

DOI: 10.7667/PSPC201862

考虑关联因素的智能化中长期电力负荷预测方法

张志¹, 杜延菱¹, 崔慧军¹, 汪洋², 贺哲², 赖晓文³

(1. 国网冀北电力有限公司, 北京 100053; 2. 北京清能互联科技有限公司, 北京 100080; 3. 清华大学, 北京 100084)

摘要: 为了解决现有中长期负荷预测方法中存在的预测精度欠优及场景适用性较差等问题, 提出一种考虑影响因素协调关系及滞后效应的新型中长期负荷预测方法。首先通过关联矩阵筛选得到影响负荷变化的强相关因素, 结合计量学的 X-12-ARIMA 模型对负荷及其影响因素进行季节分解, 得到 3 个特征分解部分。在此基础上, 通过时滞效应检验确定滞后期数, 结合主成分分析法去除数据噪声影响, 进一步提升数据纯度。最后针对月度和季度负荷进行预测算例分析, 通过对比其他时序外推方法的预测结果和非线性模型方法应用后的精度提升, 验证了所提方法的有效性和适用性。

关键词: 中长期负荷预测; 季节分解; 时滞效应; 主成分分析

Intelligent mid-long electricity load forecast method considering associated factors

ZHANG Zhi¹, DU Yanling¹, CUI Huijun¹, WANG Yang², HE Zhe², LAI Xiaowen³

(1. State Grid Jibei Electric Power Co., Ltd, Beijing 100053, China; 2. Beijing Tsintergy Technology Co., Ltd, Beijing 100080, China; 3. Tsinghua University, Beijing 100084, China)

Abstract: Since available forecast models have a lack of validity and applicability, a new electric load forecast method, considering co-integration relation and hysteresis effect of influenced factors, is proposed. Firstly, by the use of incidence matrix, this paper obtains the close associated factors which affects load change. Secondly, combined with metrology model, X-12-ARIMA, the data of load and influence factors is decomposed into three different parts. Then, based on the seasonal decomposition, it takes lag steps about time lag effects testing and uses Principle Component Analysis (PCA) for de-noising to enhance data purity further. Finally, according to monthly and seasonal load forecasting, numerical examples are studied to verify effectiveness and applicability of proposed method through comparing with other method forecast results and non-linear model forecasting accuracy improvement.

Key words: mid-long term load forecasting; seasonal decomposition; time lag effects; principle component analysis

0 引言

中长期负荷预测主要针对月度及以上的时间尺度的负荷需求进行预测, 准确的中长期负荷预测对变电站选址定容、网架结构优化及调度计划的灵活制定等具有重要的指导意义, 可有效提高电网规划的经济性及保障电力市场的有序推进^[1]。近年来, 中国改革开放进入深水区, 国内经济结构正经历深刻调整, 负荷需求变化与经济社会发展关系愈发密切^[2]。同时, 随着信息化时代的到来, 数据统计精度、发布频度与信息化程度不断提升, 数据的类型、体量及统计口径都得到有效扩充, 很多细化的行业类、景气类指数不断推出, 为中长期负荷预测提供了充足数据来源与良好的数据基础。但经济社会、

行业结构、相关政策、气象条件等多方面因素之间的耦合关系复杂, 难以明确对中长期负荷需求的影响程度, 且预测的时间跨度长、不确定因素多, 因此中长期负荷预测是个复杂的、多维的、非线性的不确定性问题。

目前, 围绕中长期负荷需求预测这一主题已开展了大量的研究, 其主要思路是通过寻找全社会电力需求的时序发展规律及其与各种相关因素的历史相关性建立数学模型, 进而通过外推的方法实现预测。一般地, 这些分析预测方法都围绕着以下两个方面来展开: 一是电力数据的时序发展规律, 如回归分析、灰度预测、差分自回归移动平均 (Autoregressive Integrated Moving Average, ARIMA) 方法以及人工智能方法等^[3-5]; 二是电力需求与国民

经济发展的相关性, 如国内生产总值 (Gross Domestic Product, GDP) 综合电耗法、弹性系数法、经济周期分析法等^[6-7]。其中: 第一类方法从电力数据自身规律出发外延推测负荷发展趋势, 但预测精度依赖输入数据本身的精度, 且方法的物理意义不够明确, 难以对预测结果解析及分析误差^[8]; 第二类方法针对影响地区负荷变化的主要因素, 对地区年度负荷预测有较高精度, 但在季度、月度负荷预测方面, 方法考虑的影响因素相对较少, 难以反映不同时段电能需求特性差异, 也没有充分考虑到不同行业电能需求的结构化特征与发展趋势, 未能从更大数据区域中挖掘出关键的影响因素, 难以保证预测方法的鲁棒性^[9]。

