

DOI:10.13409/j.cnki.jdpme.20230422002

# 基于改进 K-means 算法的大跨屋盖结构表面风荷载分区研究\*

李玉学<sup>1,3</sup>, 杨君保<sup>2</sup>, 陈铁<sup>1,3</sup>, 田玉基<sup>4</sup>

(1. 石家庄铁道大学土木工程学院, 石家庄 050043; 2. 河北省建筑材料工业设计研究院有限公司, 石家庄 050050;  
3. 石家庄铁道大学道路与铁道工程安全保障教育部重点实验室, 石家庄 050043;  
4. 北京交通大学结构风工程与城市风环境北京市重点实验室, 北京 100044;)

**摘要:** 针对 K-means 算法进行大跨屋盖结构表面风荷载分区中存在的分类数  $k$  值需凭经验事先给定以及所有初始聚类中心均需随机选取带来的分类情况数过多、从中寻找最优分类结果工作量大且效率低的问题, 提出基于改进 K-means 算法的大跨屋盖结构表面风荷载分区方法。首先, 建立分类数  $k$  与其相应测点风荷载的误差平方和 (Sum of the Squared Errors: SSE) 关系曲线, 引入手肘法基本思想, 实现最优分类数  $k_{\text{opt}}$  值的精准识别; 其次, 在首个初始聚类中心随机选取基础上, 引入轮盘法基本思想, 完成对剩余初始聚类中心的高效选取; 然后, 根据类内紧凑、类间分散的原则, 通过类内紧凑性判定指标  $S(k)$  和类间分散性判定指标  $D(k)$ , 构造并借助  $SD(k)$  值有效性检验, 得到最优的风荷载分区结果; 最后, 以北京奥林匹克网球中心大跨悬挑屋盖结构为例, 针对风洞试验所得风荷载测试结果, 采用所提方法对其表面最不利风压系数进行分区计算, 并与传统 K-means 算法进行对比, 结果表明, 所提方法能够高效实现大跨屋盖结构表面风压分区计算, 具有较好的工程应用价值。

**关键词:** 大跨屋盖结构; 风荷载分区; K-means 算法; 分类数; 聚类中心

**中图分类号:** TU393.3; TU312.1 **文献标识码:** A **文章编号:** 1672-2132(2024)05-1106-09

## Study on Wind Load Zoning of Large-span Roofs Based on Improved K-means Algorithm

LI Yuxue<sup>1,3</sup>, YANG Junbao<sup>2</sup>, CHEN Tie<sup>1,3</sup>, TIAN Yuji<sup>4</sup>

(1. School of Civil Engineering, Shijiazhuang Tiedao University, Shijiazhuang 050043, China;  
2. Hebei Building Materials Industry Design and Research Institute Co., Ltd., Shijiazhuang 050050, China;  
3. Key Laboratory of Roads and Railway Engineering Safety Control of China Ministry of Education, Shijiazhuang Tiedao University, Shijiazhuang 050043, China; 4. Beijing's Key Laboratory of Structural Wind Engineering and Urban Wind Environment, Beijing Jiaotong University, Beijing 100044, China)

**Abstract:** In the application of the K-means algorithm for the wind load zoning on the surface of large-span roof structures, the classification number  $k$  values are given in advance by experience, and all initial clustering centers are randomly selected. This often results in an excessive number of classification, leading to increased workload and low efficiency in identifying the optimal classification results.

\* 收稿日期: 2023-04-22; 修回日期: 2023-10-19

基金项目: 国家自然科学基金项目(51278314)、中央引导地方科技发展资金项目(206Z5401G)、北京交通大学“结构风工程与城市风环境北京市重点实验室”开放课题(2023-1)资助

作者简介: 李玉学(1978—), 男, 副教授, 博导, 博士。主要从事结构风工程研究。E-mail: liyuxue2000@163.com

To address these issues, this study proposed a wind load zoning method for large-span roof structures based on improved K-means algorithm. First, a relationship curve between the classification number  $k$  and the sum of the squared errors (SSE) of wind loads at the corresponding pressure taps was established, incorporating the Elbow Method to accurately determine the optimal classification number  $k_{st}$  values. Next, after randomly selecting the first initial clustering center, the Roulette Wheel method was introduced to efficiently select the remaining initial cluster centers. Following this, based on the principles of intra-cluster compactness and inter-cluster dispersion, compactness criterion  $S(k)$  and dispersion criterion  $D(k)$  were employed to construct and validate the zoning effectiveness using the SD ( $k$ ) value, which resulted in the optimal wind load zoning scheme. Finally, taking the large-span cantilevered roof structure of Beijing Olympic Tennis Center as an example, wind tunnel test results were employed to calculate the most unfavorable wind pressure coefficients on the surface of the structure. A comparison with the traditional K-means algorithm demonstrated that the proposed method efficiently achieved wind pressure zoning for large-span roof structures and holds significant engineering application value.

