Algorithm1: Pseudo-code of the improved SC algorithm

1: Input: the sample set $X=\{x_1,x_2,\cdots x_n\}$, the number of prototype clusters p, the number of clusters k.

2. Using k-means++ to acquire prototype sample set $U=\{u_1,u_2,\cdots u_p\}$

3: Calculate the similarity between the prototype points and the original points by $z_{ij}=\exp(-\frac{\|x_i-u_j\|_2^2}{2\sigma^2})$ and $W'=ZZ^{\mathrm{T}}$ and construct the new adjacency matrix W'.

4: Calculate the degree matrix D' by $D=diag\{d_1,\cdots,d_n\}$ and $d_i=\sum_{j=1}^{n}w_{ij}$.

5: Calculate the symmetric Laplacian matrix L'_{N} by equation $L'_N=D'^{-\frac{1}{2}}L'D'^{-\frac{1}{2}}=I-D'^{-\frac{1}{2}}W'D'^{-\frac{1}{2}}$.

6: Perform a k-means clustering on the row vectors after arranging the first k feature vectors by column.

7: Map the clustering result back to the original solution space.

8: Output: Cluster clustering results $C=\{C_1,C_2,\cdots,C_k\}$.

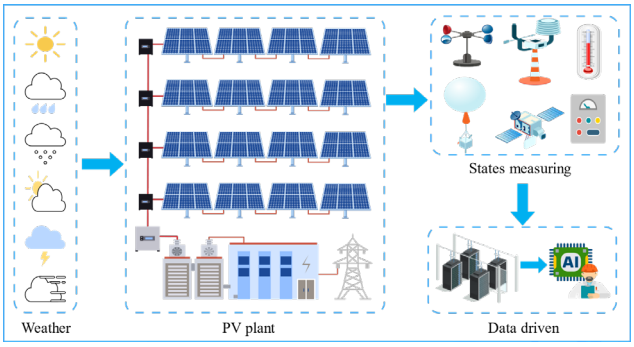

Fig. A1. Schematic diagram of the clustering of the PV plants

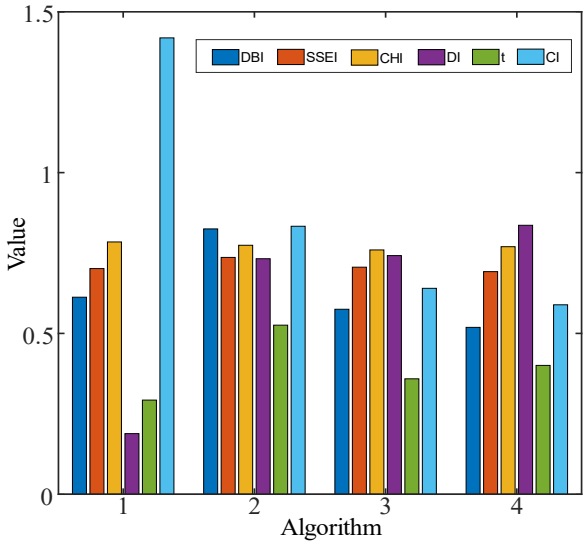

Fig. A2. Comparison of clustering indices of each algorithm