

Supplementary Material

This file includes figures, that provide the same information as those in the main text but for different runs (i.e., initializations).

Results for the WP test

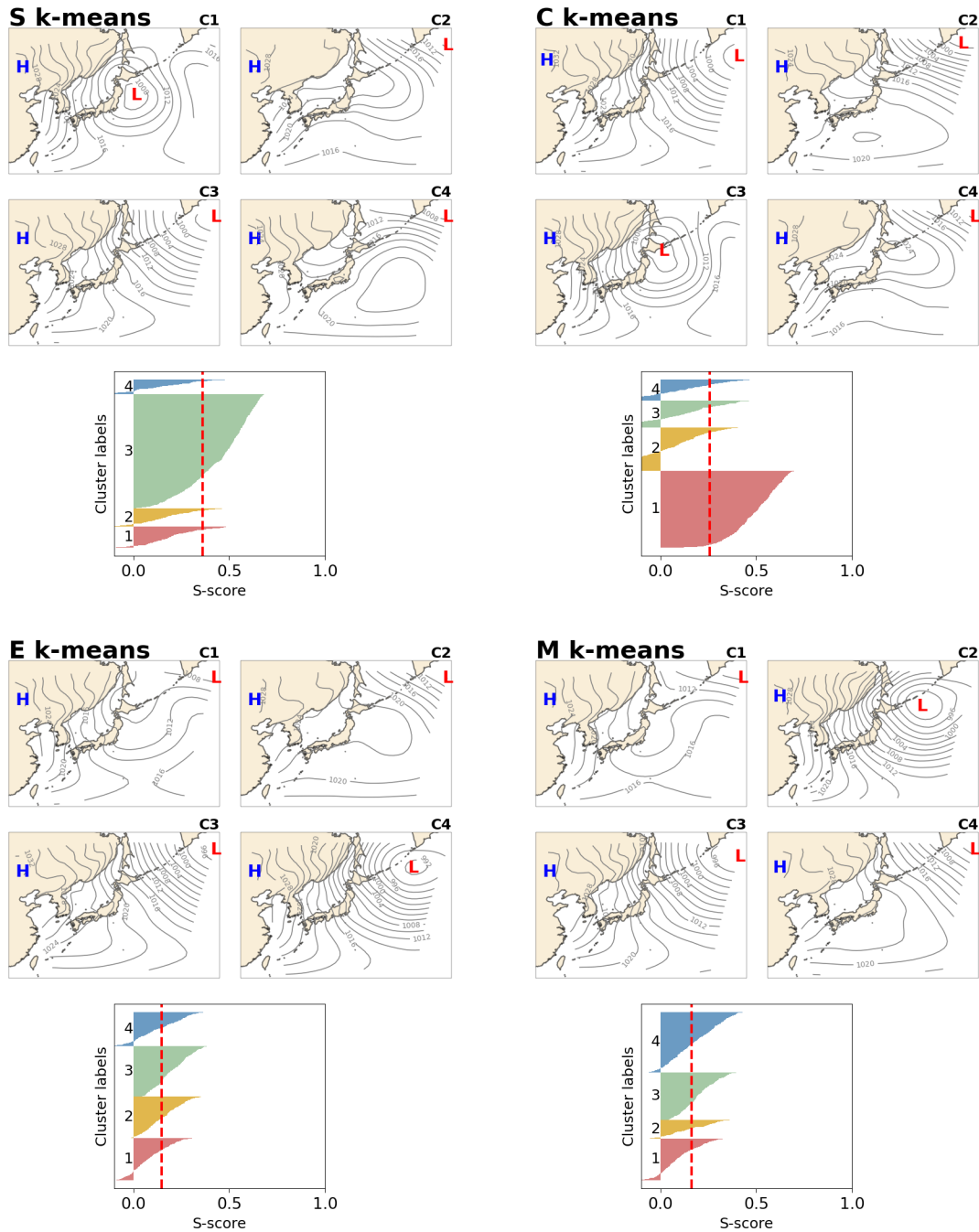


Fig. S1 Similar with Figure 3 in the main manuscript but for the results from run 1. The winter SLP pattern revealed by S, C, E, and M k-means with $k = 4$. "H" indicates the location of the high and "L" indicates the location of the low. General silhouette analysis results are shown below the maps where x-axis indicate the score, and y-axis the label of cluster numbered 1 – 4. Input data are ERA-Interim SLP data, which were re-gridded to Cartesian coordinates with a resolution of 200×200 km and grid size of 35×35 . Daily data for December, January, and February over ten year 2005 – 2014 were used.

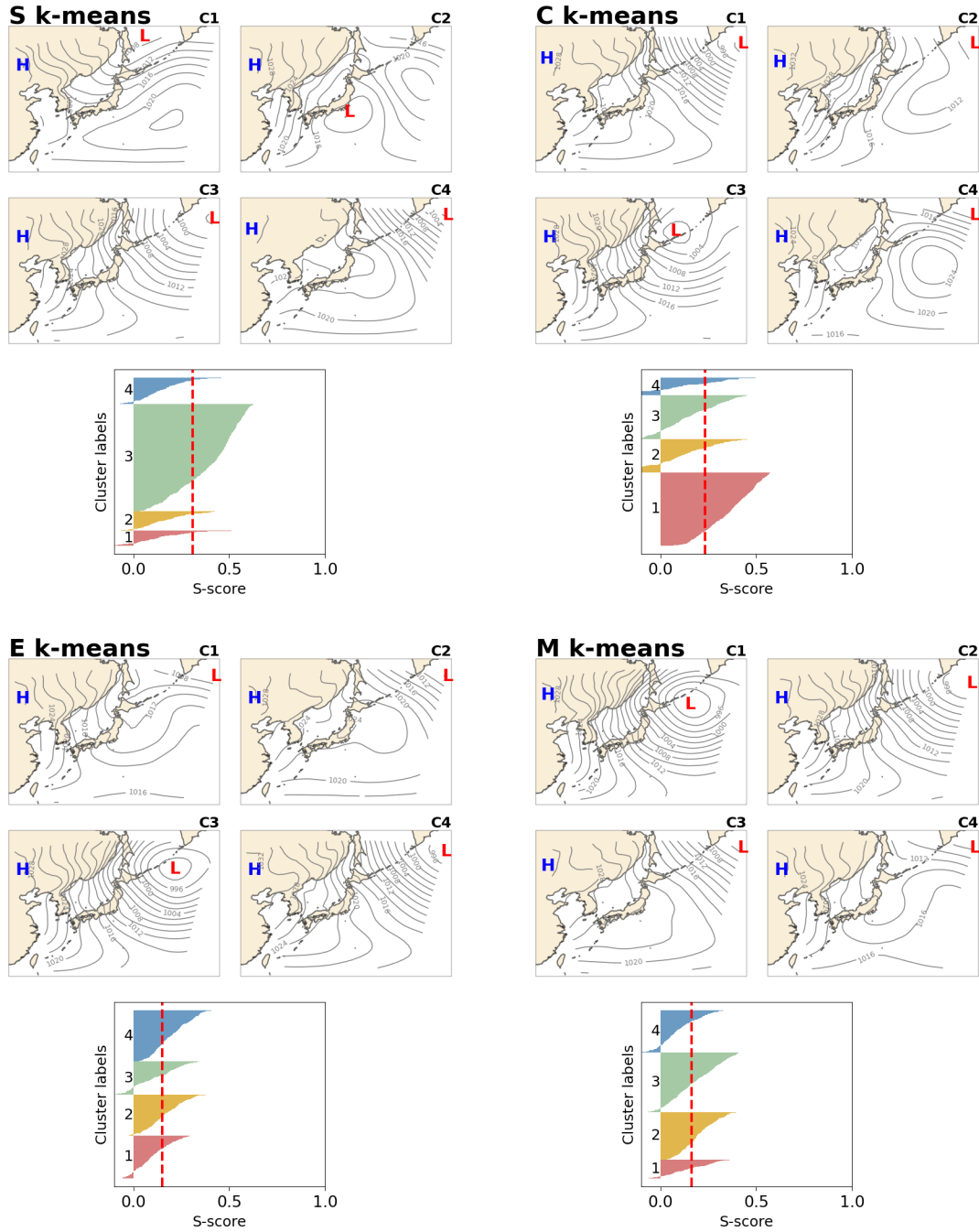


Fig. S2 Similar with Figure 3 in the main manuscript but for the results from run 2. The winter SLP pattern revealed by S, C, E, and M k-means with $k = 4$. "H" indicates the location of the high and "L" indicates the location of the low. General silhouette analysis results are shown below the maps where x-axis indicate the score, and y-axis the label of cluster numbered 1 – 4. Input data are ERA-Interim SLP data, which were re-gridded to Cartesian coordinates with a resolution of 200×200 km and grid size of 35×35 . Daily data for December, January, and February over ten year 2005 – 2014 were used.