事实上, 负荷需求发展可分为趋势部分、周期部分及其他部分。文献[10]以全社会月度用电量为预测目标, 通过季节分解的方法, 提取出不同的分量部分并结合多种经济类数据, 及分别采用针对性的模型进行预测, 所得预测电量精度较高。但全社会电量需求的发展趋势较为明确, 且宏观外部环境相对稳定, 影响因素变化相对较小。而影响地区最大负荷的相关因素类别更为繁多, 其中地区负荷与地区气象因素关系紧密。由于月度负荷预测历史样本规律遵循近大远小的规则^[1], 因此最近 3 年内的样本数据更具有参考价值, 但样本规模的减少噪声现象凸显会使得模型回归中出现过拟合现象^[12]。同时, 中长期预测的相关因素与负荷之间存在复杂的非线性关系, 前述文献多采用局部非线性模型进行拟合^[7], 对影响因素的非线性特性反映不足, 需应用多类型回归模型进行全方位描述。

综上所述, 本文在文献[10]构建的中长期负荷预测框架的基础上, 提出一种考虑关联大数据的智能化中长期电力负荷预测方法。首先, 方法针对难以甄别最大负荷主导影响因素的问题, 采用构建相关系数矩阵的方法, 筛选与最大负荷相关性较强的影响因素作为输入变量; 其次, 针对回归模型会出现过拟合现象的问题, 本文采取主成分分析法(PCA)对样本数据进行降维处理, 避免样本量与影响因素类型间矛盾带来的过拟合影响。最后, 针对局部非线性模型对影响因素非线性特征表达上的局限性, 本文应用多种非线性模型进行预测, 多方位比对分析本文方法的模型适用性及数据特征提取能力。

本文所提方法优点在于预测流程清晰, 对数据的要求相对宽松, 对模型的通用性较高, 方法应用范围广泛, 且能够有效辨析误差来源, 有助于优化预测精度和总结提升。

1 智能化中长期电力负荷预测方法流程

本文所构建的预测方法流程主要包括以下四个部分, “数据预处理及相关性分析—时间序列季节分解—分量滞后期数确定及 PCA 处理—综合预测及预测结果合成分析”, 整体流程图如图 1 所示。

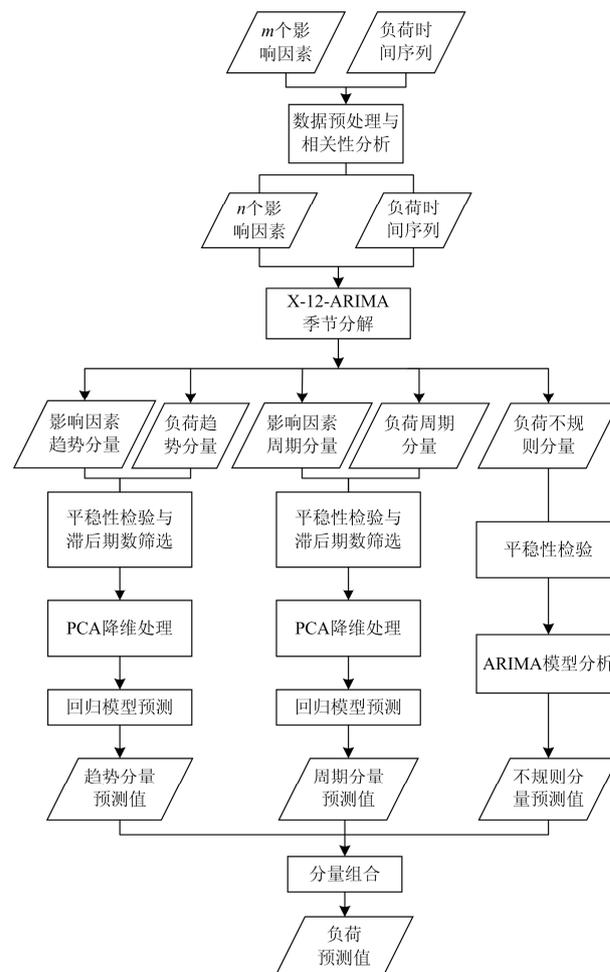


图 1 负荷预测流程图

Fig. 1 Flow chart of the forecast method

以下, 针对预测流程中的各个部分进行具体介绍。

1) 数据预处理及相关性分析

在多类型数据大量收集的基础上进行数据预处理, 针对数据集中的空缺值填充及长时间跨度合并, 使得各类影响因素数据在时间序列上呈现连续性; 同时, 对负荷与多种经济因素的历史数据进行相关性检验, 筛选出有强相关性数据; 并对预处理数据进行离群值检验, 修正错误数据, 为之后的负荷预测提供优质的数据来源。

