

基于多因素加法模型的中期电力负荷预测^①

翁金芳, 黄伟, 江育娥, 林劼

(福建师范大学 软件学院, 福州 350017)

摘要: 提前准确预测所需电力负荷, 做好电力规划是电力部门保证电力供应稳定不可或缺的重要环节. 基于欧洲智能网络(EUNITE)竞赛电力数据和北美电力数据, 提出一种多因素加法模型, 进行中期电力预测. 考虑到温度、假期、星期等因素对电力负荷产生不同的影响, 拟合出这些因素与电力负荷之间的映射关系, 相加得到电力负荷预测的函数. 还比较了业界常用的7种不同的算法模型, 使用6种不同指标对这些模型和多因素加法模型进行评估, 实验结果发现, 在这8种不同算法模型中, 多因素加法模型有着更加精确的预测性能, 运算速度比其他模型快, 并且模型更加容易理解和解释.

关键词: 负荷预测; 多因素加法模型; 中期电力预测; 人工智能算法; 数据挖掘

Multiple Factors Addictive Model for Mid-Term Electric Load Forecasting

WENG Jin-Fang, HUANG Wei, JIANG Yu-E, LING Jie

(Faculty of Software, Fujian Normal University, Fuzhou 350017, China)

Abstract: Accuracy forecasting of electric load is important for power system to make plan. A Multiple Factors Addictive (MFA) model is proposed to predict mid-term electric load based on Europe (EUNITE) competition dataset and North American electric dataset. Firstly, MFA considers factors such as temperature, holiday, and week separately to fit functions for electric loads. And then all these fitted functions are added together to a unified function, which is used to make prediction of the electric load. Seven other state-of-art algorithms which are popular in the field are also used to make forecasting. The performances of prediction models are evaluated by using 6 different metrics. Compared with 7 other kinds of different models prediction results, MFA has the advantages of more accurate forecasting performance and faster operational speed, and is simple and easy to understand.

Key words: load forecasting; multiple factors additive model; mid-term load forecast; artificial intelligence algorithms; data mining

电力负荷预测(Electrical power load forecasting)是电力系统规划和运行不可或缺的、有效的部分^[1]. 为了维护电力系统的有效运行, 对负荷的变化有一个事先的估计, 负荷预测成为电力规划乃至商业和工业正常运行的一个重要的保障. 负荷预测是指预测未来某时刻所需最大的电力负荷数据, 根据目的的不同可分为超短期、短期、中期和长期预测. 其中中期负荷预测(mid-term load forecasting)是指预测时间长度可以是未来一个月或者未来一年中的负荷量^[1,2]. 由于较长的时

间跨度, 中期负荷预测可能存在误差的累积, 使得预测精度降低, 这就会影响到电力系统的运行. 因此, 提高中期负荷预测的精度成为研究热点^[3-6].

电力负荷预测研究已有五十多年历史, 有传统经典的预测方法和现代预测方法^[2]. 经典的预测方法包括趋势外推法^[3]、回归分析法^[7]和时间序列法^[8], 其中自回归滑动平均(ARIMA)是时间序列法中广受欢迎的一种方法^[9]. 现代预测方法主要有专家系统方法^[10]、灰色系统理论^[11]、混沌理论^[12]、小波分析^[13]、神经网络

^① 基金项目:国家自然科学基金(61472082);福建省自然科学基金(2014J01220)