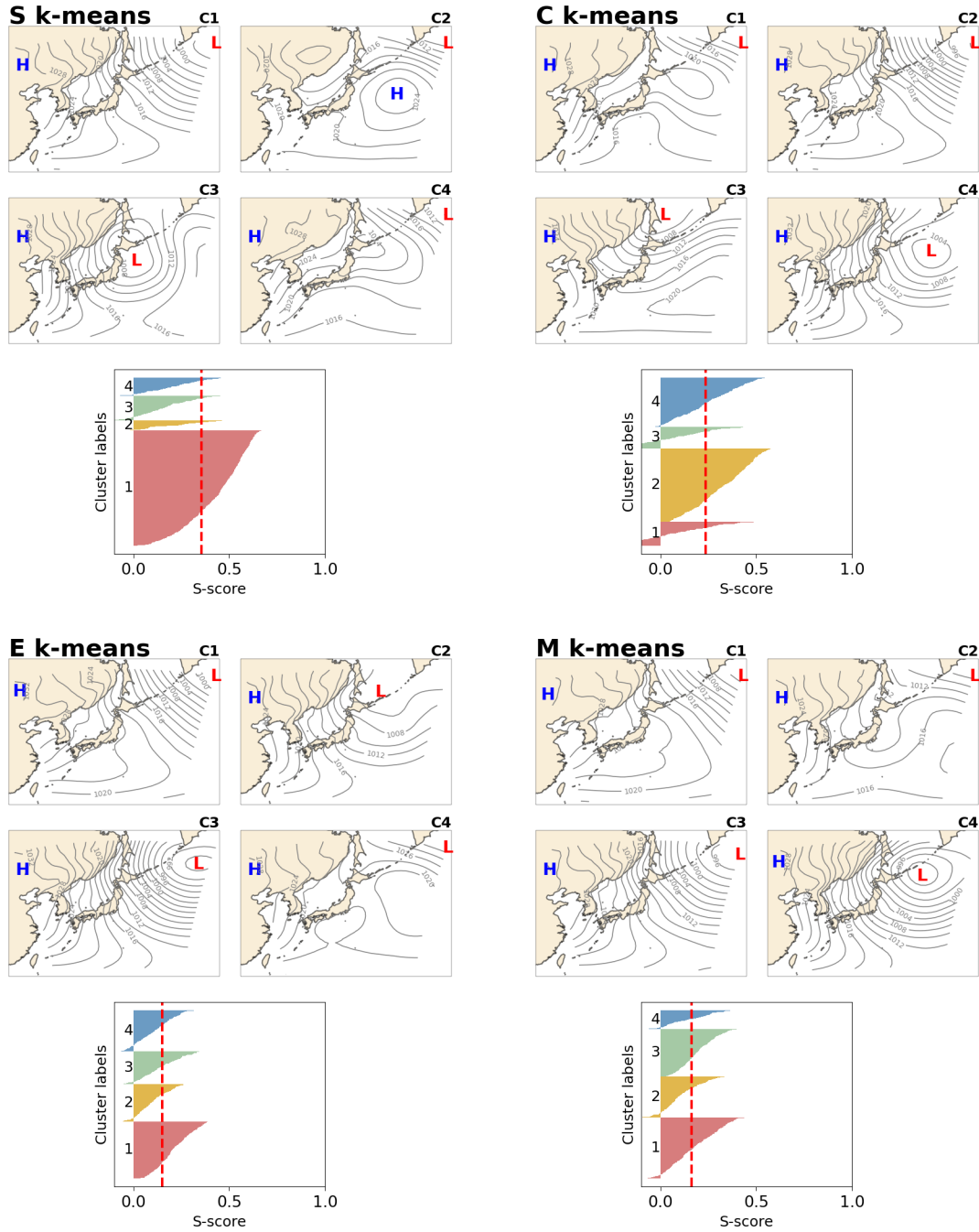


Fig. S3 Similar with Figure 3 in the main manuscript but for the results from run 3. The winter SLP pattern revealed by S, C, E, and M k-means with $k = 4$. "H" indicates the location of the high and "L" indicates the location of the low. General silhouette analysis results are shown below the maps where x-axis indicate the score, and y-axis the label of cluster numbered 1 – 4. Input data are ERA-Interim SLP data, which were re-gridded to Cartesian coordinates with a resolution of 200×200 km and grid size of 35×35 . Daily data for December, January, and February over ten year 2005 – 2014 were used.

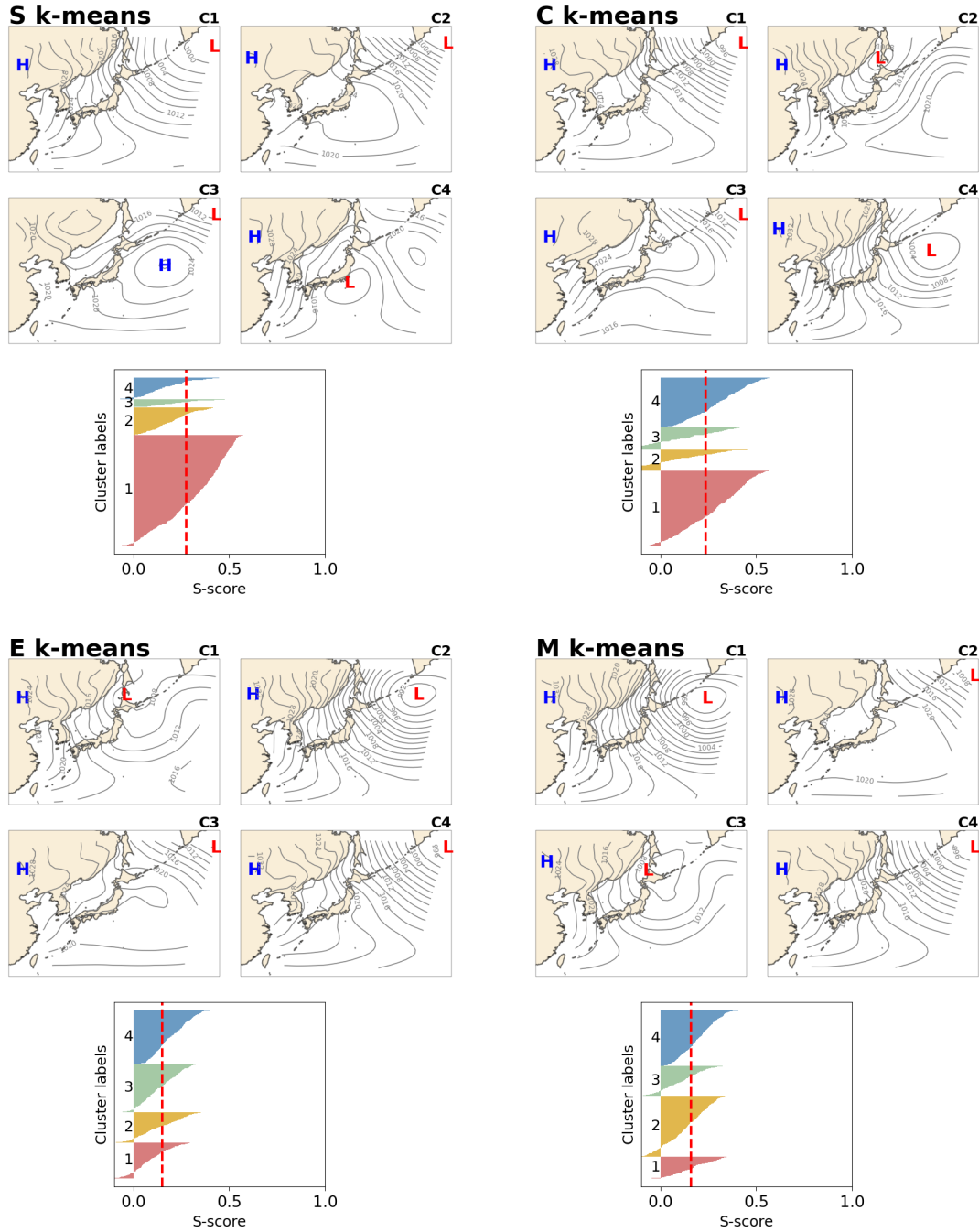


Fig. S4 Similar with Figure 3 in the main manuscript but for the results from run 4. The winter SLP pattern revealed by S, C, E, and M k-means with $k = 4$. "H" indicates the location of the high and "L" indicates the location of the low. General silhouette analysis results are shown below the maps where x-axis indicate the score, and y-axis the label of cluster numbered 1 – 4. Input data are ERA-Interim SLP data, which were re-gridded to Cartesian coordinates with a resolution of 200×200 km and grid size of 35×35 . Daily data for December, January, and February over ten year 2005 – 2014 were used.