**Keywords:** large-span roof structures; wind load zoning; K-means algorithm; classification number; clustering centers

## 0 引言

《建筑结构荷载规范》<sup>[1]</sup>规定,针对体型复杂的大跨屋盖结构,需要借助风洞试验获取用于结构设计的风荷载。实际工作中,为提高结构设计计算效率,通常将风洞试验所得结构表面多个测点风荷载处理为数量较少的区域风荷载,并根据相应区域的风荷载代表值进行结构设计计算。因此,合理、高效地进行大跨屋盖结构表面风荷载分区,对于兼顾结构设计计算精度和效率非常关键。

针对大跨屋盖结构表面风荷载分区问题,国内外不少学者进行了研究。李丹煜等<sup>[2]</sup>基于聚类分析方法,对于平屋盖、鞍形屋盖极值风压系数分区进行了研究,发现聚类分析方法可以一定程度地解决屋盖表面风荷载分区问题。聂燕松<sup>[3]</sup>基于以平均、脉动风荷载为基向量的等效静力风荷载表达式,给出了球面、柱面、鞍形屋盖为代表的大跨屋盖结构设计风荷载,并采用K-means聚类算法对此设计风荷载进行了分区计算。孙晓颖等<sup>[4]</sup>通过数值模拟得到了风向角、跨高比、矢跨比、地面粗糙度、风速等因素对鞍形屋盖风压分布的影响,并根据屋面结构形式及风压分布特点将鞍形屋面分成12个区。顾明等<sup>[5]</sup>分别针对底部开敞和封闭两种形式鞍形索网屋盖进行风洞试验测试,并将试验所得屋盖表面风荷载也划分为12个区,同时比较了两种不同形式鞍

形索网屋盖表面相应区域风荷载差异。L. Pagnini等<sup>[6]</sup>基于有无周围建筑物情况下大跨柱面屋盖表面风荷载风洞试验测试数据,得到了不同风向角下屋盖表面平均风压和极值风压的分区结果。李云等<sup>[7]</sup>提出了大跨屋盖结构表面风荷载的多目标分区方法,并据此进行了不同风向角下大跨屋盖表面风洞试验所得风荷载的分区计算。H. Alrawashdeh等<sup>[8]</sup>基于风洞试验测试结果,提出了平屋盖表面局部体型系数的分区方法,在此基础上,给出了针对美国及加拿大风荷载规范中关于平屋盖局部体型系数分区结果的修正建议。J. Cao等<sup>[9]</sup>将正方形平屋盖分为角部区域和中心区域,基于风洞试验测试结果,分别给出了 $0^\circ$ 和 $45^\circ$ 风向角下,上述两个区域的分区体型系数。E. Gavanski等<sup>[10]</sup>基于风洞试验测试结果,研究了不同坡度和跨数条件下锯齿形屋面风荷载分区结果,并与ASCE7-16规范中的分区结果进行对比,据此提出了规范修正建议。

本文基于K-means算法基本原理,并针对其用于大跨屋盖结构表面风荷载分区计算时,存在的分类情况数过多、寻找最优分区结果工作量大、计算效率低等问题,通过引入肘法和轮盘法分别实现聚类数和初始聚类中心的高效选取,建立基于改进K-means算法的大跨屋盖结构表面风荷载分区方法,并以北京奥林匹克公园网球中心大跨悬挑屋盖结构表面风荷载风洞试验测试数据为基础,对其用于结构设计的最不利风压系数进行分区计算,以对

所提方法进行说明和验证。

## 1 K-means 聚类算法

### 1.1 算法原理

K-means 聚类算法是基于迭代求解的划分型聚类算法,该方法将欧式距离作为两数据样本相似性评价指标,基于该算法将风洞试验所得大跨屋盖表面不同测点风荷载数据样本进行分区,具体为:

(1)将风洞试验测试所得大跨屋盖结构表面  $n$  个测点风荷载值表示为一个数据集  $P = \{p_i | i = 1, 2, \dots, n\}$ ,从数据集  $P$  中随机选取  $k$  个测点风荷载值作为拟划分的  $k$  个聚类中心  $c_j (j = 1, 2, \dots, k)$ ,并将任意测点风荷载值  $p_i$  与第  $j$  类聚类中心  $c_j$  间欧式距离简化表示为  $d(p_i, c_j) = |p_i - c_j|$ 。

(2)分别计算  $n$  个测点风荷载值  $p_i$  与选定的  $k$  个聚类中心  $c_j$  间欧式距离  $d(p_i, c_j) = |p_i - c_j|$ 。

(3)根据  $d(p_i, c_j)$  计算结果,依据  $d(p_i, c_j)$  最小原则,将  $n$  个测点风荷载值  $p_i$  划分为  $k$  类。

(4)将划分得到的  $k$  类内所有测点风荷载值取平均,并将其作为新的聚类中心,重复(2)~(4)步骤,直到相邻迭代两次计算得到的新聚类中心数值不再改变,并将其作为最终分类结果。

(5)选取最后的聚类中心值作为相应分区风荷载代表值,同一类内所有测点风荷载影响范围作为风荷载分类区域。

### 1.2 算法不足

由 1.1 可见,采用 K-means 聚类算法对大跨屋盖结构表面测点风荷载分区计算过程中存在两方面不足:(1)最初选取的类数  $k$  通常需要根据经验主观设定,比如,有学者提出  $k_{\max} \leq \sqrt{n}^{[11]}$  或者  $k_{\max} \leq 2\ln(n)^{[12]}$  ( $n$  为测点总数,  $\ln$  为自然对数),上述处理方法一定程度地解决了  $k$  值选取随机性问题,但都是基于经验,理论依据不足。理论上讲,类数  $k$  可以选定  $1 \sim n$  的任意值,而且选定的数量不同,测点风荷载分区结果差异较大,若分类数  $k$  取值较大,起不到在结构设计时简化风荷载施加的目的,选定分类数  $k$  值较小,会造成用于结构设计的分区风荷载值太粗糙、不准确,给结构抗风安全带来不可预见的隐患。(2)聚类的所有初始聚类中心均是随机选取,

初始聚类中心的选取结果直接决定了方法的最终分区结果,所选取的任意两初始聚类中心间欧式距离过大或过小,均不能实现最终分区结果中类内测点风荷载数值紧凑、类间测点风荷载数值分散的聚类目标。因此,面对大跨屋盖结构表面众多的测点风荷载数据,需要从中搜索较优的初始聚类中心,此过程需要耗费大量的时间,降低了方法的使用效率。