收稿时间:2015-06-07;收到修改稿时间:2015-09-24

理论^[14]、模糊集理论^[15]等方法。

随着人工智能的发展,支持向量机(SVM)、分形和数据挖掘理论等新方法被用来预测电力负荷^[2]。支持向量机(SVM),由 Vapnik 及其同事提出的,是一种基于统计学习理论(SLT)的机器学习方法^[16]。近年来,诸多学者试图改进支持向量机进行更加精确的负荷预测。翟永杰等人提出一种模糊支持向量机(FSVM)算法^[4]。刘遵雄等人采用相关向量机(RVM)进行中期电力负荷预测^[5],是一种与支持向量机(SVM)相似的稀疏概率模型。Bo-Juen Chen 等人在 MATLAB 上采用支持向量机(SVM)模型,用交叉方法来确定参数组^[6]。随着 Vapnik 提出的 ϵ -不敏感损失函数,支持向量机被扩展到解决回归问题的支持向量回归(SVR)^[16]。支持向量回归(SVR)是支持向量机的重要应用分支,SVR 解决中期负荷预测问题时具有高精度,如,方瑞明使用的基于自组织特征映射(SOFM)聚类的支持向量回归模型(SVR)^[17]。Ehab E. Elattar 等人通过局部加权回归(LWR)的支持向量回归(SVR)算法和加权分布距离算法来改进精度^[1]。Chih-Hung Wu 等人提出一种新的混合遗传算法(HGA),是提高搜索回归精度的支持向量回归(SVR)的最佳类型^[18]。人工智能技术是中期负荷预测(MTLF)重要的算法,而人工神经网络(ANN)得到了相当多的关注。Nima Amjady 等人提出神经网络与进化算法混合的预测机制方法^[19]。Babinec Stefan 等人通过反赫布型学习规则的自适应优化的回声状态神经网络,优化了神经元隐层的权值^[20]。Huina Mao 等人结合了基于贝叶斯神经网络和支持向量机的自组织模糊神经网络(SOFNN)学习方法^[21]。

本文提出一种多因素加法模型(multiple factors additive, 简称为 MFA)来预测中期电力负荷。使用的数据为欧洲智能技术网络(EUNITE)竞赛电力数据^[22]和北美电力数据^[23]。我们同时还采用了业界认可的一些常用算法:自回归积分滑动平均模型(ARIMA),线性模型(LM),随机森林(Random Forest, 简称为 RF),多元自适应样条回归(MARS),支持向量机(SVM),神经网络(NNET),和决策树(DT)等。从 6 个不同的评估指标来对这些模型与多因素加法模型进行对比后发现多因素加法模型有不少优势:(1)模型预测性能优于其他模型,(2)运行时间少,(3)模型易于理解和解释。换句话说,多因素加法模型不仅提供了一种简单容易解释的模型,而且提高了预测的准确度,降低了运行时间。

本文的结构如下。第 1 节说明了本文所用到的 7 个算法和多因素加法模型。第 2 节介绍了实验环境和实验数据,并且重点解释了多因素加法模型函数的构建过程。第 3 节阐述了评估模型的 6 个预测性能指标。第 4 节呈现了实验结果,并且进行相应的分析。第 5 节总结了本文。

1 实验方法

本章介绍了实验环境,实验数据,采用的 8 个不同算法模型,并且重点介绍了多因素加法模型。

1.1 实验环境

本文所有编程实现都以 R 语言为工具,版本为 3.1.2^[24]。所用机器配置为: Intel(R) Core(TM)i5CPU M460@2.53GHz, 4GBRAM。

1.2 实验数据

斯洛伐克东部电力公司(East-Slovakia Power Distribution Company)于 2001 年 8 月组织了一次全球性的负荷预测竞赛(EUNITE)^[6],并提供了 1997 年到 1998 年每半小时的负荷数据、日平均温度、以及假期的信息。该竞赛要求预测出 1999 年 1 月每日中每半小时的最大电力负荷。本文选取每日中每半小时的最大电量、平均温度、节假日和星期作为输入变量。我们用 1997 年、1998 年的数据作为训练集,预测 1999 年 1 月每日每半小时最大电力负荷量。

另一个数据是北美电力数据,包含有 1985 年 1 月 1 日到 1992 年 10 月 12 日每日每半小时的温度和电量。为了跟 EUNITE 竞赛数据一致,本文也收集了北美地区的法定节假日和星期信息作为另外两个输入变量。数据从 1985 年 1 月 1 日到 1992 年 10 月 12 日,共 2842 天的数据,同样包括时间、温度、星期、节假日和每日每半小时最大电量。我们用 1985 年 1 月到 1992 年 8 月的数据作为训练集,预测出 1992 年 9 月 1 日到 1992 年 10 月 12 日的负荷数据。