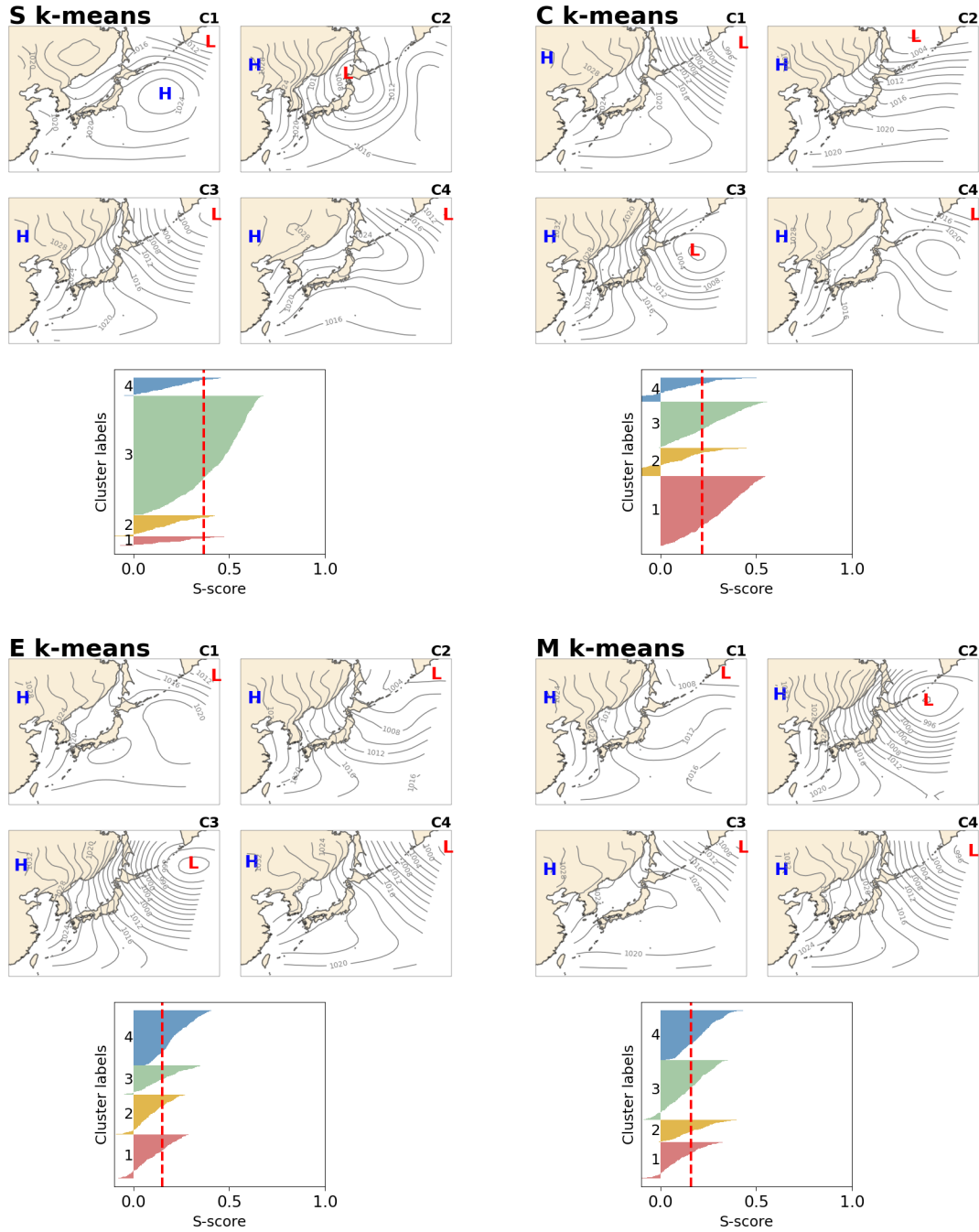


Fig. S5 Similar with Figure 3 in the main manuscript but for the results from run 5. The winter SLP pattern revealed by S, C, E, and M k-means with $k = 4$. "H" indicates the location of the high and "L" indicates the location of the low. General silhouette analysis results are shown below the maps where x-axis indicate the score, and y-axis the label of cluster numbered 1 – 4. Input data are ERA-Interim SLP data, which were re-gridded to Cartesian coordinates with a resolution of 200×200 km and grid size of 35×35 . Daily data for December, January, and February over ten year 2005 – 2014 were used.

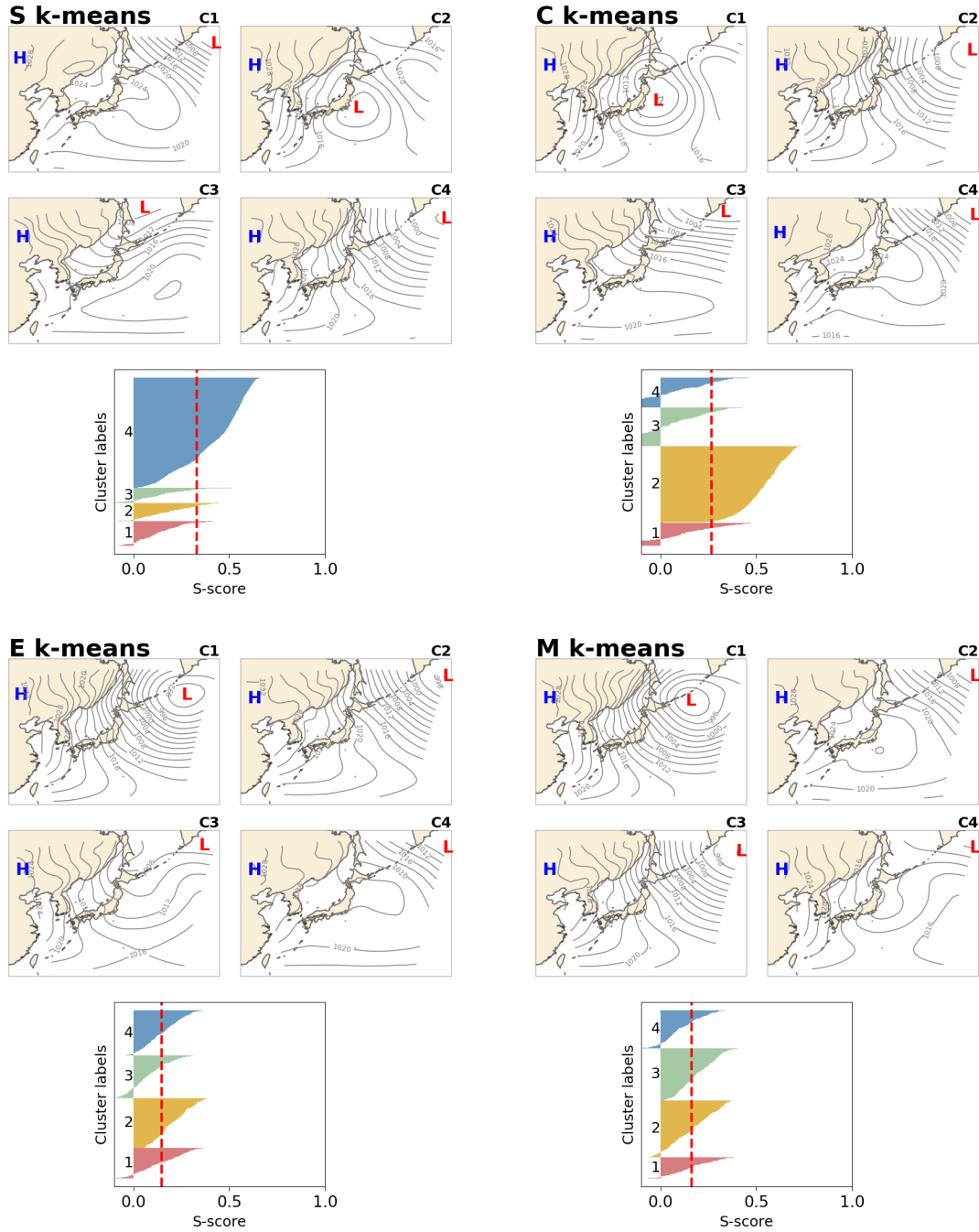


Fig. S6 Similar with Figure 3 in the main manuscript but for the results from run 6. The winter SLP pattern revealed by S, C, E, and M k-means with $k = 4$. "H" indicates the location of the high and "L" indicates the location of the low. General silhouette analysis results are shown below the maps where x-axis indicate the score, and y-axis the label of cluster numbered 1 – 4. Input data are ERA-Interim SLP data, which were re-gridded to Cartesian coordinates with a resolution of 200×200 km and grid size of 35×35 . Daily data for December, January, and February over ten year 2005 – 2014 were used.

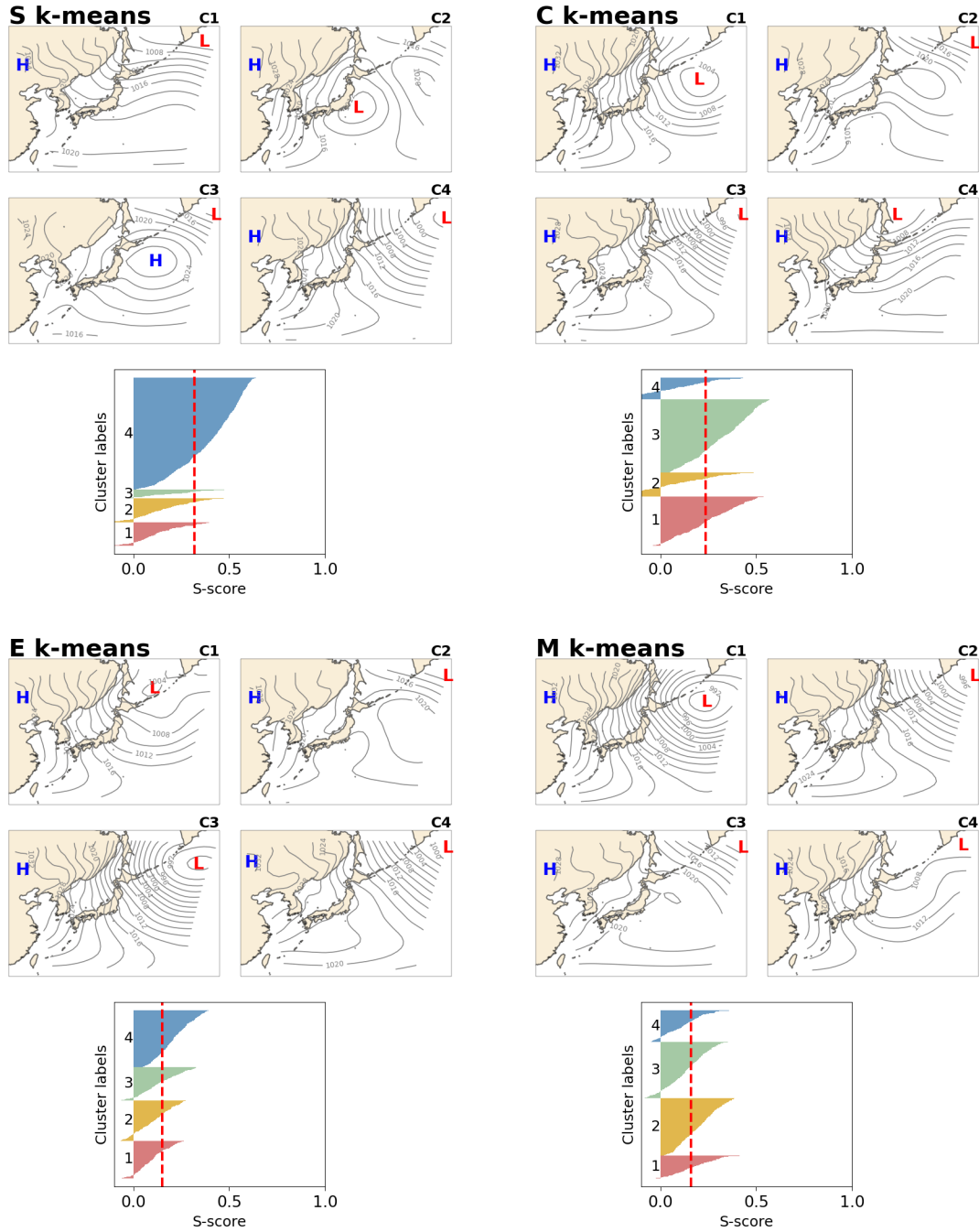


Fig. S7 Similar with Figure 3 in the main manuscript but for the results from run 7. The winter SLP pattern revealed by S, C, E, and M k-means with $k = 4$. "H" indicates the location of the high and "L" indicates the location of the low. General silhouette analysis results are shown below the maps where x-axis indicate the score, and y-axis the label of cluster numbered 1 – 4. Input data are ERA-Interim SLP data, which were re-gridded to Cartesian coordinates with a resolution of 200×200 km and grid size of 35×35 . Daily data for December, January, and February over ten year 2005 – 2014 were used.