## 2 算法改进

### 2.1 手肘法的引入

针对 1.2 节所述 K-means 算法进行大跨屋盖结构风荷载测点数据分区计算中分类数  $k$  值选取存在的不足,本文引入手肘法基本思想对其改进<sup>[13-14]</sup>,具体为:

(1)采用传统 K-means 聚类算法依次对测点风荷载进行  $k = 1, 2, \dots, n$  的分类计算,求取每个  $k$  值相应所有分类情况中的第一种情况所对应误差平方和 SSE 值(因为每个  $k$  值会对应多种分类情况,但这些情况的 SSE 值差别很小,几乎不变,如此可提高计算效率),作为该  $k$  值对应 SSE 值:

$$SSE = \sum_{r=1}^k \sum_{i=1}^{v_r} |p_{ri} - \bar{p}_r|^2 \quad (1)$$

式中,  $p_{ri}$  为第  $r$  类中第  $i$  个测点风荷载值;  $\bar{p}_r$  为第  $r$  类中所有测点风荷载平均值;  $v_r$  为第  $r$  类中所有测点数量。

(2)为方便计算和直观显示,当随着  $k$  值的增长, SSE 值几乎不再变化(通过试算根据经验连续出现 3 次 SSE 值为上个 SSE 值的 90%~110%)时,之后  $k$  值对应的 SSE 值不再计算。根据已有  $k$  值及由式(1)求得的 SSE 值,绘制  $k \sim SSE$  关系曲线,连接  $k \sim SSE$  关系曲线首、末两个端点,得到直线  $y = ak + b (k = 1, 2, \dots, l)$  ( $l$  为 SSE 值几乎不变时所对应  $k$  值),将  $k = 1, 2, \dots, l$  分别代入直线方程,得到数据集  $Y = \{y_1, y_2, \dots, y_l\}$ 。

(3)将  $k \sim SSE$  关系曲线中与  $k = 1, 2, \dots, l$  对应的 SSE 值记为数据集  $Z = \{z_1, z_2, \dots, z_l\}$ 。据此,将数据集  $Y$  和  $Z$  中相对应元素作差,得到数据集  $H = \{h_1, h_2, \dots, h_l\}$ ,选取数据集  $H$  中最大元素  $h_{\max}$ ,并将此最大元素  $h_{\max}$  对应的  $k$  值选为最优分类数  $k_{st}$ 。

对于大跨屋盖结构表面多个测点风荷载样本值而言,随着类数 $k$ 的增大,样本划分更加精细,相应的同一类中测点数据聚合程度越高,其由式(1)计算所得SSE值越小。但当类数 $k$ 超过最优分类数 $k_{st}$ 时,由 $k$ 的增加带来的数据样本聚合程度回报率会迅速变小,此时SSE值减小幅度也会明显变小,相应的 $k\sim$ SSE关系曲线也会趋于平缓。可见,采用上述(1)~(3)步骤能够较好实现最优分类数 $k_{st}$ 的识别。

需要说明的是,上述(1)~(3)步骤是基于同一类内测点风荷载数据样本聚合程度较高原则求得的最优分类数 $k_{st}$ 。同时,为了兼顾不同类间测点风荷载数据样本差别较大原则,本文在保证同一类内测点风荷载数据样本聚合程度较高基础上,将初始分类数 $k$ 的选取范围扩展为 $k_{st}-1$ 、 $k_{st}$ 和 $k_{st}+1$ 共三个值。

## 2.2 初始聚类中心的合理选取

针对1.2节所述K-means算法进行大跨屋盖结构风荷载测点数据分区中初始聚类中心随机选取存在的不足,本文引入轮盘法基本思想对其改进<sup>[15]</sup>,具体为:

(1)从 $n$ 个测点风荷载 $P=\{p_i|i=1,2,\dots,n\}$ 数据集中随机选取一个测点风荷载值作为第一个聚类中心 $c_1$ 。

(2)针对剩余的 $n-1$ 个测点风荷载数据,基于轮盘法基本思想,分别根据公式(2)计算其被选为第二个聚类中心的概率,进而选出第二个初始聚类中心 $c_2$ 。

$$G_1^i = \frac{F_1^i}{\sum_{i=1}^{n-1} F_1^i} \quad (2)$$

式中, $F_1^i=|p_i-c_1|(i=1,2,\dots,n-1)$ 表示剩余的 $n-1$ 个测点风荷载数据值与 $c_1$ 的欧式距离。

(3)选定 $c_1$ 、 $c_2$ 两个初始聚类中心后,针对剩余的 $n-2$ 个测点风荷载数据,继续利用轮盘法基本思想,分别根据公式(3)计算其被选为第三个聚类中心的概率,进而选出第三个初始聚类中心 $c_3$ 。

$$G_2^i = \frac{F_2^i}{\sum_{i=1}^{n-2} F_2^i} \quad (3)$$

式中, $F_2^i=\min\{|p_i-c_1|,|p_i-c_2|\}(i=1,2,\dots,n-2)$ 表示选定 $c_1$ 、 $c_2$ 两个初始聚类中心后,剩余的 $n-$

2个测点风荷载数据值分别与 $c_1$ 和 $c_2$ 欧式距离的最小值。

(4)继续按照步骤(2)~(3)的方法,选取其余的初始聚类中心,直到2.1节分类数 $k_{st}-1$ 、 $k_{st}$ 和 $k_{st}+1$ 对应的较优初始聚类中心分别全部选出。