1.3 使用到的 8 种算法

除了多因素加法(MFA)模型以外,本文还用了业界常用的 7 中不同的算法:自回归积分滑动平均(ARIMA)模型,线性(LM)模型、多元自适应样条回归模型(MARS)、支持向量机模型(SVM)、随机森林模型(Random Forest,RF)、决策树模型(DT)和神经网络模型(NNET)。ARIMA 和 LM 是使用 R^[24]提供的 stats 的软件包^[25],MARS 使用的是 earth 软件包^[26],SVM 使用的是

e1071 软件包^[27], RF 使用的是 randomForest 软件包^[28], DT 使用软件包 rpart^[29], NNET 使用软件包 nnet^[30], 除了 ARIMA 算法以外, 其他算法的参数都采用 R 语言的默认值. 由于 ARIMA 算法的参数很重要, 我们试了大量的参数, 选了其中表现最好的参数, 在欧洲竞赛数据中是 ARIMA(1,2,0)×(0,1,1), 北美数据为 ARIMA(1,2,1)×(0,2,1).

1.4 多因素加法模型函数

多因素加法模型的主要思路是找到影响负荷的各个因素和负荷之间的映射关系, 然后把其映射关系相加, 得到最终所要找的模型. 多因素加法模型表现为一个曲线函数, 构造此曲线分为三个步骤: 第一步, 寻找出影响目标值的因素; 第二步, 拟合出这些因素的作用曲线函数; 第三步, 将这些函数相加得到最终模型.

从欧洲(EUNITE)竞赛数据和北美的电力数据研究发现影响负荷主要因素有温度、假期、星期和其他因素. 我们分别找出这些因素与负荷之间的映射关系, 拟合成自函数, 再将其相加得到多因素加法模型 $Y(x)$. 如式(1)所示, 其中 x 代表时间天数, $temperature$ 代表由温度因素拟合的函数, $holiday$ 代表由假期因素拟合的函数, $week$ 代表由星期因素拟合的函数, $other$ 则是代表由其他因素拟合的函数. $Y(x)$ 为多因素加法模型, $\epsilon(x)$ 为模型预测值与实际的误差.

$$Y(x)=temperature(x)+holiday(x)+week(x)+other(x)+\epsilon(x) \quad (1)$$

2 多因素加法模型实现

以(EUNITE)竞赛数据为例子, 从影响电力负荷的温度因素、假期因素、星期因素和其他因素出发, 分别介绍了这些因素与负荷之间的映射关系, 来解释多因素加法模型.

2.1 温度因素

图 1 中, 上部分的图是温度随着时间的变化曲线, 下半部分的图是每天每半小时最大用电量随着时间的变化的曲线. 从图 1 中可以看出温度与负荷有完全相反的的走势情况.

图 2 为温度与电力负荷的关系图, 线性趋势明显, 斜率为负数. 计算温度与电力负荷的 Pearson 相关系数, 值为-0.84, 因此平均温度与所用电量存在强烈的负相关.

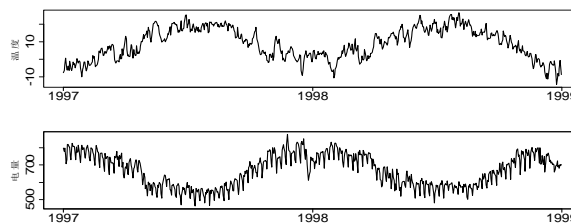


图 1 从 1997 年至 1998 年的温度和电力负荷

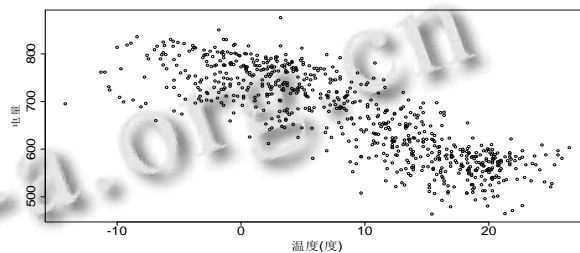


图 2 温度与电量的关系图

负荷随着季节变化而变化, 负荷数据和季节一样也具有 365 天的大周期, 夏季负荷较低, 而冬季负荷较高, 这主要由温度因素来影响负荷的, 因为冬季要使用取暖设备使得冬季有更大的用电量. 我们可以将温度拟合出一条三角函数, 来表示温度因素对负荷的影响.

