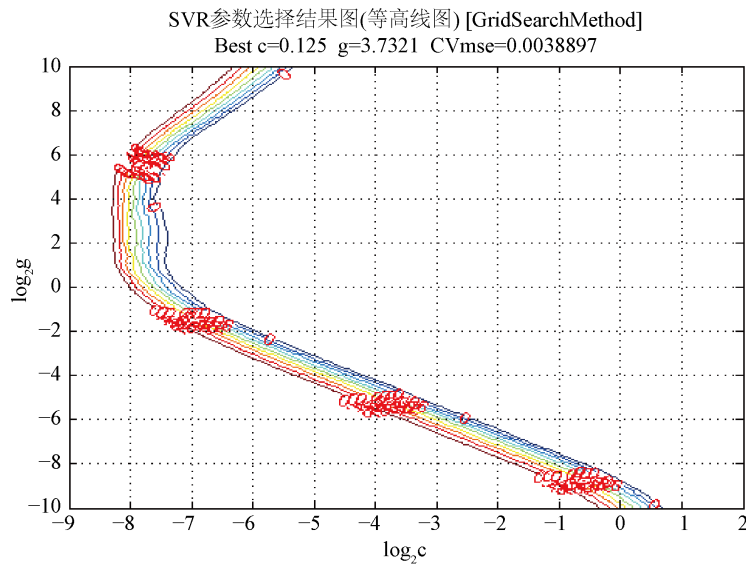


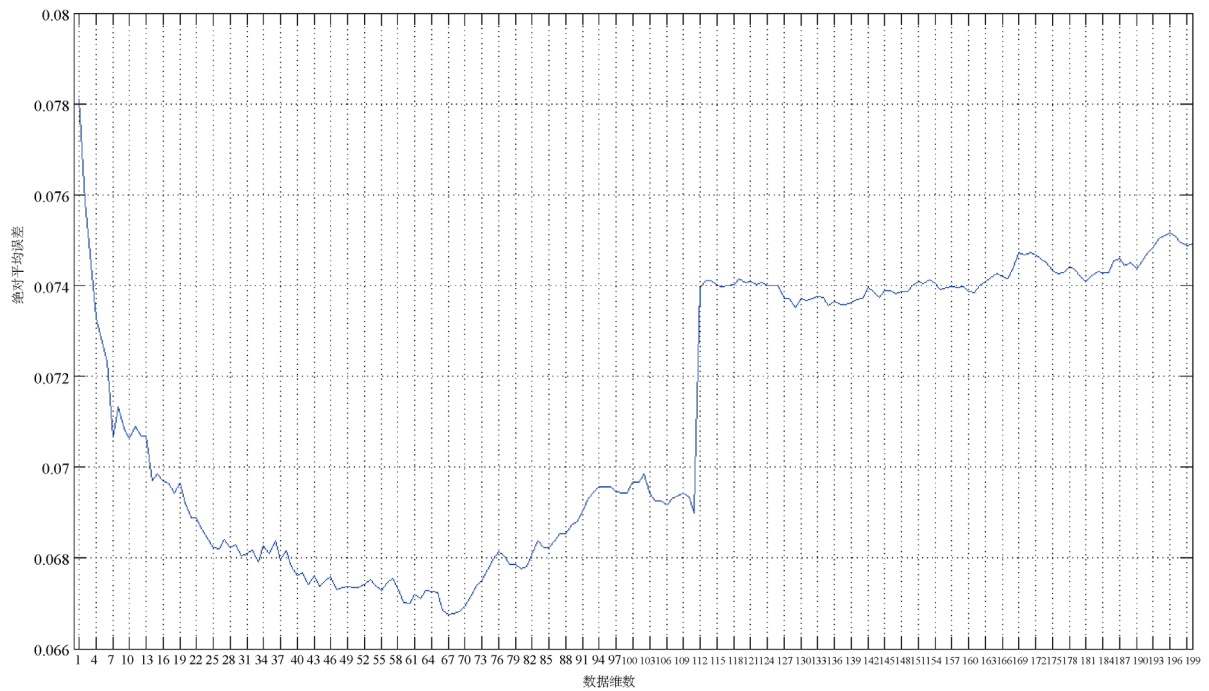
Figure 4. Precision selection of CVmse contour map with the lowest mean square error

