

基于 Copula 理论与 KPCA-GRNN 结合的区域综合能源系统多元负荷短期预测模型

马建鹏¹, 龚文杰², 张智晟¹

(1. 青岛大学电气工程学院, 山东 青岛 266071; 2. 国网青岛供电公司, 山东 青岛 266002)

摘要: 通过 Copula 理论对多元负荷之间以及多元负荷与天气因素之间的非线性相关性进行分析。基于分析结果构建模型样本集,通过核主成分分析(KPCA)对样本集进行降维解耦,采用广义回归神经网络(GRNN)模型对多元负荷进行预测,模型参数采用遗传算法进行优化。以北方某建筑级区域综合能源系统夏季多元负荷中的电负荷和冷负荷算例为例,与其他两种预测模型进行对比,验证结果表明,基于 Copula 理论与 KPCA-GRNN 结合的区域综合能源系统多元负荷短期预测模型具有较高的预测精度,该模型具有一定的应用价值。

关键词: 区域综合能源系统; 多元负荷预测; Copula 理论; 核主成分分析; 广义回归神经网络
DOI: 10. 12067/ATEEE1903001 **文章编号:** 1003-3076(2020)03-0024-08 **中图分类号:** TM743

1 引言

区域综合能源系统是一种能够同时提供人们所需的多种负荷的能源系统,由于其实现了能源的梯级利用,提高了一次能源的利用率,因此在办公以及商业建筑中得到了广泛应用。由于准确的多元负荷短期预测是区域综合能源系统运行以及调控的关键,因此其成为了国内外众多专家学者的重要研究对象。以数学理论为基础的预测方法有自适应线性时间序列法^[1]、非线性自回归法^[2],以人工智能为基础的预测法有人工神经网络^[3]、支持向量机^[4]等。上述方法皆获得了较高的预测准确度,并且具有一定的实用价值,但在预测过程中大都使用复杂的算法,且多元负荷单独预测,忽略了多元负荷之间的关联性。多元负荷时间序列具有较强的动态性以及非线性特征,多元负荷之间、负荷与影响因素之间亦存在较强耦合性。基于区域综合能源系统多元负荷的特性,本文提出了基于 Copula 理论与核主成分分析-广义回归神经网络(Kernel Principal Component Analysis-Generalized Regression Neural Network,

KPCA-GRNN)结合的区域综合能源系统多元负荷短期预测模型。首先,通过 Copula 理论对多元负荷之间以及负荷与天气因素之间的非线性耦合进行分析,使用 Spearman 相关系数量化变量之间存在的非线性耦合,由此选择合适的影响因子构造原始数据集。然后,采用 KPCA 通过核函数映射,实现对输入数据样本集中非线性耦合变量的降维解耦,克服了传统降维方式需要嵌入高维数据空间中的子空间是线性的或者近似线性的不足^[5],有效降低了输入数据样本集的冗余度。由径向基函数神经元和线性神经元组合而成的 GRNN 的网络结构具有鲁棒性,使它能够更好地对动态不确定性和强随机性的非线性时间序列进行预测^[6]。采用遗传算法优化 GRNN 神经网络的参数,有效保证了模型的泛化能力。以北方某建筑级区域综合能源系统夏季多元负荷中的电负荷和冷负荷算例为例,与其他两种预测模型进行对比,验证结果表明,基于 Copula 理论与 KPCA-GRNN 结合的区域综合能源系统多元负荷短期预测模型具有较高的预测精度,该模型具有较好的应用价值。

收稿日期: 2019-03-01
基金项目: 国家自然科学基金项目(51477078)
作者简介: 马建鹏(1994-),男,山东籍,硕士研究生,研究方向为电力系统短期负荷预测;
张智晟(1975-),男,山东籍,教授,博士后,研究方向为电力系统短期负荷预测(通讯作者)。

2 多元负荷及天气因素相关性分析

综合能源系统多元负荷由于包含气象敏感负荷, 受天气因素的影响, 故具有较强的随机性以及不确定性。为了分析多元负荷之间及天气因素对多元负荷的影响机理, 充分挖掘多变量数据中的相关信息, 有效选择多元负荷预测模型的输入变量, 必须对多元负荷之间的耦合性以及天气因素对负荷的影响进行量化分析。以电负荷为例, 图 1 为多元负荷及天气因素散点图, 由图 1 可知, 电负荷与冷负荷及各影响因素之间存在明显的非线性耦合。传统相关性分析大都采用 Pearson 法, 但 Pearson 法着重分析的是变量间的线性相关, 不能描述变量间的非线性相关^[7]。Copula 理论作为新兴的相关性分析方法, 由于其能够精确地捕捉到变量间的非线性特征, 且通过 Copula 函数导出的相关性测度具有应用范围更广、实用性更强的优点, 故在各领域获得了广泛应用^[8]。因此基于多元负荷之间以及多元负荷与影响因素之间的非线性耦合, 选择 Copula 理论对上述耦合性进行量化分析。