2.2 星期因素

除了温度因素对负荷产生一定影响(365 天的大周期), 星期因素也对负荷造成一定的影响. 从图 1 电量图(下半部分)中可以看出, 除了 365 天的大周期以外, 温度和电量还存在有小的周期(褶皱部分). 仔细探究小周期, 我们可以发现其中规律. 图 3 表示 1997 年到 1998 年间, 星期一到星期天的平均负荷. 可以看出工作日负荷较高, 周末负荷较低, 而且周日的负荷通常要略低于周六的负荷. 通过时间序列周期的研究, 我们发现负荷数据存在 7 天的小周期, 这点与星期变化的因素吻合. 我们统计出 1997 年、1998 年这两年训练集中的周一到周日的平均电量与两年平均电量的差作为星期因素函数的权值, 加到 $Y(x)$ 函数中.

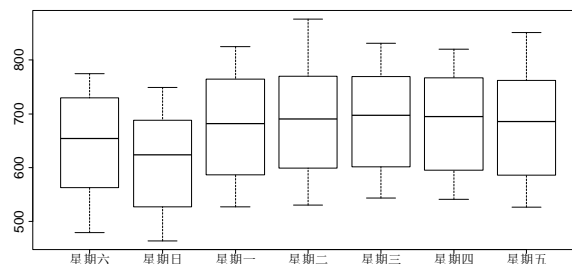


图 3 星期与电量的盒图

2.3 假期因素

我们观察到节假日的前后两天内皆会对负荷产生一定影响,只是影响程度不同.图4是1997年3月26日至4月4日的(EUNITE)竞赛数据,其中3月30日和3月31日为节假日.可以看出节假日的负荷要明显低于平日的负荷.为了找出节假日的前后两天内对负荷的具体影响,我们统计出节假日的前后两天内的平均负荷与平日的平均负荷的差,作为假期因素函数的权值,加到 $Y(x)$ 函数中.

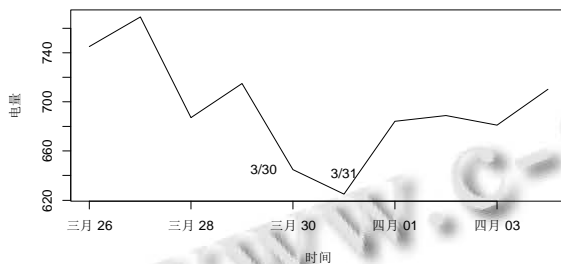


图4 假期与电量的关系图

2.4 其他因素

负荷还受其他因素的影响,进一步研究还发现,电量大小不仅看是否假期,还取决于假期的类型,不同的假期对负荷影响也不一样,圣诞节和元旦则会使负荷受较大影响.由于圣诞节和元旦这两个特殊节日的来临,电力数据会大大降低.另外,由图1温度图(上半部分)发现,1998年12月24日的温度达到了两年的最低,之后一个星期内的温度也持续走低,那么这将会对1999年1月的负荷预测也带来了一定的影响作用.由图1电量图(下半部分)可以看出在年末与年初部分,电量明显降低.因此1999年1月初负荷很有可能继续走低!因此我们在模型上加上年底至年初这段时间内的干扰,计算出这个时间段内的平均电量与两年平均电量的差作为其他因素函数的权值,作为 $other(x)$ 函数.