图 4. 精细选择结果最低均方误差 CVmse 等高线图

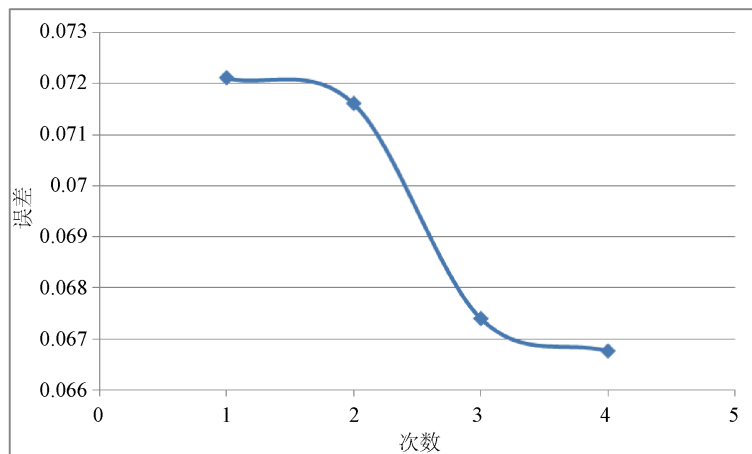
利用上面确定的 SVM 方法、核函数、参数  $C$  跟  $g$ ，对比不同输入维度  $N$  下的绝对误差平均值，并画出其图形如图 5 所示。

从结果上看，维数 7 是区间  $[1, 8]$  中的一个极小值点，因此选择 7 天的电压数据作为输入变量，具有一定的合理性。但是纵观全局，在一个更大的区间  $[1, 200]$  内，存在一个最小值点在 67 处，其绝对平均误差为 0.066766，相对于之前输入维数为 7 的时候的误差 0.0674，取得了更进一步的优化。

至此，我们已经逐步将模型的所有参数进行了优化，同时，也确定了整个模型重要参数的最好选择方式，使得预测结果的误差得到不断地缩小，几次优化的结果如图 6 所示。



**Figure 5.** Graph of absolute mean error with respect to dimension of input data  
**图 5.** 绝对平均误差关于输入数据维数的关系图



**Figure 6.** Prediction error results of several parameter optimization  
**图 6.** 几次参数优化的预测误差结果

#### 4. 模糊信息粒化

模糊信息粒化[3]就是以模糊集的形式表示信息粒子。用模糊集对时间序列进行信息粒化，主要步骤可以分为：粒化与模糊化。粒化就是将给定的时间序列划分为若干个小的子序列，每个小序列被称为一个窗口；而模糊化就是将粒化产生的每一个窗口进行模糊化，生成一个个模糊集也就是模糊粒子。将这两个步骤结合在一起，就是模糊信息化，称为 f-粒化[4]。

下面我们采用三角型模糊函数[5]以一个小时(即 4 个数据)为一个窗口的大小，一天共 24 个窗口(小序列)将 2012 年第一天电压数据的模糊信息粒化，提取主要特征：

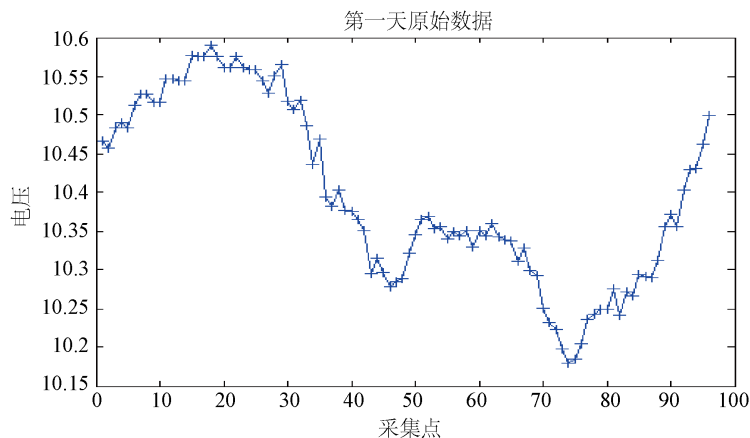


Figure 7. Original voltage data of the first day of 2012  
图 7. 2012 年第一天电压原始数据图

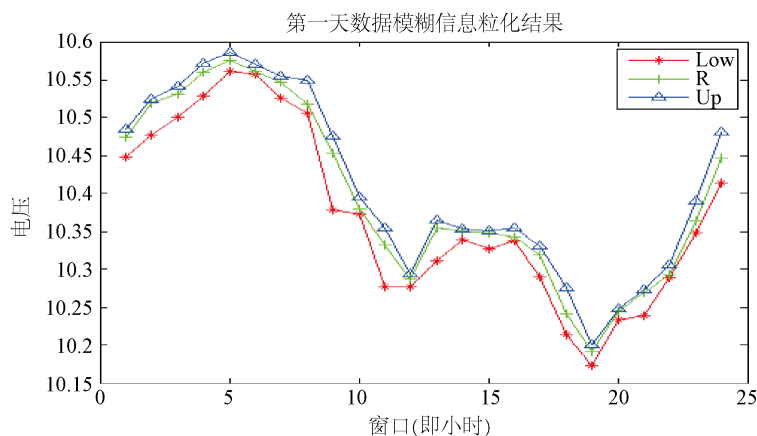


Figure 8. The first day data fuzzy information granulation results  
图 8. 第一天数据模糊信息粒化结果