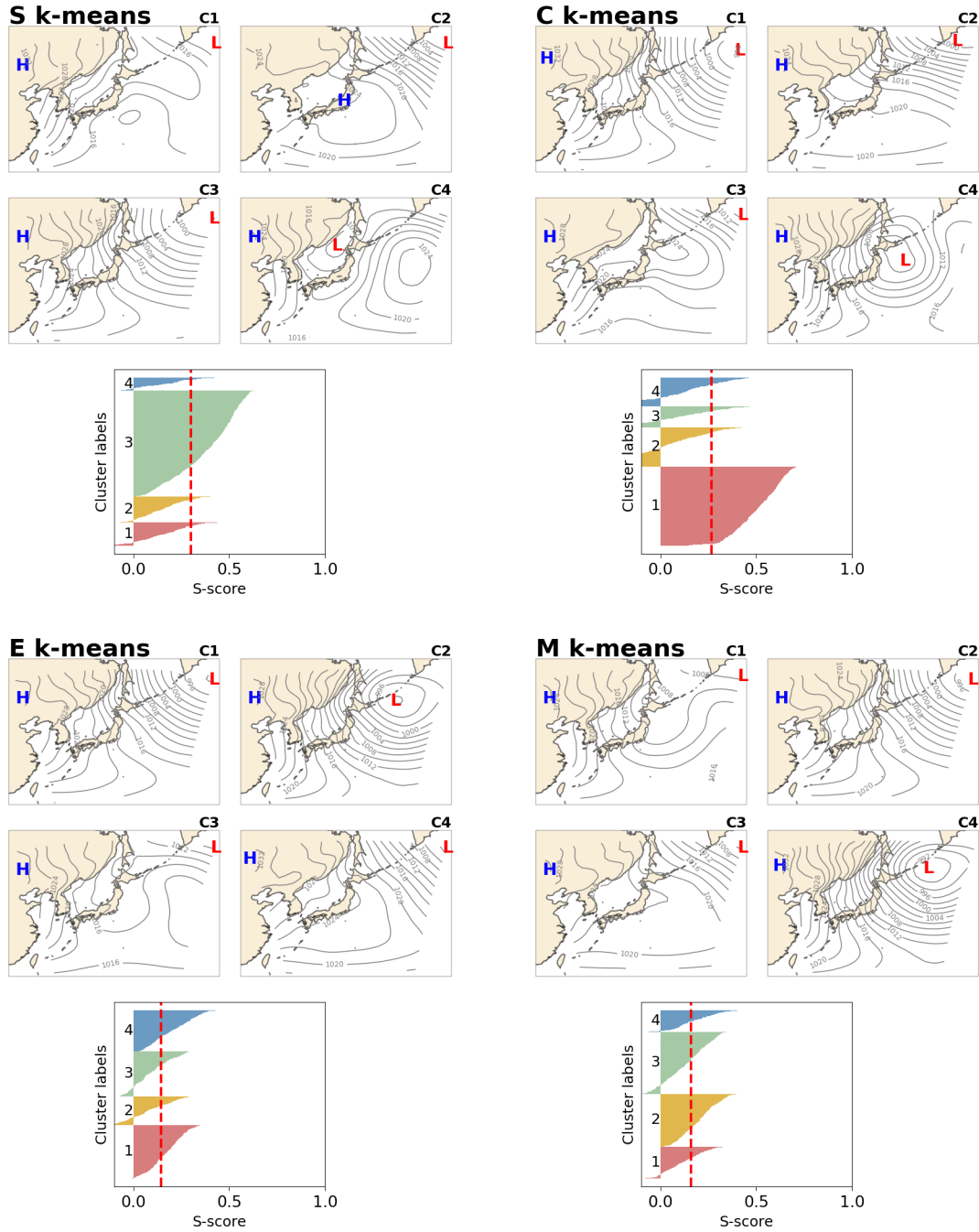


Fig. S8 Similar with Figure 3 in the main manuscript but for the results from run 8. The winter SLP pattern revealed by S, C, E, and M k-means with $k = 4$. "H" indicates the location of the high and "L" indicates the location of the low. General silhouette analysis results are shown below the maps where x-axis indicate the score, and y-axis the label of cluster numbered 1 – 4. Input data are ERA-Interim SLP data, which were re-gridded to Cartesian coordinates with a resolution of 200×200 km and grid size of 35×35 . Daily data for December, January, and February over ten year 2005 – 2014 were used.

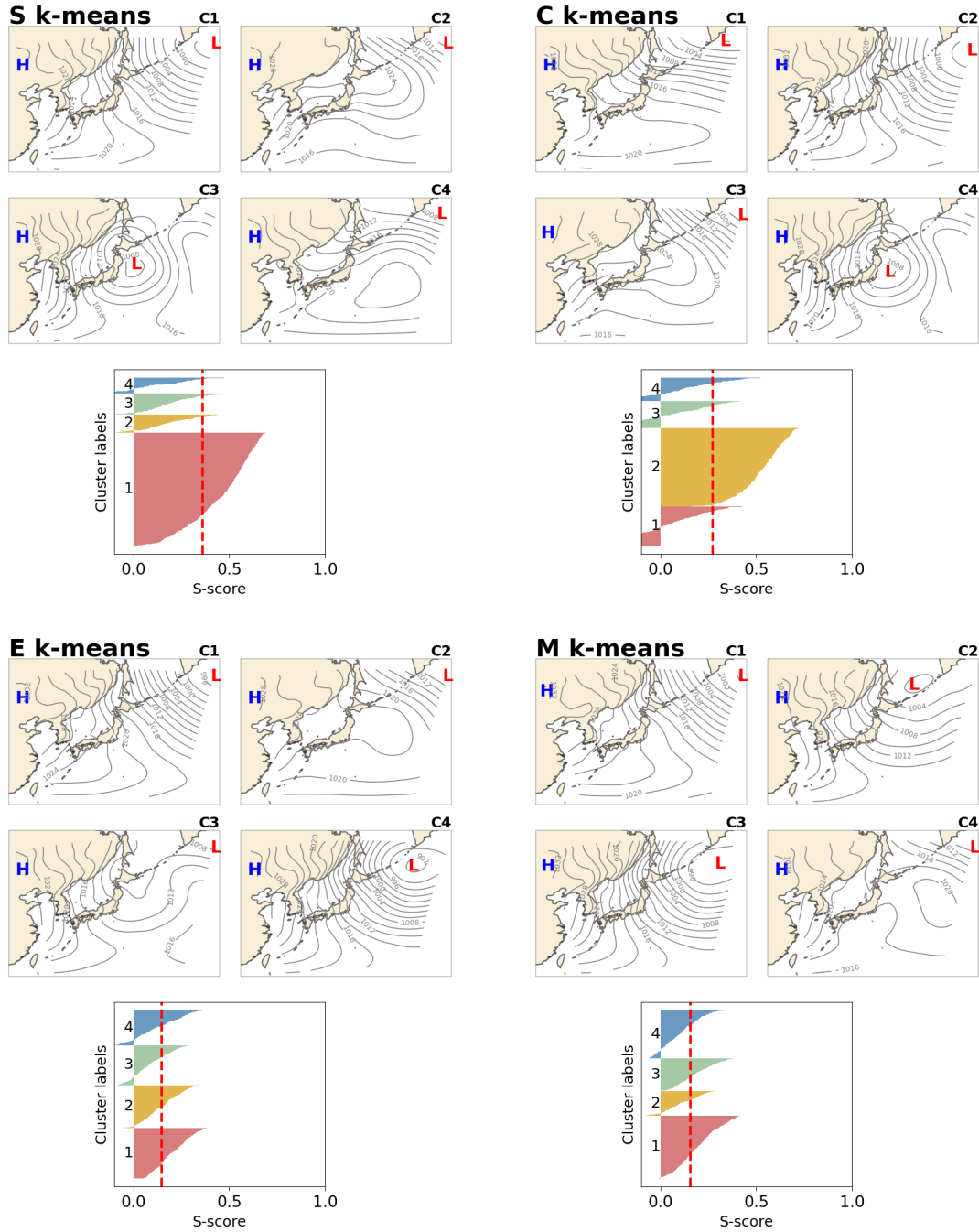


Fig. S9 Similar with Figure 3 in the main manuscript but for the results from run 9. The winter SLP pattern revealed by S, C, E, and M k-means with $k = 4$. "H" indicates the location of the high and "L" indicates the location of the low. General silhouette analysis results are shown below the maps where x-axis indicate the score, and y-axis the label of cluster numbered 1 – 4. Input data are ERA-Interim SLP data, which were re-gridded to Cartesian coordinates with a resolution of 200×200 km and grid size of 35×35 . Daily data for December, January, and February over ten year 2005 – 2014 were used.