本文用的另一个数据集是北美电力数据,如图5所示.与(EUNITE)竞赛数据不同的是,北美数据年份较多(7年),可以看出有逐年增加的趋势,因此除了以上的因素以外,对北美数据还多增加了一个线性的函数.

综合以上的各个因素,根据公式(1),我们拟合出多因素加法模型的函数 $Y(x)$.例如,图6显示了 $Y(x)$ 函数模拟的EUNITE竞赛电力数据从1997年1月至1999年1月的电量负荷.

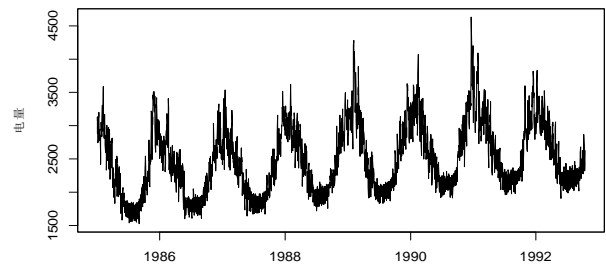


图5 北美电力数据的电量

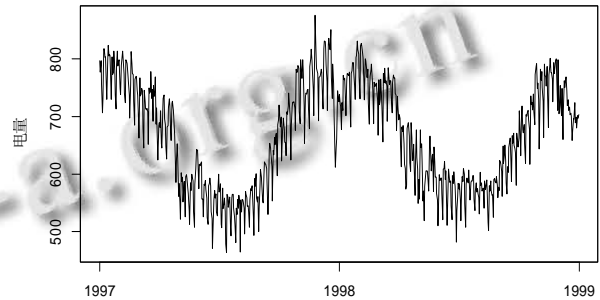


图6 MFA模型对(EUNITE)竞赛数据的拟合结果

3 评估指标

模型的预测性能是通过将目标变量的预测值与实际值比较得到的.本文使用六个评估指标,分别是平均绝对误差(MAE)、均方误差(MSE)、均方根误差(RMSE)、平均绝对百分误差(MAPE)、标准化后的均方误差(NMSE)和标准化后的平均绝对误差(NMAE).

从式(2)到式(7),是各种评估指标公式,其中, A_i 为目标变量实际值, F_i 为目标变量的预测值, N 为目标变量的个数.

$$MAE = \frac{1}{N} \sum_{i=1}^N |A_i - F_i| \quad (2)$$

均方误差(MSE)也是常用的误差度量,可以评价数据的变化程度:

$$MSE = \frac{1}{N} \sum_{i=1}^N (A_i - F_i)^2 \quad (3)$$

均方根误差(RMSE)是观测值与真值偏差的平方和观测次数 N 比值的平方根:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (A_i - F_i)^2} \quad (4)$$

平均绝对百分误差(MAPE)是衡量模型好坏最常用的一种方式:

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|A_i - F_i|}{A_i} \quad (5)$$

标准化后的均方误差(NMSE)和标准化后的平均绝对误差(NMAE),是计算模型预测性能和基准模型的预测性能之间的比率.在式(7)中 \bar{A} 为目标变量实际

值的平均值.

$$NMSE = \frac{\sum_{i=1}^N (A_i - F_i)^2}{\sum_{i=1}^N (A_i - \bar{A})^2} \quad (6)$$

$$NMAE = \frac{\sum_{i=1}^N |A_i - F_i|}{\sum_{i=1}^N |A_i - \bar{A}|} \quad (7)$$

4 实验结果与分析

我们运用前面介绍的 7 种常用算法和多因素加法模型(共 8 种不同模型), 分别应用于欧洲 EUNITE 竞赛数据和北美电力数据上. 其中欧洲 EUNITE 竞赛数据用 1997 年、1998 年的数据作为训练集, 预测 1999 年 1 月每天中每半小时最大电力负荷量. 北美电力数据用 1985 年到 1992 年 8 月的数据作为训练集, 预测出 1992 年 9 月 1 日到 1992 年 10 月 12 日期间每天中每半小时最大电力的负荷量.