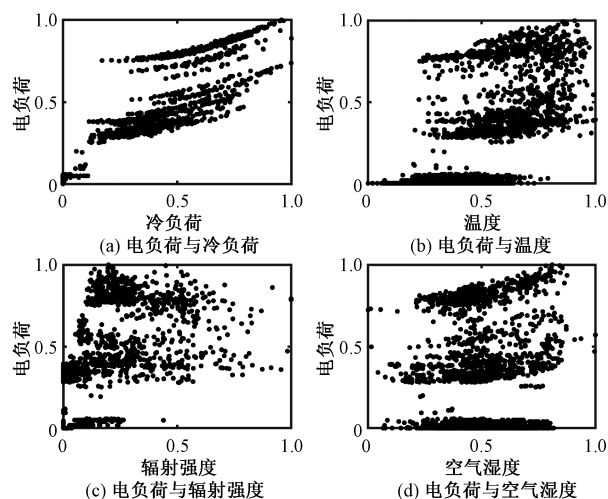


图 1 多元负荷及各影响因素散点图

Fig.1 Scatter diagram of multiple loads and various influencing factors

2.1 Copula 理论

Copula 函数作为标准均匀随机变量的连接函数, 是一种能够将边缘分布与联合分布有效联系起来的数学工具^[9], 根据 Sklar 定理, 对于随机变量 x_1, x_2, \dots, x_N , 在已知变量的边缘分布函数 $F_1(x_1), F_2(x_2), \dots, F_N(x_N)$ 的情况下, 存在一个 Copula 函数 $C(\cdot)$ 满足:

$$F(x_1, x_2, \dots, x_N) = C(F_1(x_1), F_2(x_2), \dots, F_N(x_N)) \quad (1)$$

式中, $F(x_1, x_2, \dots, x_N)$ 为随机变量 x_1, x_2, \dots, x_N 的联合分布。

依据 Sklar 定理可以通过 Copula 函数与变量的边缘分布来构建多元变量的联合分布, 将求解多元变量的联合分布问题分解为求各单变量的边缘分布及相应 Copula 函数的问题, 可有效降低问题的复杂性, 简化计算量。通过对 Copula 函数构建的联合分布函数进行求导计算, 便可以得到联合分布函数的密度函数, 通过联合分布密度函数便可以从概率的角度精确且完整地描述变量之间的相关结构。

2.2 基于 Copula 理论的相关性分析

联合密度函数虽然能够精确地刻画变量间的相关结构, 但密度函数表述较为复杂, 缺乏一定的直观性。为了更加直观地比较分析多元负荷之间以及负荷与天气因素之间的整体相关性, 选择由 Copula 函数导出的相关性测度中可以描述变量间非线性相关的 Spearman 秩相关系数 ρ 作为非线性相关的最终评价指标。

对于两个随机变量 X 和 Y , 二者对应的分布函数为 $G(x)$ 和 $H(y)$, 则一定存在一个 Copula 函数 $C(a, b)$ 使得其可以通过 Spearman 秩相关系数 ρ 来表征 X 和 Y 之间的非线性相关, 也即可以利用 Copula 函数 $C(a, b)$ 来导出 ρ , 即:

$$\begin{aligned} \rho &= 12 \int_0^1 \int_0^1 abdC(a, b) - 3 \\ &= 12 \int_0^1 \int_0^1 C(a, b) da db - 3 \end{aligned} \quad (2)$$

式中, a 和 b 分别为两个变量 X 和 Y 的边缘分布函数 $G(x)$ 和 $H(y)$ 。

常用的 Copula 函数共 5 种, 分别为 t-Copula、Gumbel-Copula、Clayton-Copula、Frank-Copula 和 Normal-Copula^[10]。由同一 Copula 函数导出的 Spearman 秩相关系数是固定且唯一的, 但是由于 Copula 函数种类繁多, 不同 Copula 函数导出的相关系数值存在一定差异, 因此基于 Copula 函数的相关性分析的准确度依赖于最优 Copula 函数的选择。

本文通过定义经验 Copula 函数 $\hat{C}(a, b)$, 并比较各备选 Copula 函数与经验 Copula 函数之间的欧氏距离来实现 Copula 函数的优选。

定义经验 Copula 函数 $\hat{C}(a, b)$ 如下^[11]: 设 x_i, y_i ($i = 1, 2, \dots, n$) 分别为取自变量 X 和 Y 的样本,