本文改进K-means算法的分析流程如图1所示。

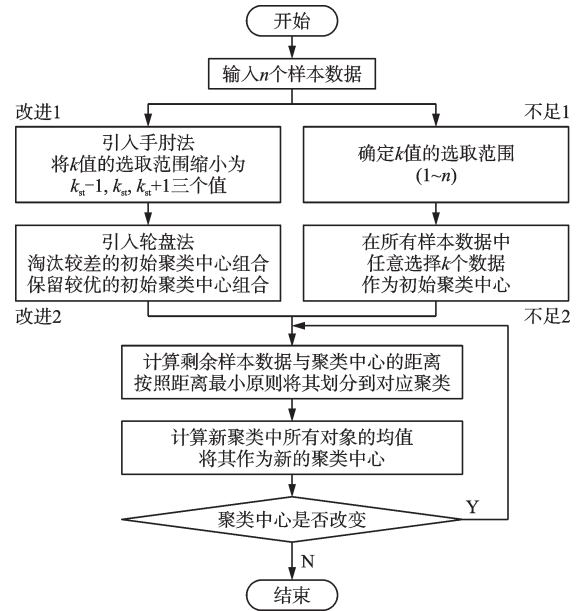


图1 改进K-means算法流程

Fig.1 Flow chart of improved K-means algorithm

## 2.3 $SD(k)$ 值有效性检验

为了找到最优分类结果,需要基于2.1和2.2节计算所得 $k_{st}-1$ 、 $k_{st}$ 和 $k_{st}+1$ 对应的所有分类情况,根据类内紧凑、类间分散的原则选出最优分类结果。本文通过 $SD(k)$ 值有效性检验实现这一过程<sup>[16]</sup>,具体为:

(1)定义类内紧凑性判定指标 $S(k)$ ,此 $S(k)$ 值越大,表示类内紧凑性越低。

$$S(k) = \frac{1}{k} \sum_{m=1}^k \frac{\sigma(p_m)}{\sigma(p)} \quad (4)$$

式中, $\sigma(p)$ 为全部测点风荷载值的方差; $\sigma(p_m)$ 为第 $m$ 类内全部测点风荷载值的方差; $k$ 为分类数,分别取2.1节所得 $k_{st}-1$ 、 $k_{st}$ 和 $k_{st}+1$ 三个值。

(2)定义类间分散性判定指标 $D(k)$ ,此 $D(k)$ 值越大,表示类间分散性越低。

$$D(k) = \frac{D_{\max}}{D_{\min}} \sum_{t=1}^k \left( \sum_{s=1}^k |c_t - c_s| \right)^{-1} \quad (5)$$

式中,  $D_{\max} = \max(|c_t - c_s|)$ ,  $D_{\min} = \min(|c_t - c_s|)$  ( $\forall t, s \in \{1, 2, \dots, k\}, t \neq s$ ), 分别为第  $t, s$  两类中心间最大和最小欧式距离;  $c_t$  为第  $t$  类内所有测点风荷载的均值;  $c_s$  为第  $s$  类内所有测点风荷载的均值; 同样,  $k$  为分类数, 分别取 2.1 节所得  $k_{st}-1$ 、 $k_{st}$  和  $k_{st}+1$  三个值。

(3) 构造兼顾类内紧凑性和类间分散性的综合指标  $SD(k)$ , 此  $SD(k)$  值越小, 表示类内越紧凑、类间越分散, 综合分类结果越理想。

$$SD(k) = \alpha S(k) + D(k) \quad (6)$$

式中,  $\alpha$  为加权因子,  $\alpha = D(k_{\max})$ ,  $k_{\max}$  为分类数的最大值, 本文取  $k_{st}+1$ 。

(4) 根据  $k_{st}-1$ 、 $k_{st}$  和  $k_{st}+1$  对应的所有分类情况, 分别由公式(6)计算其对应的  $SD(k)$  值, 并从中选定最小的  $SD(k)$  值对应的分类情况作为最优分类结果。

## 2.4 分区风荷载代表值计算

根据 2.3 节所得大跨屋盖结构表面风荷载测点最优分类结果, 以各个类内测点的从属面积为权重计算相应类内的风荷载代表值, 并以相邻两测点中线作为其从属面积分界线, 所有同一类内最外层分界线围成的面积为相应类的风荷载划分区域, 据此第  $R$  区域的风荷载代表值可以表示为:

$$P_R = \sum_{i=1}^{n_R} (p_{Ri} \cdot A_i) / A_R \quad (7)$$

式中,  $p_{Ri}$  为第  $R$  区内第  $i$  个测点风荷载值;  $A_i$  为测点  $i$  的屋面从属面积;  $A_R$  为第  $R$  区的屋面总面积;  $n_R$  为第  $R$  区内测点总数。

## 3 算例分析

为了说明和验证本文所提改进的 K-means 聚类算法对大跨屋盖结构表面测点风荷载分区计算, 以北京奥林匹克公园网球中心大跨悬挑屋盖结构为例, 基于风洞试验所得结构表面风荷载测试数据对其进行分区计算。

### 3.1 工程简介

北京奥林匹克公园网球中心大跨悬挑屋盖结构为“莲花”造型(图 2), 底部直径 79.7 m, 坐落在高

6.0 m 的平台上, 平台、看台和 21.6 m 长的看台挑棚共同组成一个“Z”字形, 整个屋盖由 12 个“Z”字形单元组成, 相邻单元之间开有较大洞口, 挑棚前檐最高点标高 23.0 m, 挑棚设有大约 3° 倾角, 以满足排水需要。屋盖结构形状特殊, 钝体绕流特征明显, 属于《建筑结构荷载规范》<sup>[1]</sup> 规定的典型风敏感结构, 其结构设计用风荷载需要借助风洞试验获取。