表 1 和表 2 分别列出了 8 种不同模型在欧洲 (EUNITE)竞赛数据和北美电力数据的 6 个不同性能评估指标. 从上节介绍中可知, 这 6 种不同性能指标都是值越小性能越好. 表 1 和表 2 中, 性能最好的模型被标为粗体.

表 1 八种不同模型在欧洲(EUNITE)竞赛数据的预测性能

	MAE	MSE	RMSE	MAPE	NMSE	NMAE
MFA	11.8	230	15.18	0.0161	0.20	0.43
ARIMA	15.3	348	18.67	0.0205	0.05	0.20
LM	18.1	461	21.48	0.0244	0.06	0.22
RF	18.1	486	22.06	0.0245	0.06	0.22
SVM	18.8	565	23.79	0.0257	0.07	0.23
MARS	19.3	549	23.43	0.0263	0.07	0.24
NNET	20.3	683	26.15	0.0276	0.09	0.25
DT	23.5	742	27.25	0.0315	0.09	0.29

表 2 八种不同模型在北美电力数据的预测性能

	MAE	MSE	RMSE	MAPE	NMSE	NMAE
MFA	124.5	20308	142	0.051	0.0034	0.051
ARIMA	119.1	26038	161	0.047	0.64	0.72
NNET	136.5	26782	163	0.057	0.66	0.82
LM	168.9	42062	205	0.069	1.04	1.02
MARS	245.9	74213	272	0.1005	1.84	1.48
SVM	246.2	74651	273	0.099	1.85	1.48
RF	259.6	80443	283	0.106	1.99	1.57
DT	284.2	97966	313	0.116	2.43	1.71

我们可以看出, 两组数据的所有模型中, 使用 6 个评估指标, 多因素加法模型(MFA)和自回归积分滑动平均(ARIMA)是表现最好的两个模型: 其中, 在 8 种情况下, MFA 表现最好, 在 4 种情况下, ARIMA 表现最佳. 也就是说多因素加法模型(MFA)和自回归积分滑动平均(ARIMA)具有最准确的预测精度.

图 7 是多因素加法模型在欧洲(EUNITE)竞赛数据中对 1999 年 1 月的每日负荷预测结果, 实线部分为真实值, 虚线部分为多因素加法模型的预测值. 图 8 是多因素加法模型对北美电力数据 1992 年 9 月 1 日至 1992 年 10 月 12 日的每日负荷预测结果, 实线部分是真实值, 虚线部分是多因素加法模型的预测值.

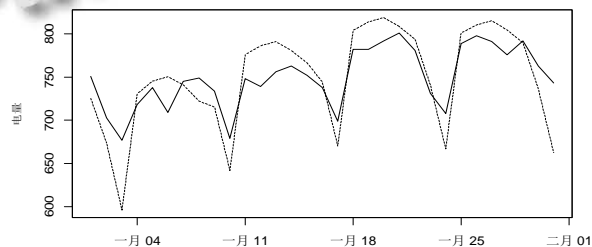


图 7 新模型预测值与 1999 年 1 月实际值比较

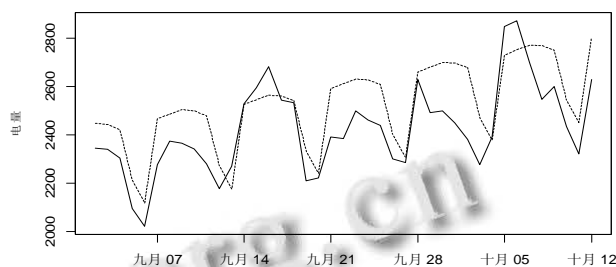


图 8 北美数据预测数据与实际数据比较

在欧洲 EUNITE 竞赛中, 荣获第一名组的 MAPE 值为 0.0195^[6]. MFA 模型的 MAPE 值为 0.0161, 比他们的性能好一些. 由于 MAPE 是业界常用的模型评价指标, 6 种指标虽然有所差异, 但是, 大体趋势一致, 因此, 以下我们就着重使用 MAPE. 图 9 为不同的模型在欧洲(EUNITE)竞赛数据和北美数据中的 MAPE 值的柱状图. 上半部分是欧洲竞赛数据, 下半部分为北美电力数据.