$F_x(x)$ 、 $F_y(y)$ 为 X 和 Y 的经验分布函数,则有:

$$\hat{C}(a,b)=\frac{1}{n}\sum_{i=1}^nI_{[F_x(x_i)\leqslant a]}I_{[F_y(y_i)\leqslant b]}\quad(3)$$

式中, $I_{(if)}$ 为示性函数,当条件 if 成立时, $I_{(if)}=1$,否则为 0。

经验 Copula 函数与各备选 Copula 函数之间的欧氏距离公式如下:

$$d_j=\sum_{i=1}^n|\hat{C}(a_i,b_i)-C_j(a_i,b_i)|\quad(4)$$

式中, a_i 为 x_i 对应的经验分布函数值 $F_x(x_i)$; b_i 为 y_i 对应的经验分布函数值 $F_y(y_i)$; $C_j(\cdot)$ 为备选 Copula 函数。

使用 Copula 理论进行相关性分析的步骤如下:首先采用核密度估计法求得变量的边缘分布,根据边缘分布使用最大似然法计算各备选 Copula 函数的未知参数;根据样本点数据,计算获得变量的经验分布函数,然后采用样条插值法获得每一个样本点对应的经验分布函数值,根据式(3)和备选 Copula 函数,由经验分布函数值计算其对应的经验 Copula 函数值与备选 Copula 函数值;根据式(4),计算各备选 Copula 函数与经验 Copula 函数之间的欧氏距离,选择欧氏距离最小者为最优 Copula 函数;根据式(2),通过最优 Copula 函数进行相应的积分计算便可导出 Spearman 相关系数。

根据上述方法,对夏季(6~8月)电负荷、冷负荷以及天气因素之间的非线性相关性进行分析,样本数据共 2208 组,样本数达到 Copula 函数的计算要求,样本数据能够较准确地反映变量的真实分布规律, E 代表电负荷, C 代表冷负荷, T 代表温度, R 代表辐射强度, M 代表空气湿度。

Spearman 秩相关系数可以由 Frank-Copula、t-Copula、Normal-Copula 三种备选 Copula 函数导出,最优 Copula 函数的选择如表 1 所示。由表 1 可知,三种备选 Copula 函数与经验 Copula 函数之间的欧氏距离不同,欧式距离越小,备选 Copula 函数越接近经验 Copula 函数,由此导出的相关系数越能准确描述变量之间的相关性。

由最优 Copula 函数导出的相关性分析结果如表 2 所示。由表 2 可知,电负荷、冷负荷之间以及多元负荷与温度、辐射强度、空气湿度之间存在高度的非线性相关,但相关程度存在差异。电负荷与冷负荷之间的非线性相关高达 0.86,同时电负荷、冷负荷与温度、辐射强度以及空气湿度之间的非线性相

表 1 欧氏距离计算结果

Tab.1 Euclidean distance calculation result

	Normal-Copula	t-Copula	Frank-Copula
E 与 C	0.399	0.216	0.187
E 与 T	0.615	0.477	0.146
E 与 R	0.927	0.610	0.205
E 与 M	0.460	0.172	0.474
C 与 T	1.153	0.095	0.114
C 与 R	0.756	0.184	0.335
C 与 M	0.161	0.298	0.286

关均大于 0.1。电负荷与冷负荷之间的耦合特性分析结果表明了区域综合能源系统多元负荷之间的系统性以及预测过程中的不可分割性,故为了提高多元负荷的预测精度,二者应综合建模预测。电负荷、冷负荷与天气因素之间的耦合特性分析结果表明,在多元负荷预测中上述影响因素对预测结果的影响是不可忽略的,为预测数据集的构造提供了理论依据。因此将温度、辐射强度、湿度作为预测模型的影响因子,与电负荷、冷负荷历史数据共同构成多元负荷预测模型的输入样本集。

表 2 相关性分析结果

Tab.2 Correlation analysis results

	E	C	T	R	M
E	1	0.86	0.80	0.65	0.19
C	0.86	1	0.83	0.76	0.32

3 多元负荷短期预测模型

为了保证预测模型准确度,与普通单负荷短期预测相比,多元负荷短期预测需要更多的数据集。而随着数据集维度的增加,将会导致预测模型输入空间维度过高且模型结构复杂,影响模型的预测性能。因此需要在保证有效信息的前提下对输入空间进行合理的降维处理,提高模型的学习能力和泛化能力^[12]。而通过第 2 节的分析可知,多元负荷之间以及多元负荷与天气因素之间存在非线性耦合,故在此基础上,摒弃基于线性降维的主成分分析(PCA),选择基于变量之间非线性耦合的降维方式 KPCA。

3.1 核主成分分析

核主成分分析的基本原理就是通过非线性函数 Φ 对原始数据 $[d_1,d_2,\cdots,d_q]$ 进行映射,降低数据的非线性,得到更高维空间中的对应数据 $[\Phi(d_1),\Phi(d_2),\cdots,\Phi(d_q)]$,然后对高维空间的数据进行