图 2 北京奥林匹克公园网球中心大跨悬挑屋盖

Fig.2 Large-span cantilevered roof of Beijing Olympic Park Tennis Center

### 3.2 风洞试验测试及数据处理

北京奥林匹克公园网球中心大跨悬挑屋盖结构模型风洞试验在北京大学力学与工程科学系的大型低速风洞中完成, 试验段长 3.65 m, 直径为 2.25 m。屋盖刚性测压模型采用 ABS 材料和有机玻璃制成, 模型缩尺比为 1:120, 阻塞比为 0.94%, 地面粗糙度取 B 类, 来流风速剖面和湍流度剖面如图 3 所示(图中  $Z$  和  $Z_r$  分别表示高度和参考点高度,  $Z_r = 23$  m;  $V$  和  $V_g$  分别表示风速和参考点风速,  $V_g = 15$  m/s;  $I_u$  表示湍流度)。

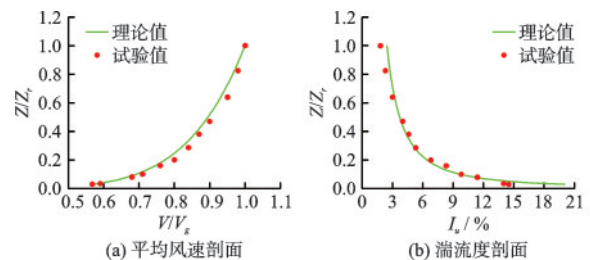


图 3 试验平均风速剖面和湍流度剖面

Fig.3 Profile of mean wind velocity and turbulence intensity

考虑到结构的对称性, 仅在第 3、4 单元上布置测点, 其中, 悬挑屋盖单元上表面布置 30 个测点, 下表面布置 25 个测点, 每个测点采样点数为 3 900, 采样频率为 40 Hz, 共测试了 18 个风向角, 定义正北方

向为0°风向角,顺时针增大,每个风向角下测试并采集数据一次,屋盖风洞试验模型、单元编号以及风向角定义如图4所示,单元上、下表面测点位置及编号如图5所示。

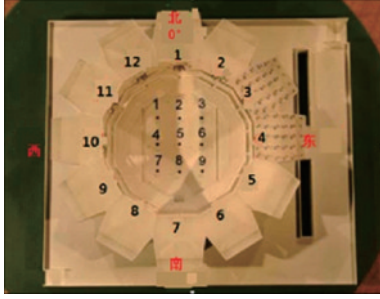


图4 屋盖风洞试验模型、单元编号及风向角定义

Fig.4 Model for wind tunnel test, unit numbers and wind angle definitions

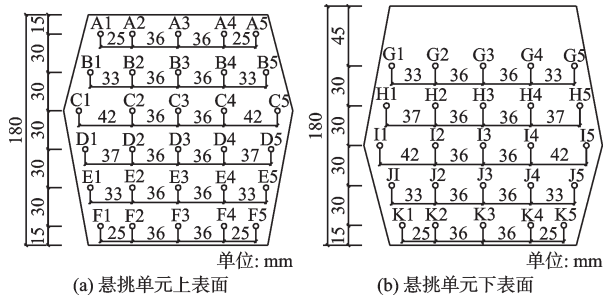


图5 悬挑屋盖单元表面测点位置及编号

Fig.5 Locations and numbers of pressure taps on the surface of cantilevered roof unit

定义风压值符号:沿悬挑屋盖表面的法线方向,远离屋盖表面为负,表现为风吸力,指向屋盖表面为正,表现为风压力。

为了表示方便,将屋盖表面测点*i*的风压值 $p_i(t)$ 处理为对应于参考点(选在屋盖最高点)处风压的无量纲压力系数(简称为风压系数),即:

$$C_{p_i}(t) = \frac{p_i(t) - p_\infty}{p_0 - p_\infty} \quad (8)$$

式中, $C_{p_i}(t)$ 为屋盖表面测点*i*的风压系数; $p_0$ 为参考点处的总压; $p_\infty$ 为参考点处的静压。

由式(9)可以得到屋盖表面测点*i*位置处,用于结构设计的极值风压系数:

$$\hat{C}_{p_i} = \bar{C}_{p_i} + \text{sign}(\bar{C}_{p_i}) g \tilde{C}_{p_i} \quad (9)$$

式中, $\bar{C}_{p_i}$ 和 $\tilde{C}_{p_i}$ 分别为屋盖表面测点*i*位置处的平均风压系数和脉动风压系数,分别由式(10)和(11)求得; $\text{sign}(g)$ 为符号函数; $g$ 为峰值因子,由极值穿越理论求得,本文统一取 $3.5^{[17]}$ 。

$$\bar{C}_{p_i} = \frac{\sum_{m=1}^N C_{p_i}(t_m)}{N} \quad (10)$$

$$\tilde{C}_{p_i} = \sqrt{\frac{\sum_{m=1}^N [C_{p_i}(t_m) - \bar{C}_{p_i}]^2}{N-1}} \quad (11)$$