Results for the CC test

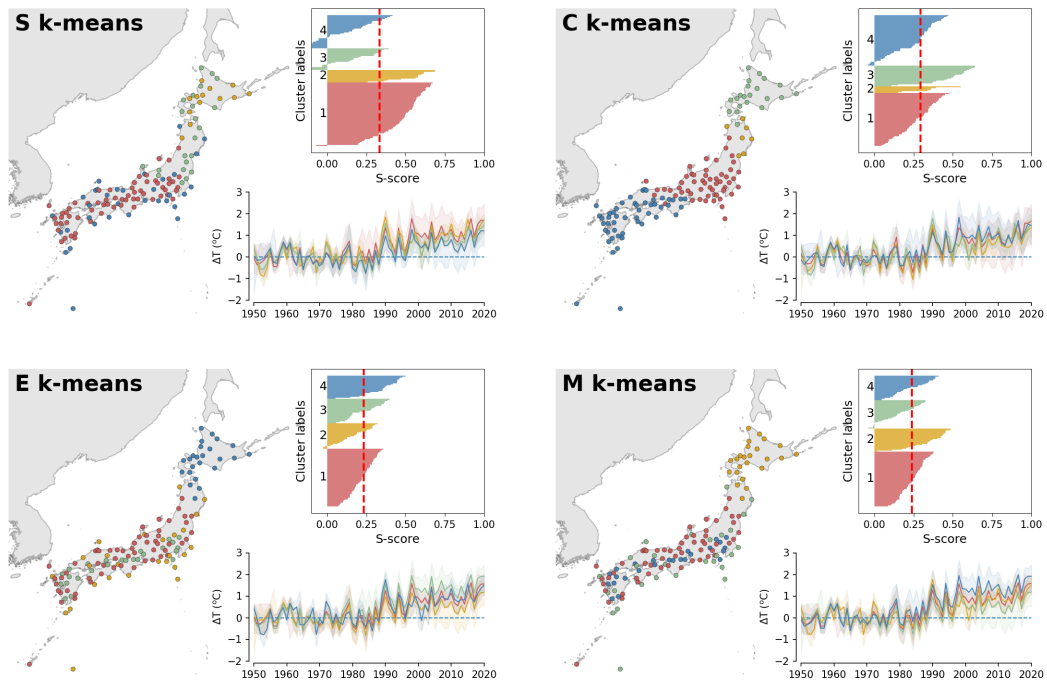


Fig. S10 Similar with Figure 4 in the main manuscript but for the results from run 1. Result for CC experiment for cluster the climate change (temperature increase) time series over 134 weather stations over the entirety of Japan. Pattern were revealed by S, C, E, and M k-means with $k = 4$. Input data correspond to annual mean data collected over 70 years from 1951 – 2020 (subtracted by the mean of the first 30 years) and observed temperature achieved at in situ weather stations (dots in map) operated by the JMA. Time series of centroids and input vectors are shown in below panels together with general silhouette analysis results, where the x-axis indicates the score (S-score) and the y-axis presents the labels of clusters numbered 1 – 4.

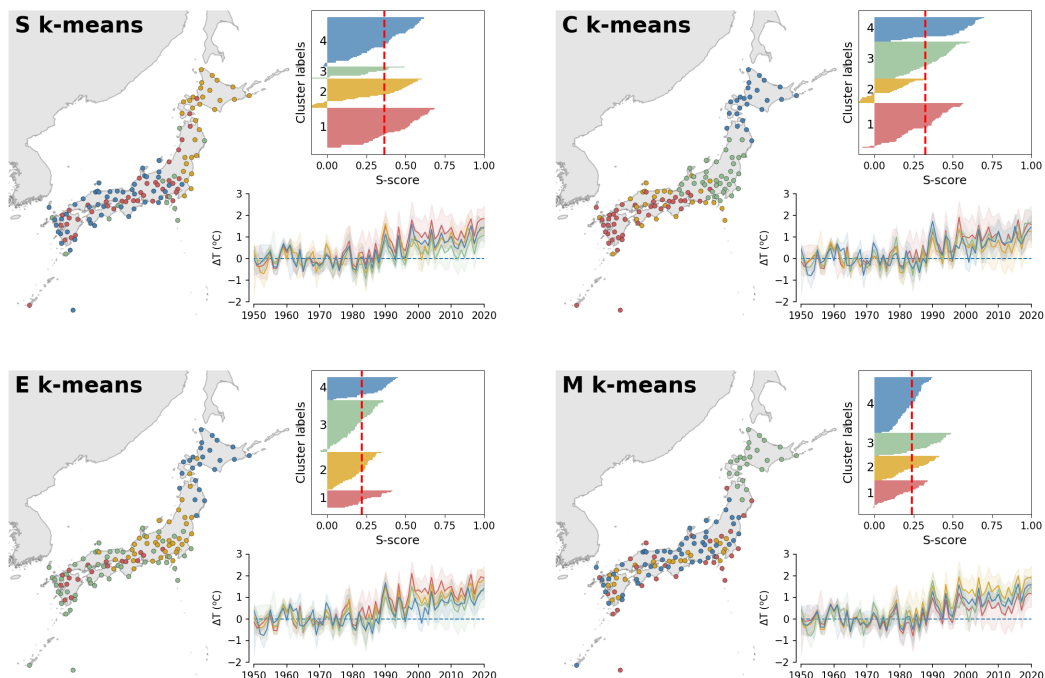


Fig. S11 Similar with Figure 4 in the main manuscript but for the results from run 2. Result for CC experiment for cluster the climate change (temperature increase) time series over 134 weather stations over the entirety of Japan. Pattern were revealed by S, C, E, and M k-means with $k = 4$. Input data correspond to annual mean data collected over 70 years from 1951 – 2020 (subtracted by the mean of the first 30 years) and observed temperature achieved at in situ weather stations (dots in map) operated by the JMA. Time series of centroids and input vectors are shown in below panels together with general silhouette analysis results, where the x-axis indicates the score (S-score) and the y-axis presents the labels of clusters numbered 1 – 4.

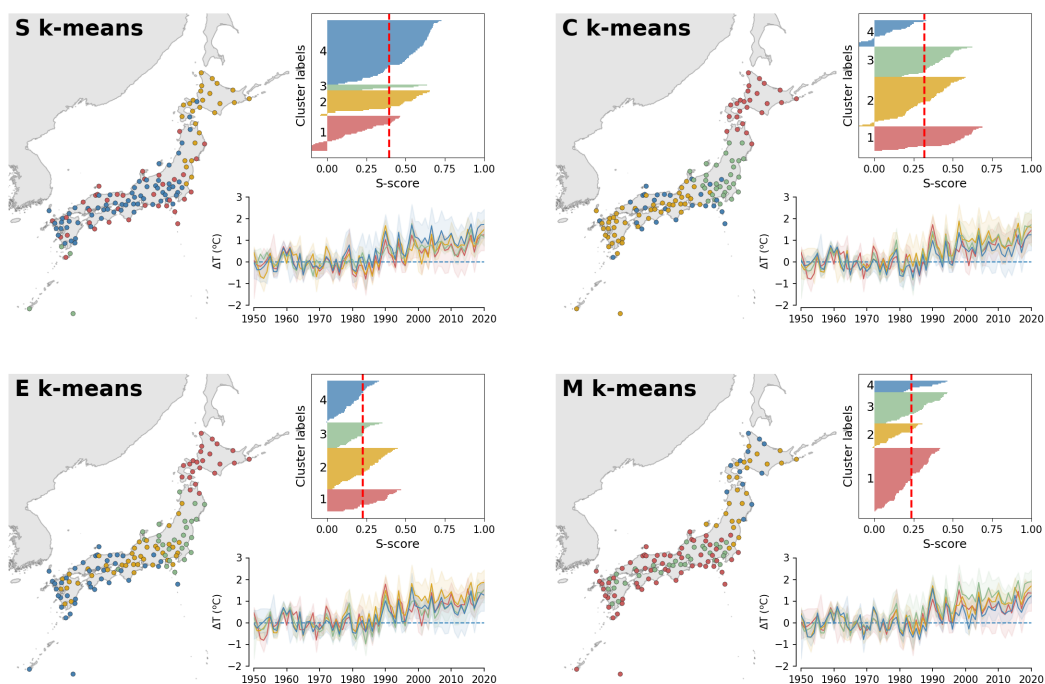


Fig. S12 Similar with Figure 4 in the main manuscript but for the results from run 3. Result for CC experiment for cluster the climate change (temperature increase) time series over 134 weather stations over the entirety of Japan. Pattern were revealed by S, C, E, and M k-means with $k = 4$. Input data correspond to annual mean data collected over 70 years from 1951 – 2020 (subtracted by the mean of the first 30 years) and observed temperature achieved at in situ weather stations (dots in map) operated by the JMA. Time series of centroids and input vectors are shown in below panels together with general silhouette analysis results, where the x-axis indicates the score (S-score) and the y-axis presents the labels of clusters numbered 1 – 4.

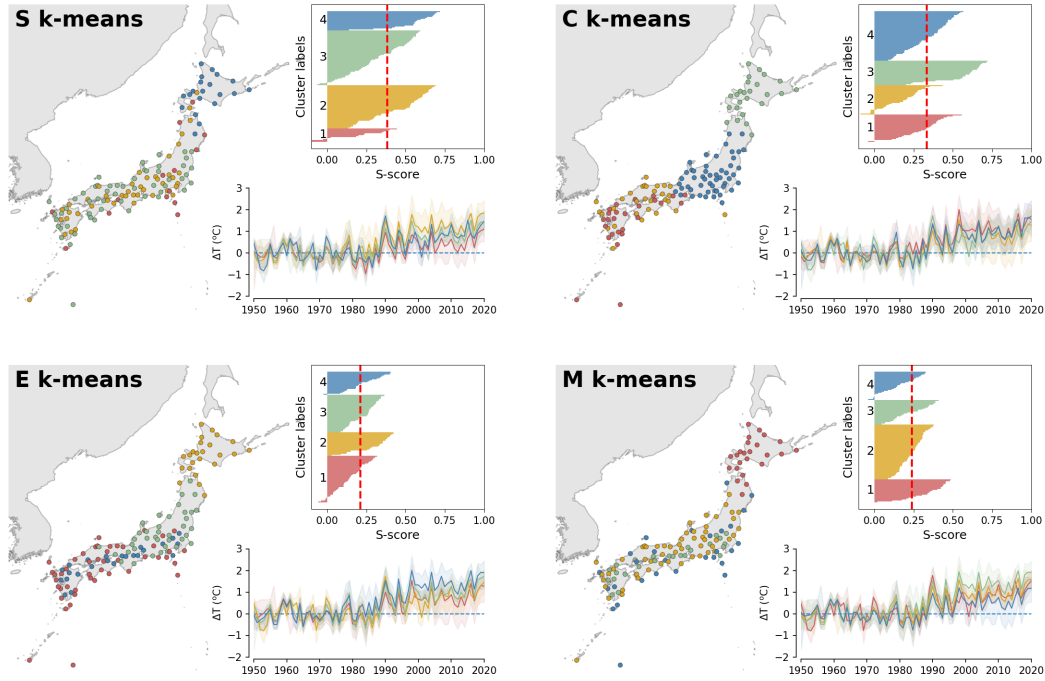


Fig. S13 Similar with Figure 4 in the main manuscript but for the results from run 4. Result for CC experiment for cluster the climate change (temperature increase) time series over 134 weather stations over the entirety of Japan. Pattern were revealed by S, C, E, and M k-means with $k = 4$. Input data correspond to annual mean data collected over 70 years from 1951 – 2020 (subtracted by the mean of the first 30 years) and observed temperature achieved at in situ weather stations (dots in map) operated by the JMA. Time series of centroids and input vectors are shown in below panels together with general silhouette analysis results, where the x-axis indicates the score (S-score) and the y-axis presents the labels of clusters numbered 1 – 4.