2) 时间序列季节分解

使用 X-12-ARIMA 模型, 对经过预处理的负荷

数据与影响因素数据进行季节分解, 将各个时间序列分解为趋势分量、周期分量以及不规则分量。同时, 考虑到分量间的独立性, 故采用加法模型进行描述。

3) 分量滞后期数确定及数据降维处理

通过季节分解后的各分量数据, 与负荷的相应分量进行协整检验, 再结合最大信息准则确定负荷分量与各影响因素分量间的滞后期数; 确定数据滞后期数后, 通过 PCA 主成分分析法对数据进行降维转换, 滤除噪声成分。

4) 分量预测及预测结果合成分析

基于 PCA 转换后数据, 应用多类型的回归模型, 对负荷趋势分量和季节日分量分别进行预测, 同时, 由于不规则分量不应与外部影响因素相关, 因此应用自回归模型进行处理, 最后将各部分预测结果进行合成, 与原负荷数据进行比对分析。

2 数据预处理及相关性分析

2.1 影响因素收集及空缺填补

本文预测对象为月度及季度的区域最大负荷, 为了充分分析负荷与各种因素的关系, 本文选取了多类型相关影响因素, 主要分为地区经济类、社会产能类及地区气象类, 包括地区生产总值、社会消费品零售总额、电力电缆铺设长度、区域最高最低气温等 27 个影响因素, 其中, 某些影响因素的统计周期为季度统计, 例如地区生产总值, 因此需要填补该类数据的月度空缺值, 本文使用拉格朗日插值法进行处理, 计算公式如式(1)所示。

$$\begin{cases} p_k(x) = \prod_{i \in B_k} \frac{x - x_i}{x_k - x_i} \\ L_n(x) = \sum_{j=0}^{n-1} y_j p_j(x) \end{cases} \quad (1)$$

式中: $B_k = \{i = 0, 1, \dots, n-1 | i \neq k\}$; x 为空缺处相应期序号; x_i, x_k 为非空缺处的期序号; $p_j(x), p_k(x)$ 为非空缺处对空缺处的取值权重系数; 为期序号对应的数值, 针对季度数据, 取两相邻期序中值为中间月份填补值, 再结合式(1)进而求解得到其余两月度填补数据。

2.2 相关性检验

进一步, 通过计算影响因素与最大负荷间的相关系数, 筛选出相关系数大于 0.8 的影响因子作为本文负荷预测流程下一阶段的输入数据, 相关系数计算公式如式(2)所示。

$$\rho_{XY} = \frac{Cov(X, Y)}{\sigma_X \sigma_Y} \quad (2)$$

式中: $Cov(X, Y)$ 为 X 与 Y 序列的协方差; σ_X, σ_Y

为 X, Y 序列的对应标准差。

3 时间序列季节分解

针对经过相关性筛选的影响因素与负荷数据, 应用 X-12-ARIMA 工具进行季节分解, 其中 X-12-ARIMA 工具包含 regARIMA 和 X-11-ARIMA 两个部分。首先, regARIMA 对输入的数据进行预处理, 然后通过 X-11-ARIMA 进行多次基于移动平均的季节调整, 筛选并得到最终分解结果。

而 X-11-ARIMA 部分的核心则是基于多次迭代的移动平均方法, 文献[11]中对序列成分分解做出了详细描述。通过应用 X-12-ARIMA 工具, 将模型时间序列分解为长期趋势分量(trend component, T)、季节分量(season component, S)以及不规则分量(irregular component, X)。由于各个分量之间被认为是相互独立, 且本文为了尽量降低各个分量预测误差对总体预测结果的影响, 因此采取加法模型形式表示各个分量间关系:

$$Y_t = T_t + S_t + I_t \quad (3)$$

4 分量滞后期数确定及 PCA 处理

4.1 分量滞后期数确定

为了防止伪回归现象出现, 通过迪基-富勒判别测试(Argument Dickey-Fuller Test, ADF)进行单位根检验, 对不满足平稳性要求的影响因素再进入协整检验, 同时, 结合赤池信息准则与斯瓦茨准则, 计算其相应 AIC 值及 SC 值, 筛选最优滞后期数^[12], 其中, 滞后期数确定流程如图 2 所示。