PCA^[13]。即求解高维空间中样本协方差矩阵的特征值以及特征向量:

$$\lambda_{\phi_i} \mu_{\phi_i} = C_{\phi} \mu_{\phi_i} \quad i = 1, 2, \dots, q \quad (5)$$

式中, λ_{ϕ_i} 为特征值; μ_{ϕ_i} 为对应特征向量; C_{ϕ} 为高维空间中样本协方差矩阵, 如式(6)所示:

$$C_{\phi} = \frac{1}{n} \sum_{i=1}^n [\Phi(d_i) \Phi(d_i)^T] \quad (6)$$

根据 Mercer 定理, 存在核矩阵 K 可将式(5)转化为式(7)求特征值的问题:

$$\lambda_{k_i} \mu_{k_i} = K \mu_{k_i} \quad i = 1, 2, \dots, q \quad (7)$$

式中, $k_{ij} = K(d_i, d_j) = \Phi(d_i) \Phi(d_j)^T$, $K = (k_{ij})_{q \times q}$ 。通过计算得到特征值以及特征向量 λ_{k_i} 和 μ_{k_i} , 对 μ_{k_i} 进行正交规范化处理得到 μ_i , μ_i 与对应核矩阵向量 k_i 的乘积即为核主成分 x_i , 最后根据对应累计贡献率大小, 选择前 q 个核主成分 $[x_1, x_2, \dots, x_n]$ ($n \leq q$) 作为预测模型的输入向量。

本文采用高斯核函数计算核矩阵, 高斯核函数表达式如下:

$$k_{ij} = \exp\left(-\frac{\|x_i - x_j\|^2}{2\delta^2}\right) \quad (8)$$

式中, δ 为高斯核函数的带宽。

3.2 KPCA-GRNN 预测模型

多元负荷作为典型的混沌时间序列, 具有明显的非线性变化特征, 而 GRNN 神经网络作为径向基函数神经网络 (Radial Basis Function Neural Network, RBFNN) 的变种, 具有高精度非线性函数的逼近能力, 因此选择 GRNN 神经网络实现对于多元负荷的短期预测。GRNN 神经网络由径向基神经元以及线性输出神经元组成, 网络具有较高的鲁棒性和容错性, 因此获得了广泛应用并在负荷预测中取得了较高的准确度^[14,15]。GRNN 神经网络由 4 层神经元构成, 分别是输入层、模式层、求和层、输出层^[16]。KPCA-GRNN 模型结构如图 2 所示。

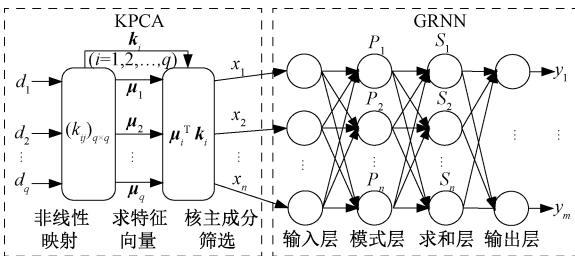


图 2 KPCA-GRNN 模型结构图

Fig.2 Structure diagram of KPCA-GRNN model

$x_2, \dots, x_n]$ ($n \leq q$) 作为 GRNN 神经网络的输入。GRNN 网络输入层与模式层的神经元个数与输入样本的维度相同, 输入层只承接输入数据, 不对数据进行处理。模式层使用高斯函数作为传递函数对输入层承接的数据进行处理, 模式层数据输出如下:

$$P_i = \exp\left[-\frac{(x - x_i)^T(x - x_i)}{2\sigma^2}\right] \quad i = 1, 2, \dots, n \quad (9)$$

式中, σ 为 GRNN 神经网络的光滑因子。

求和层神经元数与模式层神经元数相同, 该层共有两种传递函数, 分别对模式层的输出进行处理。一种是连接权值同为 1 的求和函数 S_D :

$$S_D = \sum_{i=1}^n P_i \quad (10)$$

另一种是权值为第 i 个输出样本 y_i 中的第 j 个元素 y_{ij} 的求和函数 S_{N_j} ^[17]:

$$S_{N_j} = \sum_{i=1}^n y_{ij} P_i \quad j = 1, 2, \dots, m \quad (11)$$

输出层神经元个数等于输出数据的维度, 其传递函数为比例函数:

$$y_j = \frac{S_{N_j}}{S_D} \quad j = 1, 2, \dots, m \quad (12)$$

光滑因子作为 GRNN 神经网络的唯一可调节结构参数, 能够直接影响神经网络的预测精度, 本文采用遗传算法对光滑因子进行优选。

KPCA-GRNN 模型将核主成分分析与广义回归神经网络有机结合, 通过 KPCA 有效降低了输入样本的维度, 简化了预测模型的结构, 并且 GRNN 神经网络相比于其他预测方法能够更好地预测非线性时间序列。因此, KPCA-GRNN 模型能够较好地实现综合能源系统多元负荷预测。