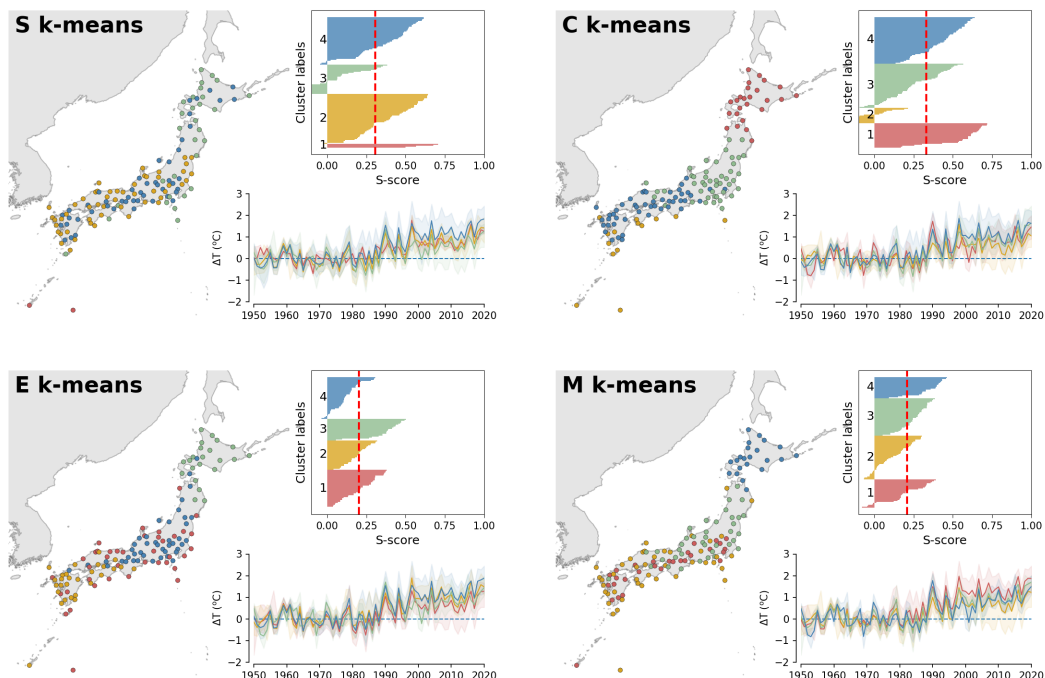


Fig. S14 Similar with Figure 4 in the main manuscript but for the results from run 5. Result for CC experiment for cluster the climate change (temperature increase) time series over 134 weather stations over the entirety of Japan. Pattern were revealed by S, C, E, and M k-means with $k = 4$. Input data correspond to annual mean data collected over 70 years from 1951 – 2020 (subtracted by the mean of the first 30 years) and observed temperature achieved at in situ weather stations (dots in map) operated by the JMA. Time series of centroids and input vectors are shown in below panels together with general silhouette analysis results, where the x-axis indicates the score (S-score) and the y-axis presents the labels of clusters numbered 1 – 4.

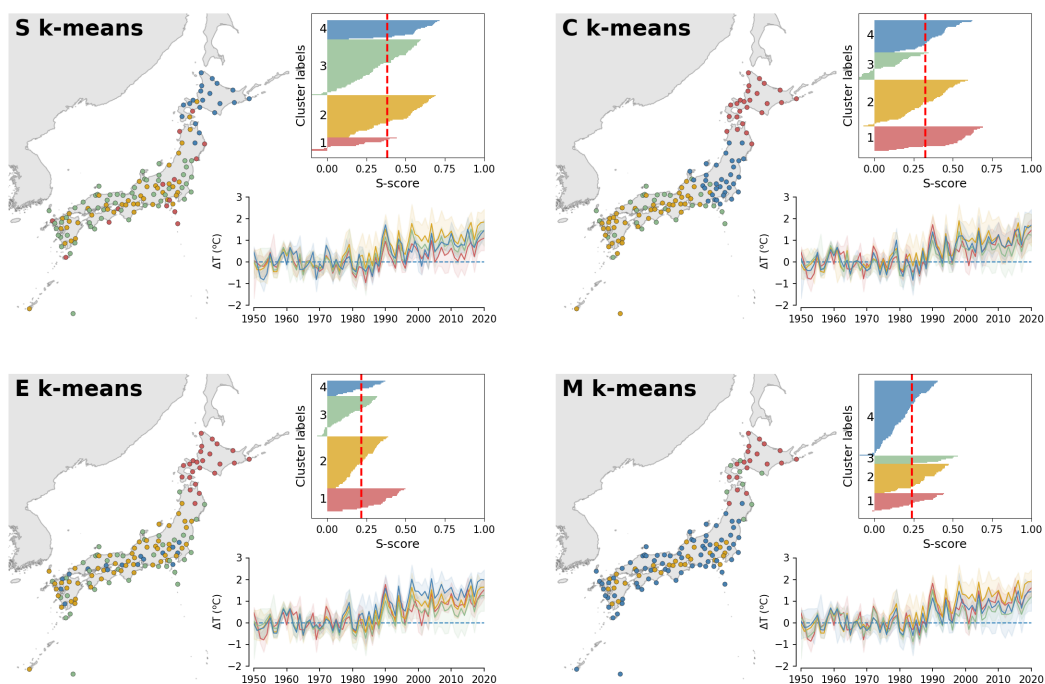


Fig. S15 Similar with Figure 4 in the main manuscript but for the results from run 6. Result for CC experiment for cluster the climate change (temperature increase) time series over 134 weather stations over the entirety of Japan. Pattern were revealed by S, C, E, and M k-means with $k = 4$. Input data correspond to annual mean data collected over 70 years from 1951 – 2020 (subtracted by the mean of the first 30 years) and observed temperature achieved at in situ weather stations (dots in map) operated by the JMA. Time series of centroids and input vectors are shown in below panels together with general silhouette analysis results, where the x-axis indicates the score (S-score) and the y-axis presents the labels of clusters numbered 1 – 4.

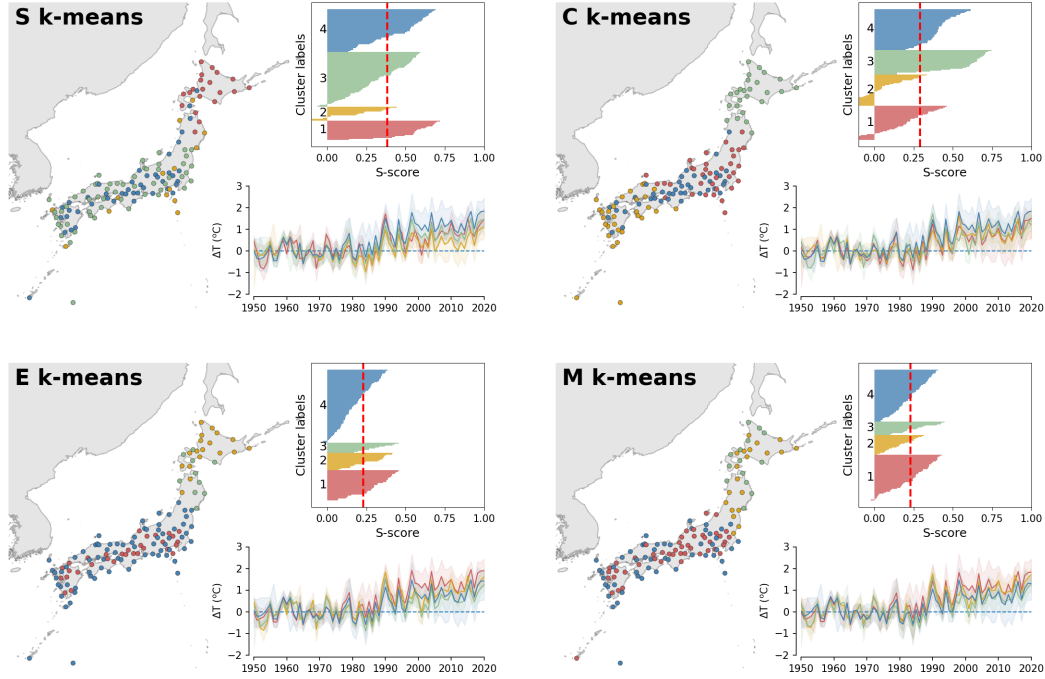


Fig. S16 Similar with Figure 4 in the main manuscript but for the results from run 7. Result for CC experiment for cluster the climate change (temperature increase) time series over 134 weather stations over the entirety of Japan. Pattern were revealed by S, C, E, and M k-means with $k = 4$. Input data correspond to annual mean data collected over 70 years from 1951 – 2020 (subtracted by the mean of the first 30 years) and observed temperature achieved at in situ weather stations (dots in map) operated by the JMA. Time series of centroids and input vectors are shown in below panels together with general silhouette analysis results, where the x-axis indicates the score (S-score) and the y-axis presents the labels of clusters numbered 1 – 4.