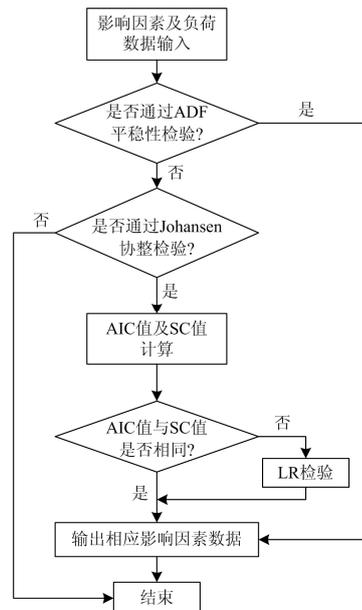


图 2 滞后期数确定流程图

Fig. 2 Flow chart of time lag period determination

4.2 数据降维处理

为了避免由于数据类型繁多及回归模型复杂引起的过拟合对预测精度的影响, 需要对各分量数据进行降维处理。应用主成分分析法(PCA)对通过协整检验的各因素进行处理, 首先, 选用 Z-score 对各类因素进行标准化处理, 然后构建协方差矩阵, 其次对协方差矩阵进行特征值求解, 采用累积方差百分比(Cumulative Percent Variance, CPV), 选择特征值贡献率超过 0.95 的影响因素确定降解维度数, 最后应用 PCA 降维算法把原影响因素矩阵重构出新的降维矩阵。

5 负荷分量预测模型选择

不同类型回归模型具有不同特点及适用范畴, 因此需要结合负荷特性进行模型甄选, 本文根据中长期负荷发展具有一定历史存续的特征, 针对稳定性较好的趋势分解部分和周期分解部分采用多元线性回归模型进行预测分析。

$$\tilde{f}_t = \omega_1 p_{1,t} + \dots + \omega_l p_{l,t} + \dots + \omega_k p_{k,t} + b_t = \mathbf{W}^T \tilde{\mathbf{P}}_t \quad (4)$$

式中: \tilde{f}_t 为预测负荷值; $p_{i,t}$ 为降维转换后数值; k 为转换后样本维度数; $\mathbf{W}_t^T = [\omega_1, \omega_2, \dots, \omega_k, b_t]$ 为增广系数向量; $\tilde{\mathbf{P}}_t = [p_{1,t}, p_{2,t}, \dots, p_{k,t}, 1]^T$ 则相应扩展向量。定义模型预测结果的损失函数为

$$e_t = \sum_{t=1}^m (f_t - \mathbf{W}^T \tilde{\mathbf{P}}_t)^2 = (\mathbf{F} - \mathbf{W}^T \tilde{\mathbf{P}})^T (\mathbf{F} - \mathbf{W}^T \tilde{\mathbf{P}}) \quad (5)$$

针对式(5), \mathbf{F} 为实际负荷向量, 再通过求解其损失函数的偏导数, 进而得到相应权系数向量 \mathbf{W}^T 。

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{W}^T} &= 2 \tilde{\mathbf{P}}^T (\mathbf{F} - \mathbf{W}^T \tilde{\mathbf{P}}) = 0 \\ \Rightarrow \mathbf{W}^* &= (\tilde{\mathbf{P}}^T \tilde{\mathbf{P}})^{-1} \tilde{\mathbf{P}}^T \mathbf{F} \end{aligned} \quad (6)$$

值得注意的是, 式(6)是参数 \mathbf{W} 最优解的闭式解, 只有在 $\tilde{\mathbf{P}}^T \tilde{\mathbf{P}}$ 为满秩矩阵或者正定矩阵时, 才使得式(6)计算成立。但在实际中, 往往不能满足应用要求, 这样的话会出现多解情况, 并且这多个解都能使均方误差最小化, 但不是所有解都适合做预测结果, 因为某些解可能会产生过拟合问题^[12]。

针对不规则分量, 考虑序列自身的内相关性, 采用时序自回归模型进行预测分析。同时, 为了展现本文方法的适用性, 应用多种非线性回归模型进行预测分析, 包括: 支持向量机(Support Vector Machine, SVM)回归模型^[13]、逻辑回归模型^[14]和随机森林模型^[15]等。