从上面两个图，图 7、图 8，我们可以看出来经过模糊信息粒化之后，数据的基本特征本都被保留下来了。其中 Low, R, Up 分别对应着原始数据在每个窗口内变化的最小值，平均值和最大值。但是，数据由原来的 96 个，变成了由 Low、R、Up 各 24 点组成的特征向量组，起到了一定的提取特征、降维的作用[6]。

定义第  $i$  天与第  $j$  天电压序列的距离为：

$$D(i, j) = D\_Low(i, j) + D\_R(i, j) + D\_Up(i, j) \tag{4.1}$$

其中， $D\_Low(i, j)$ 、 $D\_R(i, j)$ 、 $D\_Up(i, j)$  分别为第  $i$  天与第  $j$  天电压序列模糊信息粒化后 Low、R、Up 的距离。显然， $D(i, j)$  越小，第  $i$  天与第  $j$  天电压序列就越相似。

利用前面经过优化的 SVM 回归方法，先求出要预测那天的电压数据，然后以该序列作为查询序列，采用模糊信息粒化之后，跟模糊信息粒化后历史序列去匹配，最后以匹配结果作为最终预测结果。依然从 301 天开始作为测试数据集，在该天前的数据全部作为历史数据集。

用前面的方法预测第 301 天的电压数据，以及使用模糊信息粒化匹配，得到第 301 天电压结果如下图 9 所示：

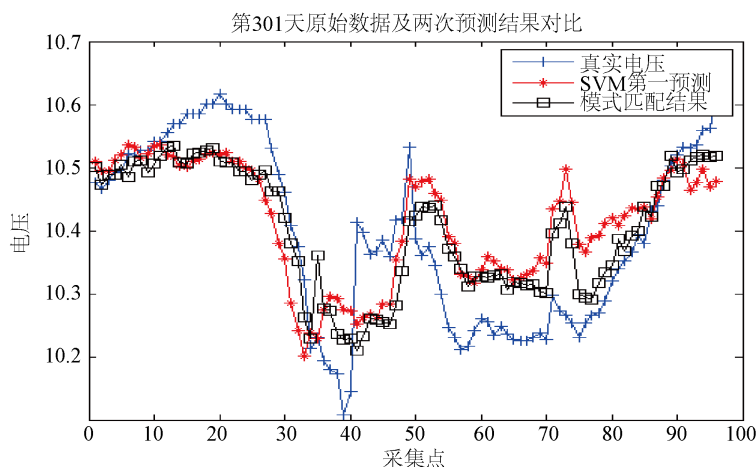


Figure 9. Real voltage and two prediction results on the 301st day

图 9. 第 301 天真实电压及两次预测结果

从图 9 上看, 结果还是比较令人满意的, 如果我们同样求第 301 天到第 366 天的绝对平均误差, 会发现, 其值为 0.068251, 相比之下, 还是组合方法稍微精确一点。

## 5. AVC 系统调压阈值设置策略

根据我们预测出来的结果, 如果发现预测值只是超过警戒线, 并没有超过合格线, 那么可以在该预测时间点提前将警戒线调整至更靠近合格线的位置。比如, 当预测值为 10.63 的时候, 可以将上警戒线调整到 10.7; 当预测值为 10.04 时, 将下警戒线调整到 10.0。

这样做, 我们就可以避免一些不必要的调整动作, 避免电压的频繁调节, 从而保证了理想的电压合格率。

## 6. 结束语

本文通过参数优化, 利用支持向量机回归模型较准确地预测出了未来的电压值。结合模糊信息粒化方法, 定义模式匹配算法, 得到最终的电压预测方案, 取得了令人满意的结果, 并根据预测结果给出了阈值的设定策略。

## 基金项目

资助项目: 中国南方电网科技项目《基于 BART 算法和超吸收壁 Brown 运动的 AVC 系统定值智能学习与过程控制技术研究》。

## 参考文献

- [1] Vapnik, V.N. 统计学习理论的本质[M]. 张学工, 译. 北京: 清华大学出版社, 2000.
- [2] Chang, C.-C. and Lin, C.-J. (2019) LIBSVM—A Library for Support Vector Machines. <http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html>
- [3] Zadeh, L.A. (1997) Toward a Theory of Fuzzy Information Granulation and Its Centrality Inhuman Reasoning and Fuzzy Logic. *Fuzzy Sets and Systems*, **90**, 111-117. [https://doi.org/10.1016/S0165-0114\(97\)00077-8](https://doi.org/10.1016/S0165-0114(97)00077-8)
- [4] Zadeh, L.A. (1979) Fuzzy Sets and Information Granularity. North Holland, Amsterdam.
- [5] 李洋, 史峰, 王小川, 等. MATLAB 神经网络 30 个案例分析[M]. 北京: 北京航空航天大学出版社, 2010: 112-152.
- [6] 喻胜华, 肖雨峰. 基于信息粒化和支持向量机的股票价格预测[J]. 财经理论与实践(双月刊), 2011, 32(6): 44-47.