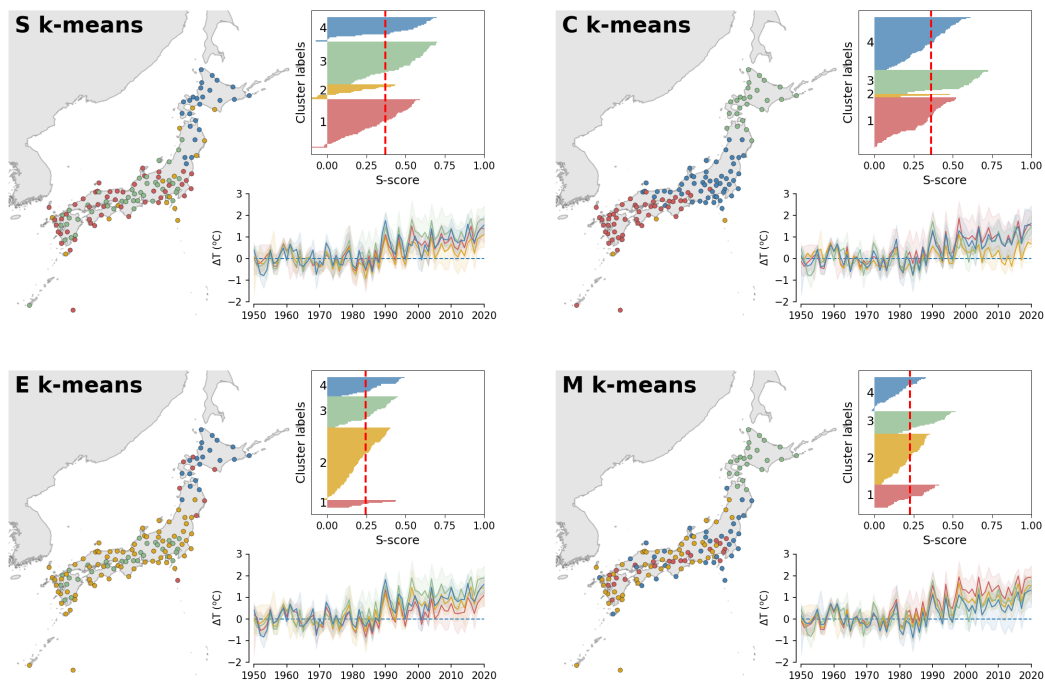


Fig. S17 Similar with Figure 4 in the main manuscript but for the results from run 8. Result for CC experiment for cluster the climate change (temperature increase) time series over 134 weather stations over the entirety of Japan. Pattern were revealed by S, C, E, and M k-means with $k = 4$. Input data correspond to annual mean data collected over 70 years from 1951 – 2020 (subtracted by the mean of the first 30 years) and observed temperature achieved at in situ weather stations (dots in map) operated by the JMA. Time series of centroids and input vectors are shown in below panels together with general silhouette analysis results, where the x-axis indicates the score (S-score) and the y-axis presents the labels of clusters numbered 1 – 4.

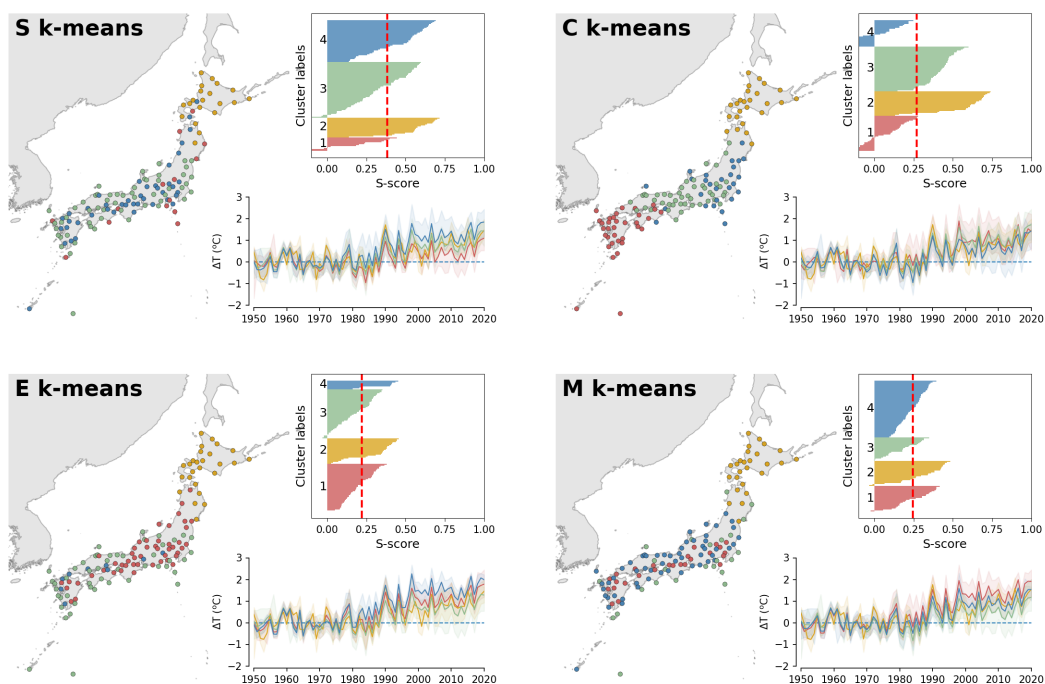


Fig. S18 Similar with Figure 4 in the main manuscript but for the results from run 9. Result for CC experiment for cluster the climate change (temperature increase) time series over 134 weather stations over the entirety of Japan. Pattern were revealed by S, C, E, and M k-means with $k = 4$. Input data correspond to annual mean data collected over 70 years from 1951 – 2020 (subtracted by the mean of the first 30 years) and observed temperature achieved at in situ weather stations (dots in map) operated by the JMA. Time series of centroids and input vectors are shown in below panels together with general silhouette analysis results, where the x-axis indicates the score (S-score) and the y-axis presents the labels of clusters numbered 1 – 4.

Results for the TC test

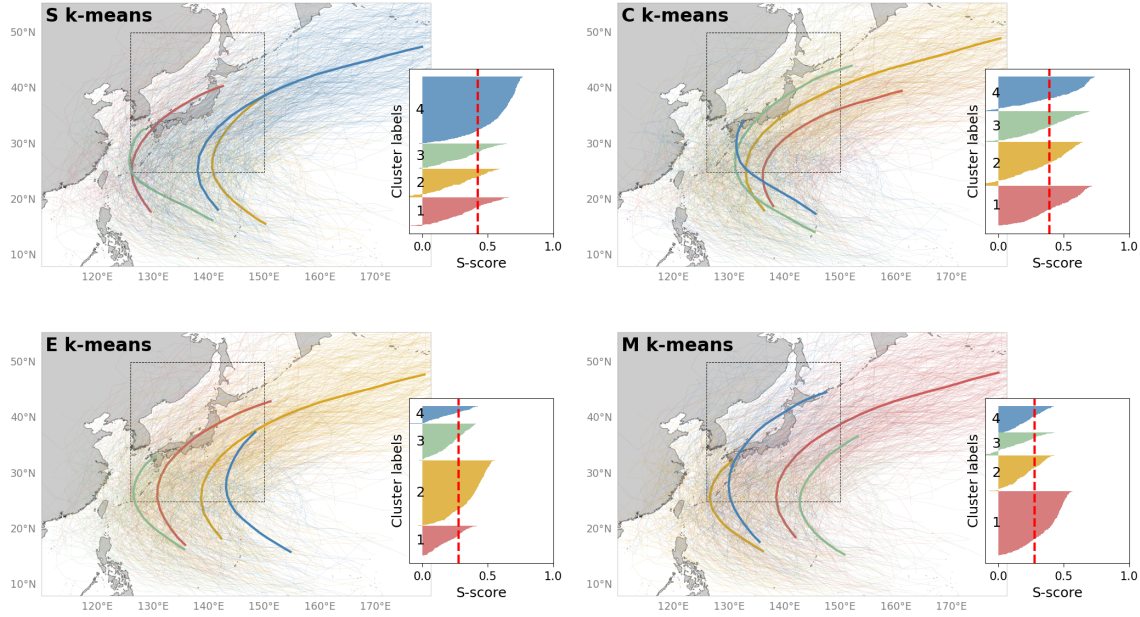


Fig. S19 Similar with Figure 5 in the maintext but for run 1. Results of the TC experiment for clustering tropical cyclone paths. The pattern was revealed by S, C, E, and M k-means, with $k = 4$. Input data are the best TC tracks obtained by the JMA from 1951 – 2020. Only TCs that passed the dashed box in the map are used to feed the k-means. Thus, a total of 863 TC tracking data points are used. The left side of each panel show the general silhouette analysis results, where the x-axis indicates the score (S-score) and y-axis presents the labels of clusters numbered 1 – 4. The centroid TC path is illustrated by the bold line, and the color is consistent with that in the silhouette diagram.

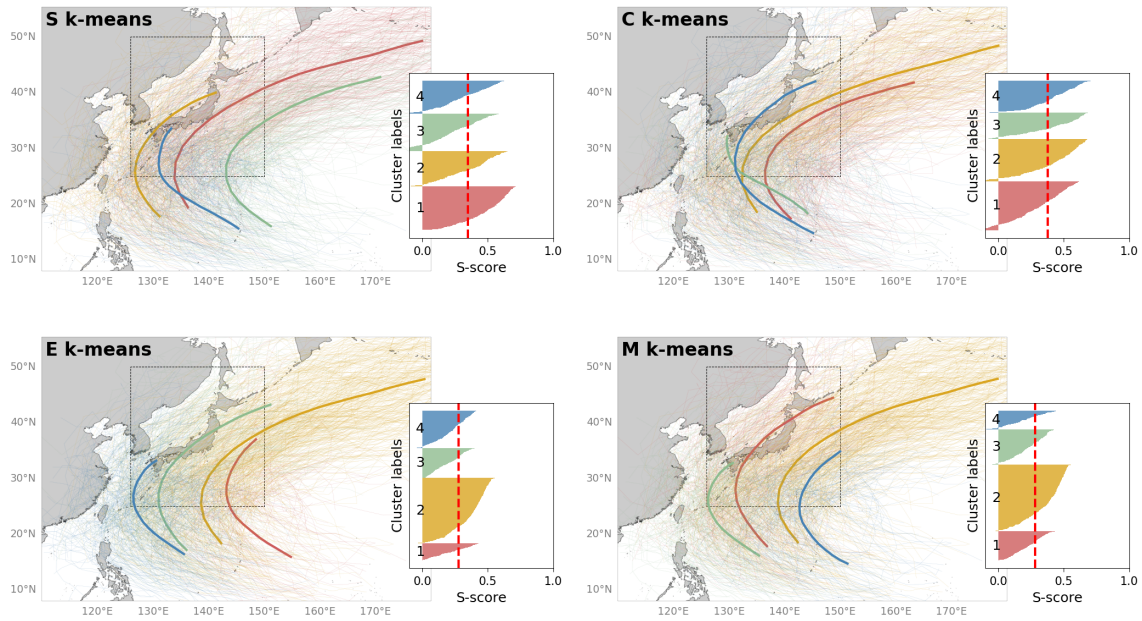


Fig. S20 Similar with Figure 5 in the maintext but for run 2. Results of the TC experiment for clustering tropical cyclone paths. The pattern was revealed by S, C, E, and M k-means, with $k = 4$. Input data are the best TC tracks obtained by the JMA from 1951 – 2020. Only TCs that passed the dashed box in the map are used to feed the k-means. Thus, a total of 863 TC tracking data points are used. The left side of each panel show the general silhouette analysis results, where the x-axis indicates the score (S-score) and y-axis presents the labels of clusters numbered 1 – 4. The centroid TC path is illustrated by the bold line, and the color is consistent with that in the silhouette diagram.