6 算例分析

6.1 基本数据

以 2013 年至 2017 年河北省的最大负荷月度/季度数据作为分析标的, 通过数据预处理及相关性分析筛选出有较强相关性的影响因素, 包括: 地区生产总值、电力电缆新增铺设、月度/季度发电量、固定资产投资、规模以上工业增加值、社会消费品零售总额、省内最大/最小平均气温等 15 种数据。

6.2 数据处理

对包括月度/季度最大负荷在内的 15 组相关因素数据进行季节分解, 得到各类影响因素的趋势分量 T_t 、季节分量 S_t 以及不规则分量 I_t 。以月度最大负荷数据为例, 可得到如图 3 的分解图。

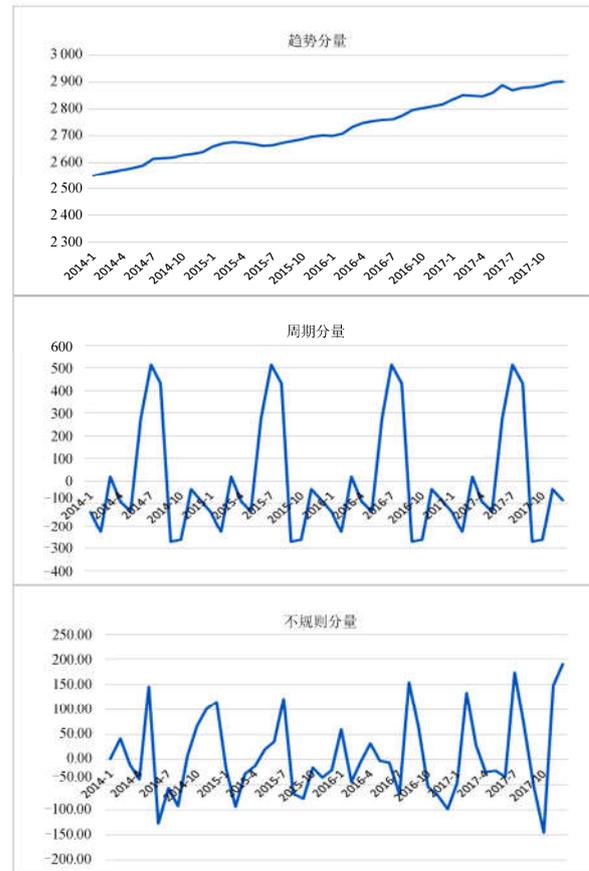


图 3 月度负荷分解结果

Fig. 3 Seasonal decomposition result of monthly peak load

针对趋势分量和季节分量分别进行协整检验以及滞后期数确定, 其中月度最大负荷的相关影响因素最大信息量滞后期数(以趋势分量为例)如表 1 所示, 而季度最大负荷的相关因素的滞后期数为 0, 即最大信息量为当期数据。

表 1 月度相关因素滞后期数表

Table 1 Time lag period of monthly associated factors

相关影响因素	滞后期数
地区生产总值	6
固定资产投资	2
规模以上工业增加值	0
社会消费品零售总额	2
软饮料产量/万吨	5
服装产量/万件	6
钢材产量/万吨	5
电力电缆新增铺设/km	2
太阳能电池新增容量/万 kW	6
最小平均气温/(°C)	0
最大平均气温/(°C)	0
月发电量/(GW·h)	0

确定最大信息滞后期数后, 针对模型变量类型较多引起的过度解析, 采用主成分分析法(PCA)做降维处理。首先, 计算影响因素的协方差矩阵, 然后通过求解矩阵的特征值, 保留特征值累计贡献率在 0.95 以上的对应特征向量, 并以此构成新的数据集作为模型输入。其中, 月度及季度的趋势分量和周期分量的前 5 个特征值贡献率如表 2 所示。

表 2 PCA 特征值贡献率表

Table 2 Eigenvalue ratio after PCA

序号	月度		季度	
	趋势部分	周期部分	趋势部分	周期部分
1	0.999 9	0.999 8	0.999 9	0.999 7
2	9.76e ⁻⁰⁶	0.000 16	5.28e ⁻⁰⁶	0.000 29
3	7.18e ⁻⁰⁹	3.66e ⁻⁰⁸	9.96e ⁻⁰⁷	1.37e ⁻⁰⁶
4	8.8e ⁻¹⁰	7.27e ⁻⁰⁹	6.64e ⁻⁰⁹	1.6e ⁻³⁴
5	2.956e ⁻¹⁰	3.3e ⁻⁰⁹	1.25e ⁻¹⁰	4.37e ⁻³⁷