在欧洲电力数据中, MFA 的 MAPE 值(0.0161)最低(最优). 在北美电力数据中, ARIMA 的 MAPE 值(0.047)最低, MFA 的 MAPE 值(0.0517)比 ARIMA 稍高, 但是低于其他 6 种算法. 从图 9 中可以看出, 无论是 (EUNITE)竞赛数据还是北美电力数据, 多因素加法模

型(MFA)表现都不错。

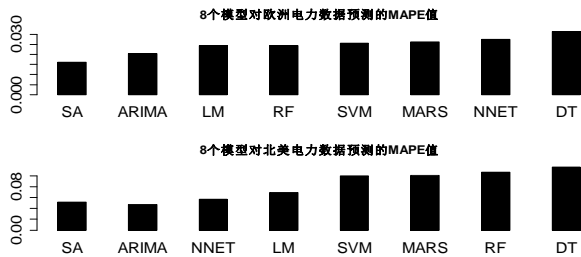


图9 不同模型对两个数据预测的MAPE值比较

除了预测性能,我们还统计了不同算法的运行时间.表3是不同模型分别在两个数据集上的运行时间.由于预测的天数比较少,因此多因素加法模型的运行时间几乎为零.从表3中,可以看出多因素加法模型在运行时间上也有优势.

表3 不同模型在不同数据上的运行时间(单位:秒)

模型	EUNITE 数据	北美数据
自回归积分滑动平均模型(ARIMA)	2.5	9.01
任意树模型(RF)	1.6	8.73
支持向量机模型(SVM)	0.31	4.17
神经网络模型(NNET)	0.38	0.39
多元自适应样条回归模型(MARS)	0.25	0.18
决策树模型(DT)	0.05	0.09
线性(LM)模型	0.02	0.09
多因素加法模型(MFA)	<1-e06	<1-e06

综上所述,这些实验结果说明了多因素加法模型(MFA)比其他常用算法在预测中期电力负荷中有着更精确的预测值与更高效的计算时间.

5 结语

准确预测每天的电力高峰负荷是一个重要的问题.在这篇文章中,我们使用了8种不同算法,应用两个实际的电力数据,欧洲EUNITE竞赛电力数据和北美电力数据,并且使用了6种不同评估指标来预测中期电力负荷.通过实验我们发现,跟其他7种不同模型相比,我们提出的多因素加法模型(MFA)在大多情况下(2个数据×6个不同评估指标=12种不同情况,其中有8次)有着更精确的预测性能,而且运行速度最快(计算性能高),并且至为重要的是该模型也比其他模型更加简单易懂.其中,ARIMA模型在12种不同情况下有4次变现也不错,也是其中最好的一种模型,但是其模型不容易理解且运行时间最长.

本文的局限性主要表现为数据的局限性.使用的数据为欧洲(EUNITE)竞赛电力数据和北美的电力数据,缺少中国本地数据.而且,欧洲EUNITE竞赛数据的年份也不够多,时间也不够新.因此未来应该收集更多更新的数据,来建立更加精确和通用的模型.