式中, $C_{p_i}(t_m)$ 为*m*时刻屋盖表面测点*i*位置处的压力系数; $N$ 为总的采样点数。

### 3.3 风压分区计算

根据3.2节风洞试验所得悬挑屋盖单元上下表面风荷载测试数据,考虑所有风向角,利用公式(9)计算得到了用于结构设计的最不利正压系数和最不利负压系数如图6所示。

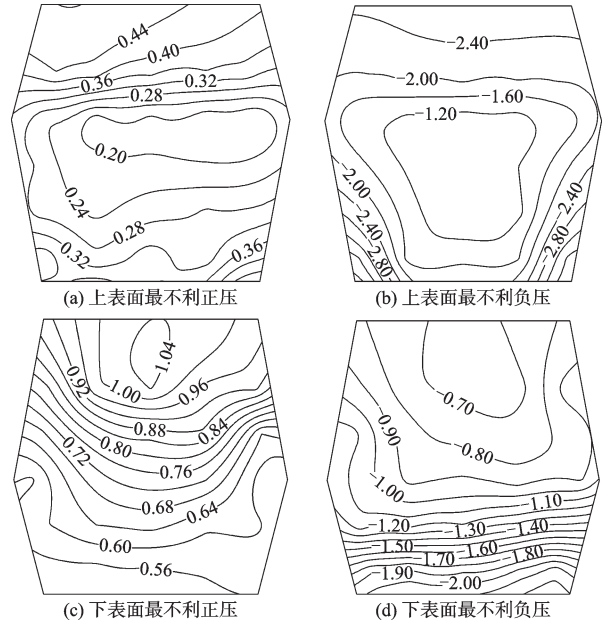


图6 屋盖单元表面最不利风压系数分布

Fig.6 Distribution of the most unfavorable wind pressure coefficients on the surface of the roof unit

由图6可见,悬挑屋盖单元上、下表面的最不利风压系数均具有一定的区域分布特性,表现为边、角以及端部区域位置数值较大,内部区域位置数值较小,此与屋盖体型及其导致的表面风场绕流形式密切相关。

为了明确阐述本文所提方法进行屋盖表面风压分区计算的具体过程并验证其有效性,随机选取悬挑屋盖单元下表面最不利正压(图6(c)),对其分区计算详细说明。

### 3.3.1 分类数 $k$ 值计算

根据图 6(c) 所示悬挑屋盖单元下表面 25 个测点最不利正风压系数计算结果, 采用 2.1 节所提方法依次计算  $k(k=1, 2, \dots, 25)$  值对应的 SSE 值, 并绘制  $k \sim SSE$  关系曲线, 计算发现, 当  $k=8$  时, 曲线接近水平, 几乎不变,  $k > 8$  以后的 SSE 值不再计算, 截取  $k \leq 8$  范围内的  $k \sim SSE$  关系曲线如图 7 所示。

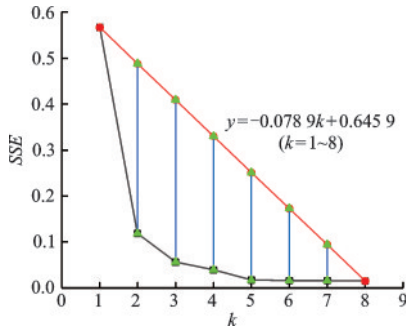


图 7  $k \sim SSE$  关系曲线

Fig.7 Relationship curves of  $k \sim SSE$

根据图 7 所示  $k \sim SSE$  关系曲线, 按照 2.1 节所述方法, 计算得到了每个  $k$  值对应的数据集  $H = \{h_1, h_2, \dots, h_8\} = \{0, 0.37, 0.35, 0.29, 0.23, 0.16, 0.08, 0\}$ , 其中数据集  $H$  中最大元素为 0.37, 其对应的  $k$  值为 2, 据此选定分类数  $k_{st} = 2$ 。

### 3.3.2 分类情况数量统计

由选定的分类数  $k_{st} = 2$ , 结合 2.2 节所述方法, 计算得到相应的悬挑屋盖单元下表面最不利正风压系数分类情况, 见表 1。

由表 1 可见, 根据选定的分类数  $k_{st} = 2$ , 共计算得到 6 种分类情况, 在此基础上, 根据类内紧凑、类间分散的原则, 利用 2.3 节所述  $SD(k)$  值有效性指标, 从中选定最优分类结果。

### 3.3.3 $SD(k)$ 值计算

根据公式 (4)~(6) 分别计算表 1 所列 6 种分类情况的  $SD(k)$  值, 见表 2。

由表 2 可见, 情况 2 对应的  $SD(k)$  值为 9.799 5, 在所有 6 种分类情况中其  $SD(k)$  值最小。据此, 选定情况 2 为最优分类结果, 对应的共得到两个分区, 相应测点见表 1。

### 3.3.4 分区风压系数代表值计算

根据 3.3.3 节选定的最优分类结果情况 2, 利用公式 (7) 计算得到了对应的两个分区最不利风压系数代表值  $P_1 = 0.923 7$  和  $P_2 = 0.621 2$ 。

表 1 分类情况统计

Table 1 Statistics of classification results

分类数	分类情况	分类结果及相应测点
$k_{st}-1=1$	情况 1	类 1: {G1, G2, G3, G4, G5, H1, H2, H3, H4, H5, I1, I2, I3, I4, I5, J1, J2, J3, J4, J5, K1, K2, K3, K4, K5}
	情况 2	类 1: {G1, G2, G3, G4, G5, H2, H3, H4}
$k_{st}+1=3$	情况 3	类 1: {G2, G3, G4, G5} 类 2: {G1, H2, H3, H4, I2, I3, I4} 类 3: {H1, H5, I1, I5, J1, J2, J3, J4, J5, K1, K2, K3, K4, K5}
	情况 4	类 1: {G2, G3, G4, G5, H3, H4} 类 2: {G1, H1, H2, I2, I3, I4} 类 3: {H5, I1, I5, J1, J2, J3, J4, J5, K1, K2, K3, K4, K5}
	情况 5	类 1: {G2, G3, G4, G5} 类 2: {G1, H2, H3, H4, I3, I4} 类 3: {H1, H5, I1, I2, I5, J1, J2, J3, J4, J5, K1, K2, K3, K4, K5}
6	情况 6	类 1: {G2, G3, G4, G5, H4} 类 2: {G1, H2, H3, I2, I3, I4} 类 3: {H1, H5, I1, I5, J1, J2, J3, J4, J5, K1, K2, K3, K4, K5}