4 算例分析

4.1 数据预处理

本文以为北方某写字楼进行供能的区域综合能源系统为例, 采用 DeST-C 软件对写字楼进行模拟, 得到的负荷数据以及天气因素数据作为多元负荷短期预测所需数据集。为验证本文提出的基于 Copula 理论与 KPCA-GRNN 结合的多元负荷短期预测模型的有效性, 对北方某建筑级综合能源系统的夏季电负荷和冷负荷进行预测仿真。预测模型的输入矩阵由电负荷、冷负荷历史数据以及温度、辐射强度、空气湿度组成, 根据节假日与否将样本划分为工作日

经过 KPCA 降维解耦处理后的样本数据 $[x_1,$

与休息日两部分,并对各变量进行归一化处理,本文自定义了一种按序列举的方式对输入矩阵进行描述,举例如下:

矩阵 $\mathbf{M} = [\mathfrak{S}_{i=1}^2(a_i, b_i)] = [a_1, b_1, a_2, b_2]$, 表示矩阵 \mathbf{M} 由变量 a, b 在第 $i (i=1, 2)$ 天的值组成,其中 \mathfrak{S} 为本文自定义的按序列举方式。

$\mathbf{N} = [\mathfrak{S}_{i=1}^2(\mathfrak{S}_{j=1}^3(a_i^j))] = [a_1^1, a_1^2, a_1^3, a_2^1, a_2^2, a_2^3]$, 表示矩阵 \mathbf{N} 由变量 a 在第 $i (i=1, 2)$ 天中第 $j (j=1, 2, 3)$ 时刻的值组成。

假设对第 d 天 t 时刻的电负荷、冷负荷进行预测,则输入矩阵 \mathbf{I}_d^t 表达式如下:

$$\mathbf{I}_d^t = [\mathbf{E}_d^t, \mathbf{C}_d^t, \mathbf{T}_d, \mathbf{R}_d, \mathbf{M}_d]^T \quad (13)$$

式中, \mathbf{E}_d^t 为电负荷历史数据矩阵; \mathbf{C}_d^t 为冷负荷历史数据矩阵; \mathbf{T}_d 为温度矩阵; \mathbf{R}_d 为辐射强度矩阵; \mathbf{M}_d 为空气湿度矩阵。

$$\mathbf{E}_d^t = [\mathfrak{S}_{i=d-3}^{d-1}(\mathfrak{S}_{j=t-1}^{t+1}(eload_i^j))] \quad (14)$$

$$\mathbf{C}_d^t = [\mathfrak{S}_{i=d-3}^{d-1}(\mathfrak{S}_{j=t-1}^{t+1}(cload_i^j))] \quad (15)$$

$$\mathbf{T}_d = [\mathfrak{S}_{i=d-3}^d(T_{\max,i}, T_{\min,i}, T_{\text{ave},i})] \quad (16)$$

$$\mathbf{R}_d = [\mathfrak{S}_{i=d-3}^d(R_{\max,i}, R_{\min,i}, R_{\text{ave},i})] \quad (17)$$

$$\mathbf{M}_d = [\mathfrak{S}_{i=d-3}^d(M_{\max,i}, M_{\min,i}, M_{\text{ave},i})] \quad (18)$$

式中, $eload_i^j, cload_i^j$ 分别为第 i 天 j 时刻的电负荷和冷负荷的历史数值,共 18 维; $T_{\max,i}, T_{\min,i}, T_{\text{ave},i}$ 分别为第 i 天温度的最大值、最小值和平均值,共 12 维; $R_{\max,i}, R_{\min,i}, R_{\text{ave},i}$ 分别为第 i 天辐射强度的最大值、最小值和平均值,共 12 维; $M_{\max,i}, M_{\min,i}, M_{\text{ave},i}$ 分别为第 i 天空气湿度的最大值、最小值和平均值,共 12 维。总输入矩阵 \mathbf{I}_d^t 共 54 维。

输出矩阵 \mathbf{O}_d^t 表达式如下:

$$\mathbf{O}_d^t = [eload_d^t, cload_d^t]^T \quad (19)$$

式中, $eload_d^t, cload_d^t$ 分别为第 d 天 t 时刻电负荷预测数值和冷负荷预测数值。

4.2 区域综合能源系统负荷预测

为验证本文预测模型的有效性,分别采用本文提出的 KPCA-GRNN 模型、PCA-GRNN 模型与 GRNN 模型对电负荷和冷负荷进行预测并比较三种模型的预测准确度。首先对原始数据进行核主成分分析与主成分分析,为保证所选主成分能够更好地代替原始数据,选择累计贡献率大于 98% 的主成分构成新的数据集,结果如表 3 所示。其中, d 表示所