从表 2 可以看出, 月度和季度各分量的最大特征值贡献率均超过 0.95 的阈值, 故原多维数据集可通过 PCA 方法转化为 1 维数据向量, 同时保留超过 99% 的原数据信息。进一步分析, 首特征值占比较大表征各类数据间存在紧密联系, 各影响因素在物理空间中变化趋势一致, 可通过降维处理来防止模型对影响因素的重复拟合。

6.3 最大负荷预测

本文选取 2017 年数据作为预测对象, 以 2013—2016 年数据为预测样本, 其中 2013 年数据作为备用数据, 应对滞后期数确定引起的样本数据偏移。对河北省 12 个月最大负荷进行虚拟预测, 并对各回归模型的预测结果进行误差分析, 以常用的平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)作为评判标准。

针对月度/季度最大负荷, 应用多元线性回归模

型进行虚拟预测, 所得结果如表 3 和表 4 所示。可以看出, 相较于常用的时序外推模型, 本文所提方法在预测精度上有很大提升。其中, 在月度最大负荷预测方面, 其他方法的最优精度为 7.81%, 而本文方法为 3%, 预测精度提升了 61.6%; 而在季度最大负荷预测方面, 其他方法的最优精度为 5.81%, 而本文方法为 1.95%, 预测精度大幅提升。

表 3 多种方法月度预测误差对比

Table 3 Monthly error comparison of several forecast methods

月份	灰色预测法/MW	ARIMA/MW	指数平滑法/MW	本文方法/MW
1	2 653.01	2 693.48	2 660.9	2 735.44
2	2 922.44	2 703.43	2 662.88	2 651.75
3	2 918.54	2 755.14	2 664.86	2 901.26
4	2 914.65	2 825.2	2 666.84	2 799.26
5	2 910.76	2 742.71	2 668.82	2 776.9
6	2 906.87	2 729.78	2 670.8	3 213.18
7	2 903	2 941.12	2 672.78	3 417.63
8	2 899.126	3 150.96	2 674.76	3 344.13
9	2 895.25	3 060.46	2 676.74	2 650.63
10	2 891.39	2 656.68	2 678.72	2 666.13
11	2 887.53	2 617.9	2 680.7	2 901.63
12	2 883.68	2 880.95	2 682.68	2 858.13
预测误差	8.19%	7.81%	9.16%	3.1%

表 4 多种方法季度预测误差对比

Table 4 Quarterly error comparison of several forecast methods

季度	灰色预测法/MW	ARIMA/MW	指数平滑法/MW	本文方法/MW
1	2 945.3	2 898.02	2 839.56	2 971.65
2	2 963.4	3 080.11	2 885.93	3 173.51
3	2 985.2	3 231.37	2 932.3	3 504.53
4	3 024.6	3 391.8	2 978.67	2 933.47
预测误差	5.86%	5.81%	7.08%	1.95%

6.4 本文模型预测效果分析

由上一节可知, 本文所提方法的预测精度相较于传统时序外推模型有大幅提升, 以下本文对影响因素的非线性特征的学习能力, 应用多个非线性模型进行虚拟预测, 观察分析其相应预测结果。

使用同一负荷数据集, 比对各非线性模型在应用本文方法前后的预测精度提升, 从而反映本文方法对非线性预测领域的适用性, 相关结果如表 5 所示。

表 5 各非线性模型预测结果

Table 5 Comparison of forecast result of non-linear models

模型类别	月度预测		季度预测	
	应用前精度	应用后精度	应用前精度	应用后精度
支持向量机	10.28%	3.2%	8.65%	2.82%
逻辑回归	11.6%	3.2%	10.1%	2.72%
随机森林	6.51%	3.25%	6.9%	3.56%

从表 5 可知, 各非线性模型的预测精度在应用本文方法后有较大提升, 表明本文方法能够有效提取数据的非线性特征, 应用领域较广。

进一步, 针对季节分解后的趋势分量和周期分量应用以上模型分别进行虚拟预测, 各分量预测精度如表 6 所示。其中, 针对不规则分量, 文献[10]采用白噪声近似处理, 但本文通过观察分析得到, 其与影响因素无显著关联, 且满足平稳性要求, 故采用 ARIMA 时序模型进行回归分析。

表 6 趋势分量及季节分量预测结果

Table 6 Predict result of trend component and seasonal component

模型类别	月度数据			季度数据		
	趋势分量	周期分量	不规则分量	趋势分量	周期分量	不规则分量
多元线性	0.75%	1.05%	10.1%	0.85%	0.97%	9.87%
支持向量机	1.85%	0.55%	10.1%	2.69%	0.37%	9.87%
逻辑回归	2.47%	0.55%	10.1%	2.7%	0.37%	9.87%
随机森林	2.1%	2.5%	10.1%	3%	4.12%	9.87%