注: 表中各测点位置及编号如图 5 所示

表 2 6 种分类情况对应的  $SD(k)$  值计算结果

Table 2 Calculation results of  $SD(k)$  corresponding to six classification cases

分类情况	$SD(k)$ 值	分类情况	$SD(k)$ 值
情况 1	12.276 1	情况 4	17.461 2
情况 2	9.799 5	情况 5	14.422 3
情况 3	13.574 8	情况 6	14.802 9

按照同样的思路, 分别计算得到了图 6 所示悬挑屋盖单元上表面最不利正、负风压系数和下表面最不利负风压系数分区结果, 分别如图 8 所示。

由图 8 可见, 本文方法计算所得悬挑屋盖单元上、下表面最不利正、负风压系数分区结果较好地吻合了图 6 所示悬挑屋盖单元上、下表面最不利正、负风压系数分布, 结合所得分区最不利正、负风压系数代表值能够实现结构设计阶段简化风荷载施加的目的。

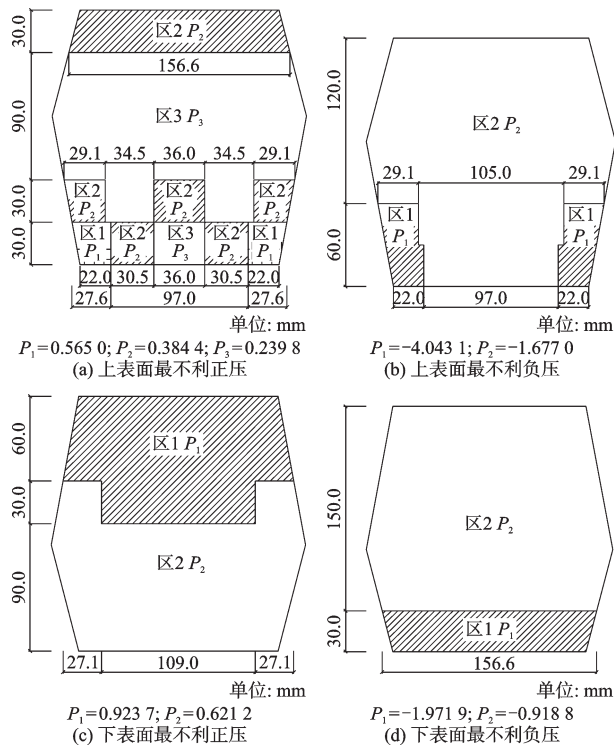


Fig.8 Zoning results of the most unfavorable wind pressure coefficients on the surface of the roof unit

### 3.4 本文方法与传统 K-means 算法对比

为了对比本文所提方法与传统 K-means 算法进行屋盖表面风荷载分区的计算效率,利用传统 K-means 算法对风洞试验所得悬挑屋盖单元上、下表面最不利正、负风压系数进行了分区计算,其中初始分类数  $k$  值按照文献[11]中  $k_{max} \leq \sqrt{n}$  选取,根据悬挑屋盖单元上、下表面测点数量为 30 和 25,  $k_{max}$  值分别取为 5 和 5,相应的初始聚类中心均随机选取,风荷载代表值也采用本文的面积加权法计算而得,其中每个  $k$  值对应的分类情况数目、分类情况总数及最优分区对比结果见表 3。

由表 3 可见,采用传统 K-means 算法和本文方法对算例中悬挑屋盖单元上、下表面最不利正、负风压系数进行分区计算,在保证能得到最优分区结果的前提下,计算过程中本文方法的分类情况总数明显减少,比如差别最大的悬挑屋盖单元下表面最不利正风压系数分区计算,同样得到了图 8(c) 所示最优分区结果。传统算法和本文算法每种分类情况的计算时间是一致的,但传统 K-means 算法需要计算 57 种分类情况,本文方法只需要计算 6 种分类

表 3 本文方法与传统 K-means 算法计算对比

Table 3 Comparison of calculations between the proposed method and the traditional K-means algorithm

风压测点位置及类型	传统 K-means 算法		本文方法	
	$k$ 值	分类情况	$k$ 值	分类情况
	1	1	1	1
正压	2	2	2	2
	3	10	3	6
	4	10	—	—
	5	28	—	—
屋盖单元上表面	1	1	—	—
负压	2	5	2	4
	3	1	3	1
	4	6	4	3
	5	23	—	—
正压	1	1	1	1
	2	1	2	1
	3	7	3	4
	4	16	—	—
屋盖单元下表面	1	1	—	—
负压	2	2	2	2
	3	3	3	1
	4	5	4	3
	5	16	—	—

情况,计算效率得到大幅提升,较好表明了本文方法的优势。

## 4 结 语

本文针对 K-means 算法进行大跨屋盖结构表面风荷载分区中存在的不足,提出基于改进 K-means 算法的大跨屋盖结构表面风荷载分区方法,主要结论有:

(1) 基于建立的分类数  $k$  与其相应测点风荷载的 SSE 关系曲线,通过引入手肘法基本思想,能够有效实现最优分类数  $k_{opt}$  值的精准识别,较好解决了 K-means 算法中分类数  $k$  需凭经验事先给定、缺乏理论依据的问题;