选主成分维数, r 表示累计贡献率。

表 3 KPCA 与 PCA 结果比较

Tab.3 Comparison of KPCA and PCA results

日期类型	KPCA		PCA	
	d	$R(\%)$	d	$R(\%)$
工作日	8	98.17	15	98.24
休息日	9	98.08	17	98.15

从表 3 结果可看出,通过 PCA 进行降维之后,原有的 54 维输入矩阵,分别降低到工作日的 15 维以及休息日的 17 维,而经过 KPCA 降维之后,原有的 54 维输入矩阵,分别降低到工作日的 8 维以及休息日的 9 维。由此可知,在累计贡献率基本相同的情况下, KPCA 在 PCA 的基础之上再次降低了 7 维和 8 维,故 KPCA 比 PCA 具有更加良好的降维效果,表明了 KPCA 对于处理存在非线性耦合的数据的优越性。

采用遗传算法对三种模型光滑因子进行优化,在保证算法性能以及效率的基础上,通过实验确定遗传算法的参数设置为:选择电负荷、冷负荷均方根差的和作为遗传算法的适应度函数,种群数为 40,最大遗传代数为 100,交叉概率 0.7,变异概率 0.01,寻优区间为 $[0.001, 1]$ 。利用含最优光滑因子的 GRNN 神经网络对电负荷、冷负荷进行预测,负荷预测相对误差曲线如图 3 和图 4 所示,其中模型 1 为本文模型,模型 2 为 PCA-GRNN 模型,模型 3 为 GRNN 模型。

由图 3 可知,对于电负荷预测,本文模型相对误差基本在 $[0, 4\%]$ 范围内变化,而 PCA-GRNN 模型和 GRNN 模型相对误差则分别在 $[0, 5\%]$ 和 $[0, 6\%]$ 范围内变化,故本文模型的误差更小;由图 4 可知,对于冷负荷预测,本文模型相对误差基本在 $[0, 5\%]$ 范围内变化,而其余两种模型相对误差在 $[0, 7\%]$ 范围内变化,故本文模型的误差更小。综上所述,本文模型与 PCA-GRNN 模型的相对误差均小于 GRNN 模型,表明了降维解耦处理可以有效降低预测误差,而本文模型与 PCA-GRNN 模型相比,本文模型误差更小,证明了对于含有非线性相关数据的降维解耦处理, KPCA 比 PCA 更加有效。

为了更直观地评价三种模型的优劣并验证模型的稳定性,对未来 3 天的多元负荷进行预测,以平均相对误差、最大相对误差以及均方根差作为衡量标准,对三种预测模型进行比较,表 4 为电负荷预测误差,表 5 为冷负荷预测误差。其中, E_{MAPE} 为平均相

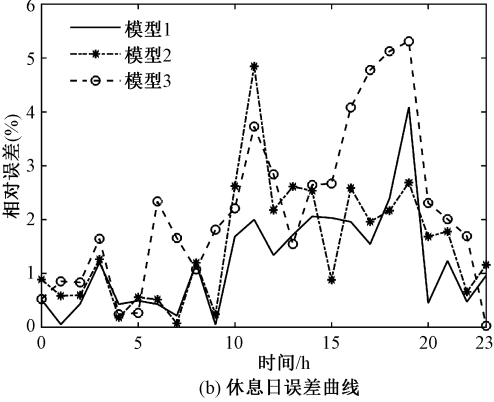
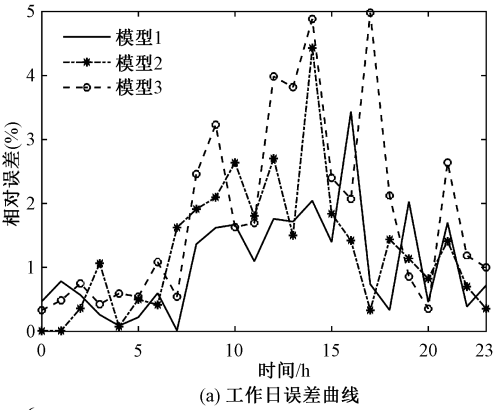