综上, 由于趋势分量值远大于其余分量值, 因此趋势分量的预测精度对最终结果影响最大; 其中, 多元线性回归模型在趋势分量预测精度最好, 故其终合成结果的精度最高, 而随机森林模型对各分量的拟合效果最差, 因此其最终合成预测误差较大; 同时, 由于支持向量机、逻辑回归及随机森林对训练样本量比较敏感, 因此样本量相对较大的月度分量预测时, 较于季度分量预测时会表现较好。对于不规则分量, 其预测误差较大, 月度、季度分别达到 10.1%和 9.87%, 由于不规则分量的波动特征明显, 导致预测值难以反映尖峰值, 故在尖峰值的位置产生较大误差, 但不规则分量数值相对较小, 因此对最后的负荷预测精度影响有限。

7 结论

本文提出了一种新型的中长期负荷预测方法, 涵盖了负荷预测框架及数据处理流程, 通过 X-12-ARIMA 季节分解工具将相关因素分为相互独立的趋势部分、季节周期部分以及不规则部分, 针对各个分量进行特征处理, 并应用相契合模型使得负荷预测精度得到显著提升。本文方法重点考虑了关联因素的时滞效应以及过拟合现象, 通过滞后期数确定和 PCA 降维处理, 有效凸显了数据的历史存续效应以及去除数据样本内的噪声影响。最后的算例结果表明, 本文方法相较于传统时序外推方法有更好的预测表现, 且方法对模型的适用性较好, 能够优化模型对样本非线性特征的学习能力。同时, 从方法预测结果分析可知, 本文方法能够分辨误差来源,

起到指引降低误差的效果。

参考文献

- [1] 康重庆, 夏清, 刘梅. 电力系统负荷预测[M]. 北京: 中国电力出版社, 2007: 2-4.
- [2] LI W Y. Framework of probabilistic power system planning[J]. CSEE Journal of Power and Energy Systems, 2015, 1(1): 1-8.
- [3] PENG C Y, LEI S B, HOU Y H, et al. Uncertainty management in power system operation[J]. CSEE Journal of Power and Energy Systems, 2015, 1(1): 28-35.
- [4] MISHRA A, KUMAR G V N. Congestion management of power system with interline power flow controller using disparity line utilization factor and multi-objective differential evolution[J]. CSEE Journal of Power and Energy Systems, 2015, 1(3): 76-85.
- [5] 李钜, 李敏, 刘涤尘. 基于改进回归法的电力负荷预测[J]. 电网技术, 2006, 30(1): 99-104.
LI Po, LI Min, LIU Dichen. Power load forecasting based on improved regression[J]. Power System Technology, 2006, 30(1): 99-104.
- [6] CETINKAYA N, CETINKAYA N. Long-term electrical load forecasting based on economic and demographic data for Turkey[C] // IEEE 14th International Symposium on Computational Intelligence and Informatics (CINTI), 2013, Budapest : 219-223.
- [7] SHAO Z, GAO F, ZHANG Q, et al. Multivariate statistical and similarity measure based semiparametric modeling of the probability distribution: a novel approach to the case study of mid-long term electricity consumption forecasting in China[J]. Applied Energy, 2015(156): 502-518.
- [8] KAVAKLIOGLU K. Modeling and prediction of Turkey's electricity consumption using support vector regression[J]. Applied Energy, 2011, 88(1): 368-375.
- [9] HU Z, BAO Y, CHIONG R, et al. Mid-term interval load forecasting using multi-output support vector regression with a memetic algorithm for feature selection[J]. Energy, 2015(84): 419-431.
- [10] 郭鸿业, 陈启鑫, 夏清, 等. 考虑经济因素时滞效应的月度负荷预测方法[J]. 电网技术, 2016, 40(2): 514-520.
GUO Hongye, CHEN Qixin, XIA Qing, et al. Study on mid-term electricity load forecast considering time lag effects of economic factors[J]. Power System Technology, 2016, 40(2): 514-520.
- [11] 中国人民银行调查统计司. 时间序列 X-12-ARIMA 季节调整——原理与方法[M]. 北京: 中国金融出版社, 2006: 121-144.
- [12] WOOLDRIDGE, JEFFREY. Introductory econometrics:

- a modern approach[M]. Cengage Learning, 2012: 105-133.
- [13] 王宁, 谢敏, 邓佳梁, 等. 基于支持向量机回归组合模型的中长期降温负荷预测[J]. 电力系统保护与控制, 2016, 44(3): 92-97.
WANG Ning, XIE Min, DENG Jialiang, et al. Mid-long term temperature-lowering load forecasting based on combination of support vector machine and multiple regression[J]. Power System Protection and Control, 2016, 44(3): 92-97.
- [14] 毛毅, 陈稳霖, 郭宝龙, 等. 基于密度估计的逻辑回归模型[J]. 自动化学报, 2014, 40(1): 62-72.
MAO Yi, CHEN Wenlin, GUO Baolong, et al. A novel logistic regression model based on density estimation[J]. Acta Automatica Sinica, 2014, 40(1): 62-72.
- [15] 张雷, 王琳琳, 张旭东, 等. 随机森林算法基本思想及其在生态学中的应用——以云南松分布模拟为例[J]. 生态学报, 2014, 34(3): 650-659.
ZHANG Lei, WANG Linlin, ZHANG Xudong, et al. The basic principle of random forest and its applications in ecology: a case study of Pinus Yunnanensis[J]. Acta Ecologica Sinica, 2014, 34(3): 650-659.
- [16] DE FELICE M, ALESSANDRI A, CATALANO F. Seasonal climate forecasts for medium-term electricity demand forecasting[J]. Applied Energy, 2015(137): 435-444.
- [17] 代亮, 许宏科, 陈婷, 等. 基于MapReduce的多元线性回归预测模型[J]. 计算机应用, 2014, 34(7): 1862-1866.
DAI Liang, XU Hongke, CHEN Ting, et al. Multivariate linear regression forecasting model based on MapReduce[J]. Journal of Computer Applications, 2014, 34(7): 1862-1866.
- [18] 张栋梁, 严健, 李晓波, 等. 基于马尔可夫链筛选组合预测模型的中长期负荷预测方法[J]. 电力系统保护与控制, 2016, 44(12): 63-67.
ZHANG Dongliang, YAN Jian, LI Xiaobo, et al. Mid-long term load forecasting based on Markov chain screening combination forecasting models[J]. Power System Protection and Control, 2016, 44(12): 63-67.
- [19] 陈毅波, 郑玲, 姚建刚. 基于粗糙集理论与 D-S 证据理论改进的多元回归负荷预测方法研究[J]. 电力系统保护与控制, 2016, 44(12): 63-67.
CHEN Yibo, ZHENG Ling, YAO Jianguang. Improved multiple regression load forecasting method based on rough set theory and D-S evidence theory[J]. Power System Protection and Control, 2016, 44(12): 63-67.
- [20] 刘云, 张杭, 张爱民. 需求侧响应下基于负荷特性的改进短期负荷预测方法[J]. 电力系统保护与控制, 2018, 46(13): 126-133.
LIU Yun, ZHANG Hang, ZHANG Aimin. Improved load forecasting method based on load characteristics under demand-side response[J]. Power System Protection and Control, 2018, 46(13): 126-133.
- [21] 方八零, 李龙, 赵家铸, 等. 动态相似与静态相似相结合的短期负荷预测方法[J]. 电力系统保护与控制, 2018, 46(15): 29-35.
FANG Baling, LI Long, ZHAO Jiatao, et al. Short-term load forecasting based on the combination of dynamic similarity and static similarity[J]. Power System Protection and Control, 2018, 46(15): 29-35.
- [22] 陈丽娜, 张智晟, 于道林. 基于广义需求侧资源聚合的电力系统短期负荷预测模型[J]. 电力系统保护与控制, 2018, 46(15): 45-51.
CHEN Lina, ZHANG Zhisheng, YU Daolin. Short-term load forecasting model of power system based on generalized demand side resources aggregation[J]. Power System Protection and Control, 2018, 46(15): 45-51.
- [23] KIRANMAI S A, LAXMI A J. Data mining for classification of power quality problems using WEKA and the effect of attributes on classification accuracy[J]. Protection and Control of Modern Power Systems, 2018, 3(3): 303-314. DOI: 10.1186/s41601-018-0103-3.

收稿日期: 2018-06-21; 修回日期: 2018-10-15

作者简介:

杜延菱(1969—), 女, 通信作者, 硕士, 高级工程师, 研究方向为电力系统调度运行。E-mail: 498315041@qq.com

(编辑 张爱琴)