(2) 在首个初始聚类中心随机选取基础上,引入轮盘法基本思想,能够完成对剩余初始聚类中心的高效选取,较好解决了 K-means 算法中所有初始

聚类中心均需随机选取带来的分类情况数过多、从中寻找最优分类结果工作量大且效率低的问题；

(3)根据类内紧凑、类间分散的原则,通过类内紧凑性判定指标  $S(k)$  和类间分散性判定指标  $D(k)$ ,构造并借助  $SD(k)$  值有效性检验,能够有效得到大跨屋盖结构表面风荷载最优分区结果；

(4)以北京奥林匹克公园网球中心悬挑屋盖结构为算例,基于风洞试验所得风荷载测试结果,采用本文所提方法,对其表面最不利风压系数进行分区计算,并与传统 K-means 算法进行对比,说明了所提方法的有效性和工程应用价值。

### 参考文献:

[1] 建筑结构荷载规范:GB 50009—2012[S].北京:中国建筑工业出版社,2012.

[2] 李丹煜,杨庆山,田玉基.基于 K-means 聚类的风压系数快速分区方法[J].工程力学,2014,31(12):164-172. Li D Y, Yang Q S, Tian Y J. A quick zoning method for wind pressure coefficients based on K-means clustering [J]. Engineering Mechanics, 2014, 31(12): 164-172. (in Chinese)

[3] 聂燕松.大跨屋盖结构分区设计风荷载研究[D].哈尔滨:哈尔滨工业大学工程硕士学位论文,2018. Nie Y S. Research on partition design wind load of large-span roof[D]. Harbin: Thesis of Master of Harbin Institute of Technology, 2018. (in Chinese)

[4] 孙晓颖,武岳,沈世钊.鞍形屋盖平均风压分布特性的数值模拟研究[J].工程力学,2006,23(10):7-14. Sun X Y, Wu Y, Shen S Z. Numerical simulation of mean wind pressure distribution on saddle roof[J]. Engineering Mechanics, 2006, 23(10): 7-14. (in Chinese)

[5] 顾明,李方慧,潘钧俊,等.不同风场下开、闭鞍形屋盖脉动风压特性分析[J].同济大学学报(自然科学版),2010,38(7):969-975. Gu M, Li F H, Pan J J, et al. Fluctuating wind pressure characteristics on saddle roof with and without enclosed wall under different terrain conditions [J]. Journal of Tongji University (Natural Science), 2010, 38(7): 969-975. (in Chinese)

[6] Pagnini L, Torre S, Freda A, et al. Wind pressure measurements on a vaulted canopy roof[J]. Journal of Wind Engineering & Industrial Aerodynamics, 2022, 223: 104934.

[7] 李云,张先龙,刘洋,等.屋盖结构分区-多目标等效静风荷载方法研究[J].钢结构(中英文),2022,37(2):

22-29.

Li Y, Zhang X L, Liu Y, et al. Research on partition-multi-objective equivalent static wind load method for roof structure [J]. Steel Construction (Chinese & English), 2022, 37(2): 22-29. (in Chinese)

- [8] Alrawashdeh H, Stathopoulos T. Wind pressures on large roofs of low buildings and wind codes and standards [J]. Journal of Wind Engineering & Industrial Aerodynamics, 2015, 147: 212-225.
- [9] Cao J, Tamura Y, Yoshida A. Wind pressures on multi-level flat roofs of medium-rise buildings [J]. Journal of Wind Engineering & Industrial Aerodynamics, 2012, 103(1): 1-15.
- [10] Gavanski E, Nishimura H. Wind loads on multi-span roof buildings [J]. Journal of Wind Engineering & Industrial Aerodynamics, 2022, 220: 104824.
- [11] Pal N R, Bezdek J C. On cluster validity for the fuzzy c-means model [J]. IEEE Transactions on Fuzzy Systems, 1995, 3(3): 370-379.
- [12] 范九伦,裴继红,谢维信.模糊相关度与聚类有效性[J].西安电子科技大学学报,1998,25(1):13-16. Fan J L, Pei J H, Xie W X. Fuzzy relation degree and cluster validity [J]. Journal of Xidian University, 1998, 25(1): 13-16. (in Chinese)
- [13] 刘爱国,沈一逸,侯显国,等.基于场景法的有源配电网动态无功优化[J].广东电力,2021,34(9):18-26. Liu A G, Shen Y Y, Hou X G, et al. Dynamic reactive power optimization of distribution network with DGs based on scenario method [J]. Guangdong Electric Power, 2021, 34(9): 18-26. (in Chinese)
- [14] 吴广建,章剑林,袁丁.基于 K-means 的手肘法自动获取 K 值方法研究[J].软件,2019,40(5):167-170. Wu G J, Zhang J L, Yuan D. Automatically obtaining K value based on K-means elbow method [J]. Computer Engineering & Software, 2019, 40(5): 167-170. (in Chinese)
- [15] 马洁莹.基于轮盘赌策略的混沌萤火虫算法研究[D].西安:西安电子科技大学硕士学位论文,2018. Ma J Y. Research on chaotic firefly algorithm based on roulette wheel selection strategy [D]. Xi'an: Thesis of Master of Xidian University, 2018. (in Chinese)
- [16] Halkidi M, Vazirgiannis M, Batistakis Y. Quality scheme assessment in the clustering process [J]. Lecture Notes in Computer Science Volume, 2000, 1910(4): 265-276.
- [17] 张相庭.结构风工程理论·规范·实践[M].北京:建筑工业出版社,2006:99-118.

(本文编辑:张艺)