图 3 电负荷预测相对误差曲线

Fig.3 Relative error curve of electrical load prediction

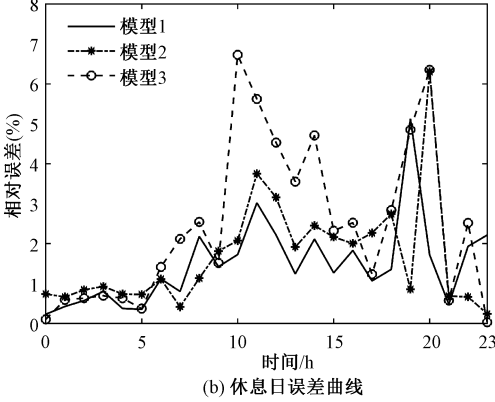
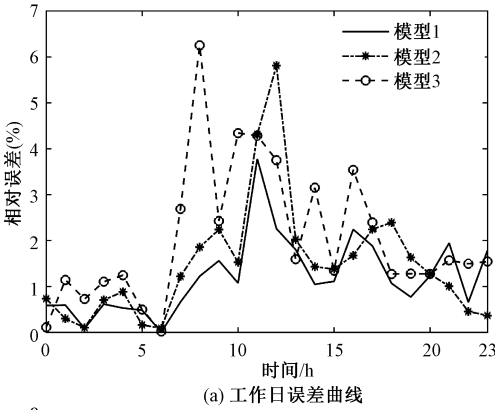


图 4 冷负荷预测相对误差曲线

Fig.4 Relative error curve of cooling load prediction

对误差, E_{MAX} 为最大相对误差, $RMSE$ 为均方根差。从表 4 和表 5 可以看出,对于未来 3 天的电负荷预测,本文预测模型平均相对误差为 1.52% 和 1.81%,最大相对误差为 3.51%和 4.12%,相比其余两种模型误差均有降低,均方根差也小于其余两种模型;对于未来 3 天的冷负荷预测,本文预测模型平均相对误差为 1.73%和 2.04%,最大相对误差为 3.86%和 5.38%,相比其余两种模型有较大幅度的

降低,同时均方根差也小于其余两种模型。通过以上数据对比,进一步证明了本文模型对夏季多元负荷预测的高精度以及稳定性,具有一定的实用性。

5 结论

本文提出了基于 Copula 理论与 KPCA-GRNN 结合的区域综合能源系统多元负荷短期预测模型。该模型在 Copula 理论的基础上分析了多元负荷之

表 4 电负荷预测误差

Tab.4 Electrical load prediction errors

日期类型	$E_{MAPE}(\%)$			$E_{MAX}(\%)$			$RMSE/kW$		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
工作日	1.52	1.94	2.63	3.51	4.82	5.71	0.77	0.98	1.38
休息日	1.81	2.37	3.09	4.12	5.46	6.93	0.48	0.59	0.80

表 5 冷负荷预测误差

Tab.5 Cooling load prediction error

日期类型	$E_{MAPE}(\%)$			$E_{MAX}(\%)$			$RMSE/kW$		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
工作日	1.73	2.46	3.15	3.86	6.23	7.38	1.18	1.83	2.15
休息日	2.04	2.75	3.42	5.38	6.72	7.69	0.89	1.40	2.32

间以及负荷与天气因素之间的相关性,强调了在进行预测过程中多元负荷间的不可分割性和各影响因子的不可忽略性。然后通过 KPCA 对数据进行简化降维,配合遗传算法优化 GRNN 神经网络的光滑因子,提高模型的预测精度。算例结果表明,相比于 PCA-GRNN 模型与 GRNN 模型,本文模型具有较高的预测精度,对于区域综合能源系统的经济稳定运行具有一定实际意义。本文模型按照日期类型将预测分为两部分,在一定程度上割裂了负荷之间的连续性,在今后工作中将对日期类型进行合理量化,令其作为负荷预测影响因子,增加模型的合理性。

参考文献 (References):