参考文献

- Elattar EE, Goulermas J, Wu HQ. Electric load forecasting based on locally weighted support vector regression. IEEE Trans. on Systems Man & Cybernetics Part C Applications & Reviews, 2010, 40(4): 438-447.
- 陈瑾然,徐文令.电力系统短期负荷预测方法的进展综述.电源技术应用,2013,(9):112.
- Rhynne R. Technological forecasting within alternative whole futures projections. Technological Forecasting & Social Change, 1974, 6(2): 133-162.
- 翟永杰,王静娴,周黎辉.基于模糊支持向量机的电力系统中期负荷预测.华北电力大学学报(自然科学版),2008,35(2): 70-73.
- 刘遵雄,张德运,孙钦东,徐征.基于相关向量机的电力负荷中期预测.西安交通大学学报,2004,38(10):1005-1008.
- Chen BJ, Chang MW, Lin CJ. Load forecasting using support vector machines: a study on EUNITE competition 2001. IEEE Trans. on Power Systems, 2004, 19(4): 1821-1830.
- 贺静,韦钢,熊玲玲.负荷预测线性回归分析法的模糊改进.华东电力,2003,32(11):21-23.
- Thomopoulos NT. Applied forecasting methods. Applied Forecasting Methods, 1980.
- George B, Gwilym J. Time series analysis, forecasting and control. Holden-day Series in Time Series Analysis, Revised Ed. San Francisco: Holden-day, 1970: 14.
- Peter J. Introduction to expert systems. Plant/Operations Progress, 1998, 9(1): 1-5.
- Deng JL. Control problems of grey systems. Systems & Control Letters, 1982, 1: 288-294.
- 李天云,刘自发.电力系统负荷的混沌特性及预测.中国电机工程学报,2000,20(11):36-40.
- Norman R. Wavelet contraction, wavelet expansion and the control of seismic resolution. Geophysics. 1953,18 (4): 769.
- McCulloch W, Walter P. A logical calculus of ideas immanent in nervous activity. Bulletin of Mathematical

- Biophysics, 1943, 5(4): 115–133.
- 15 Zadeh LA. Fuzzy sets. *Information and Control*, 1964, 8(65): 338–353.
- 16 Vladimir NV. *Statistical learning theory*. Wiley Interscience, 1998.
- 17 方瑞明. 支持向量机理论及其应用分析. 北京: 中国电力出版社, 2007.
- 18 Wu CH, Tzeng GH, Lin RH. A novel hybrid genetic algorithm for kernel function and parameter optimization in support vector regression. *Expert Systems with Applications*, 2009, 36(3): 4725–4735.
- 19 Nima A, Keynia F. Mid-term load forecasting of power systems by a new prediction method. *Energy Conversion and Management*, 2008, 49(10): 2678–2687.
- 20 Babinec S, Pospichal J. Optimization of echo state neural networks for electrical load forecasting. *Neural Network World*, 2007, 17(2): 133–152.
- 21 Mao H, Zeng X, Leng G, Zhai Y, Keane J. Short-term and midterm load forecasting using a bilevel optimization model. *IEEE Trans. on Power Systems*, 2009, 24(2): 1080–1090.
- 22 Company behind East-Slovakia Power Distribution Company. Worldwide competition within the EUNITE network, EUNITE competition report. <http://neuron-ai.tuke.sk/competition/>. 2001.
- 23 El-Sharkawi MA, Marks RJI, Oh S, Brace CM. Data partitioning for training a layered perceptron to forecast electric load. *Proc. of the Second International Forum on Applications of Neural Networks to Power Systems*. 1993, 4, 66–68.
- 24 R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing. <http://www.R-project.org/>. 2014.
- 25 R Core Team and Contributors Worldwide. Stats: the R stats package. R Foundation for Statistical Computing. <http://cran.r-project.org/>. 2014.
- 26 Stephen M. Earth: Multivariate adaptive regression splines. R Foundation for Statistical Computing. <http://CRAN.R-project.org/package=earth>. 2015.
- 27 Meyer D, Dimitriadou E, Hornik K, Leisch F, Weingessel A. E1071: Misc functions of the department of statistics (e1071), TU Wien. R Foundation for Statistical Computing. <http://CRAN.R-project.org/package=e1071>. 2014.
- 28 Andy L, Matthew W. Classification and regression by random forest. *R News*, 2002, 2(3): 18–22.
- 29 Terry T, Beth A, Brian R. Rpart: Recursive partitioning and regression trees. R Foundation for Statistical Computing. <http://CRAN.R-project.org/package=rpart>. 2015.
- 30 Brian R, William V. *Modern Applied Statistics with S*. Springer, 2010.