- [1] Bacher P, Madsen H, Nielsen H A. Online short-term heat load forecasting for single family houses [J]. *Energy & Buildings*, 2013, 65 (4): 101-112.
- [2] Powell K M, Sriprasad A, Cole W J, et al. Heating, cooling, and electrical load forecasting for a large-scale district energy system [J]. *Energy*, 2014, 74: 877-885.
- [3] Chou J S, Bui D K. Modeling heating and cooling loads by artificial intelligence for energy-efficient building design [J]. *Energy & Buildings*, 2014, 82: 437-446.
- [4] Izadyar N, Ghadamian H, Ong H C, et al. Appraisal of the support vector machine to forecast residential heating demand for the district heating system based on the monthly overall natural gas consumption [J]. *Energy*, 2015, 93: 1558-1567.
- [5] 苏雅茹 (Su Yaru). 高维数据的维数约简算法研究 (Research on dimensionality reduction of high-dimensional data) [D]. 合肥: 中国科学技术大学 (Hefei: University of Science and Technology of China), 2012.
- [6] 李聪 (Li Cong). 基于 GRNN 网络的短期与超短期负荷预测 (Short-term and ultra-short-term load forecasting based on GRNN network) [D]. 长春: 东北电力大学 (Changchun: Northeast Electric Power University), 2010.
- [7] 韦艳华, 张世英 (Wei Yanhua, Zhang shiying). 金融市场的相关性分析——Copula-GARCH 模型及其应用 (Dependence analysis of finance markets: Copula-GARCH model and its application) [J]. *系统工程 (Systems Engineering)*, 2004, 22 (4): 7-12.
- [8] 王涛, 王淳, 李成豪 (Wang Tao, Wang Chun, Li Chenghao). 基于 Copula 函数及 Rosenblatt 变换的含相关性概率潮流计算 (Probabilistic load flow calculation based on Copula function and Rosenblatt transformation considering correlation among input variables) [J]. *电力系统保护与控制 (Power System Protection and Control)*, 2018, 46 (21): 24-30.
- [9] 姜潮, 张旺, 韩旭 (Jiang Chao, Zhang Wang, Han Xu). 基于 Copula 函数的证据理论相关性分析模型及结构可靠性计算方法 (A Copula function based evidence theory model for correlation analysis and corresponding structural reliability method) [J]. *机械工程学报 (Journal of Mechanical Engineering)*, 2017, 53 (16): 199-209.
- [10] 钟嘉庆, 李茂林, 江静, 等 (Zhong Jiaqing, Li Maolin, Jiang Jing, et al.). 基于 Copula 理论的风/光出力预测误差分析方法的研究 (Method of wind/solar output forecast error analysis based on Copula theory) [J]. *电工电能新技术 (Advanced Technology of Electrical Engineering and Energy)*, 2017, 36 (6): 39-46.
- [11] 谢中华 (Xie Zhonghua). MATLAB 统计分析与应用 (MATLAB statistical analysis and application) [M]. 北京: 北京航空航天大学出版社 (Beijing: Beihang University Press), 2015.
- [12] 王飞 (Wang Fei). 并网型光伏电站发电功率预测方法与系统 (Power forecasting approach and system of grid-connected PV plant) [D]. 北京: 华北电力大学 (Beijing: North China Electric Power University), 2013.
- [13] 黄志强, 任鸿飞, 梁芳敏, 等 (Huang Zhiqiang, Ren Hongfei, Liang Fangmin, et al.). 核主成分分析在地球化学元素组合异常圈定中的应用 (The application of kernel principal component analysis to the anomaly delineation of geochemistry element association) [J]. *物探化探计算技术 (Computing techniques for geophysical and geochemical exploration)*, 2016, 38 (4): 560-565.
- [14] 贾花萍 (Jia Huaping). GRNN 神经网络在电力系统负荷预报中的应用 (GRNN neural network in the application of power system load forecasting) [J]. *电子设计工程 (Electronic Design Engineering)*, 2012, 20 (3): 14-16.
- [15] 邹红波, 伏春林, 喻圣 (Zou Hongbo, Fu Chunlin, Yu Sheng). 基于 Akima-LMD 和 GRNN 的短期负荷预测 (Short-term load forecasting based on Akima-LMD and GRNN) [J]. *电工电能新技术 (Advanced Technology of Electrical Engineering and Energy)*, 2018, 37 (1): 51-56.
- [16] 张淑清, 任爽, 姜安琦, 等 (Zhang Shuqing, Ren Shuang, Jiang Anqi, et al.). PCA-GRNN 在综合气象短期负荷预测中的应用 (Application of PCA-GRNN in integrated meteorological short-term load forecasting)

[J]. 计量学报 (Acta Metrologica Sinica), 2017, 38(3): 340-344.

[17] 周俊文, 宋晓霞 (Zhou Junwen, Song Xiaoxia). 基于 GRNN 神经网络的面料热阻预测模型研究 (Study on thermal resistance prediction model of fabrics based on general regression neural network) [J]. 丝绸 (Silk), 2018, 55 (8): 41-46.

Short-term multiple load prediction model for regional integrated energy system based on Copula theory and KPCA-GRNN

MA Jian-peng¹, GONG Wen-jie², ZHANG Zhi-sheng¹

- (1. College of Electrical Engineering, Qingdao University, Qingdao 266071, China;
2. State Grid Qingdao Electric Power Supply Company, Qingdao 266002, China)

Abstract: The nonlinear correlation between multiple loads and between multiple loads and weather factors is analyzed by Copula theory. Based on the analysis results, the model sample set is constructed. The sample set is decomposed by KPCA. The GRNN model is used to predict the multiple load and the model parameters are optimized by genetic algorithm. Taking the calculation example of electrical load and cooling load in a building-grade regional integrated energy system in north China as an example, compared with the other two prediction models, the verification results show that the model in this paper has higher prediction accuracy and better application value.

Key words: regional integrated energy system; multiple load forecasting; Copula theory; kernel principal component analysis; generalized regression neural network