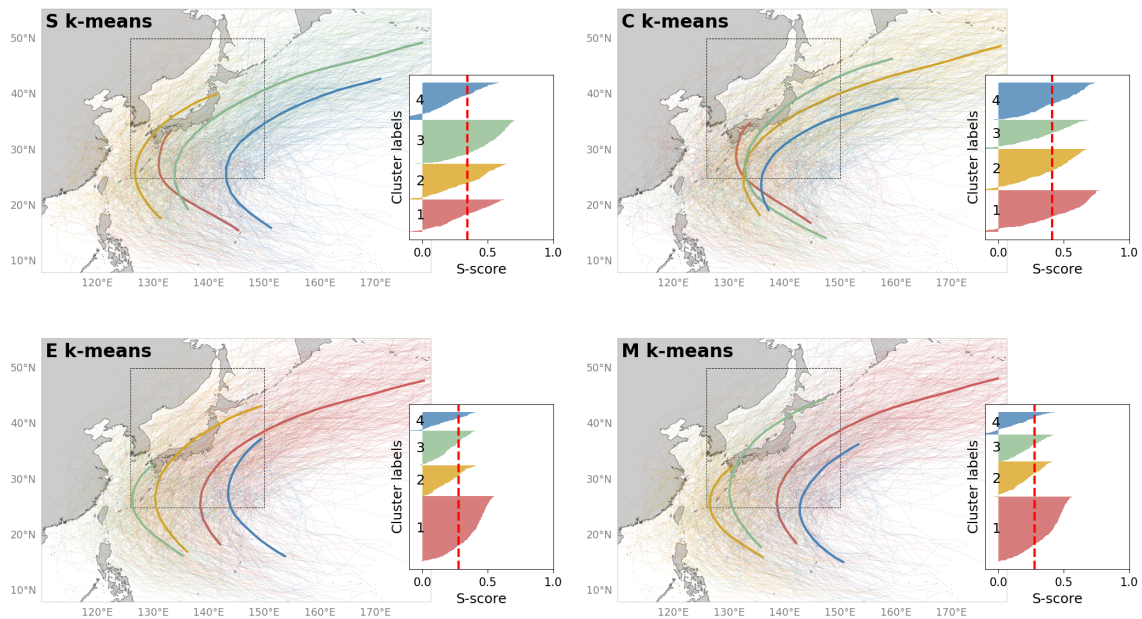


Fig. S21 Similar with Figure 5 in the maintext but for run 3. Results of the TC experiment for clustering tropical cyclone paths. The pattern was revealed by S, C, E, and M k-means, with $k = 4$. Input data are the best TC tracks obtained by the JMA from 1951 – 2020. Only TCs that passed the dashed box in the map are used to feed the k-means. Thus, a total of 863 TC tracking data points are used. The left side of each

panel show the general silhouette analysis results, where the x-axis indicates the score (S-score) and y-axis presents the labels of clusters numbered 1 – 4. The centroid TC path is illustrated by the bold line, and the color is consistent with that in the silhouette diagram.

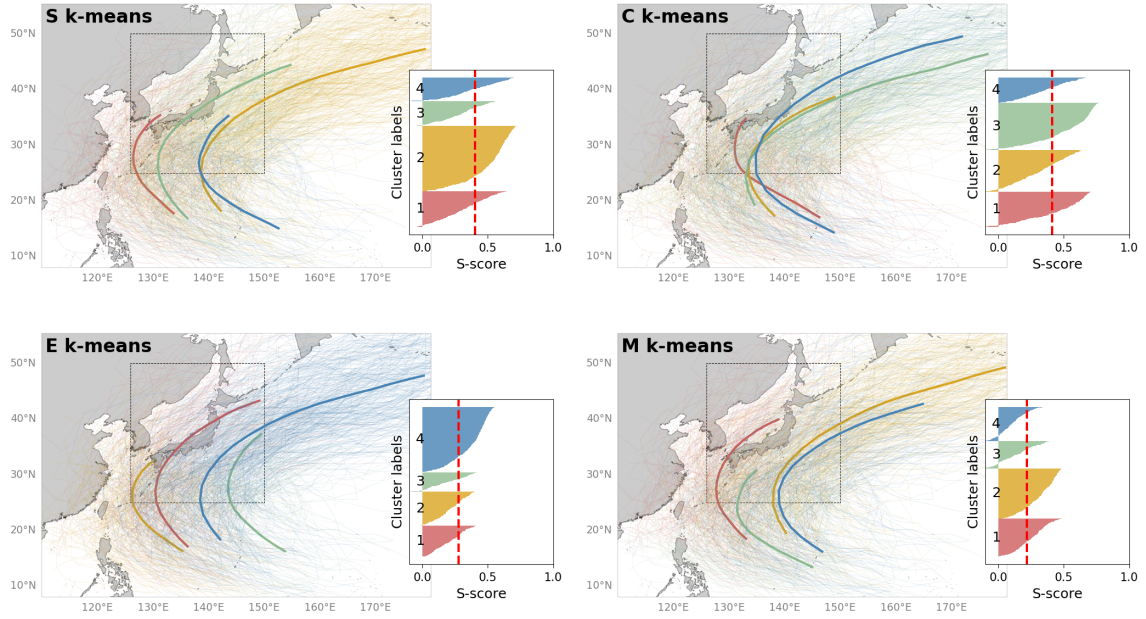


Fig. S22 Similar with Figure 5 in the maintext but for run 4. Results of the TC experiment for clustering tropical cyclone paths. The pattern was revealed by S, C, E, and M k-means, with $k = 4$. Input data are the best TC tracks obtained by the JMA from 1951 – 2020. Only TCs that passed the dashed box in the map are used to feed the k-means. Thus, a total of 863 TC tracking data points are used. The left side of each panel show the general silhouette analysis results, where the x-axis indicates the score (S-score) and y-axis presents the labels of clusters numbered 1 – 4. The centroid TC path is illustrated by the bold line, and the color is consistent with that in the silhouette diagram.

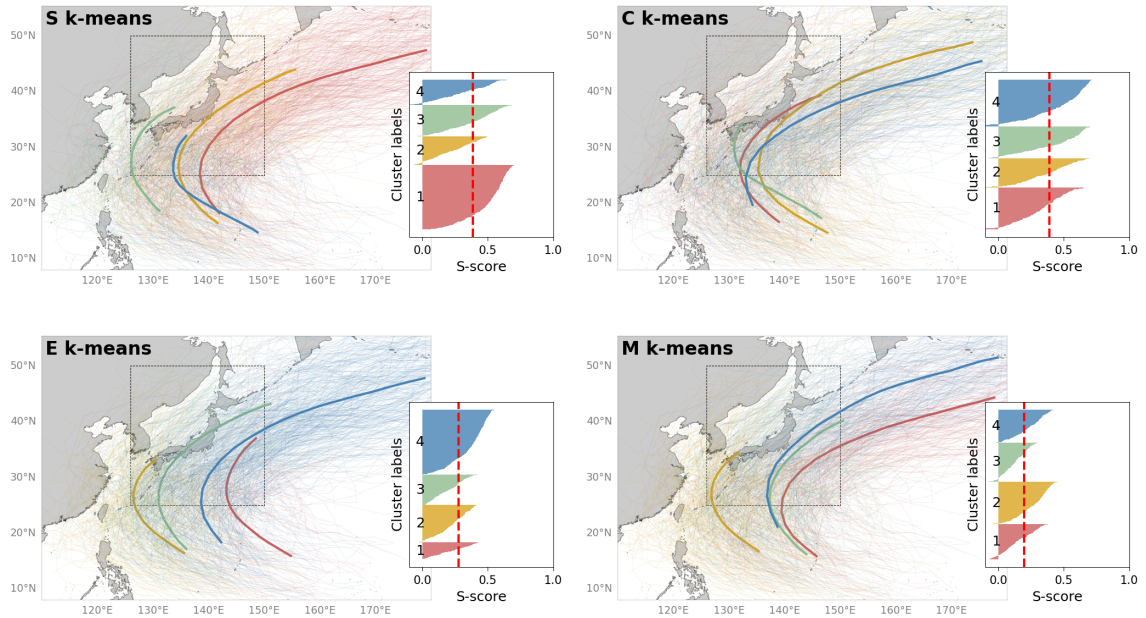


Fig. S23 Similar with Figure 5 in the maintext but for run 5. Results of the TC experiment for clustering tropical cyclone paths. The pattern was revealed by S, C, E, and M k-means, with $k = 4$. Input data are the best TC tracks obtained by the JMA from 1951 – 2020. Only TCs that passed the dashed box in the map are used to feed the k-means. Thus, a total of 863 TC tracking data points are used. The left side of each panel show the general silhouette analysis results, where the x-axis indicates the score (S-score) and y-axis presents the labels of clusters numbered 1 – 4. The centroid TC path is illustrated by the bold line, and the color is consistent with that in the silhouette diagram.

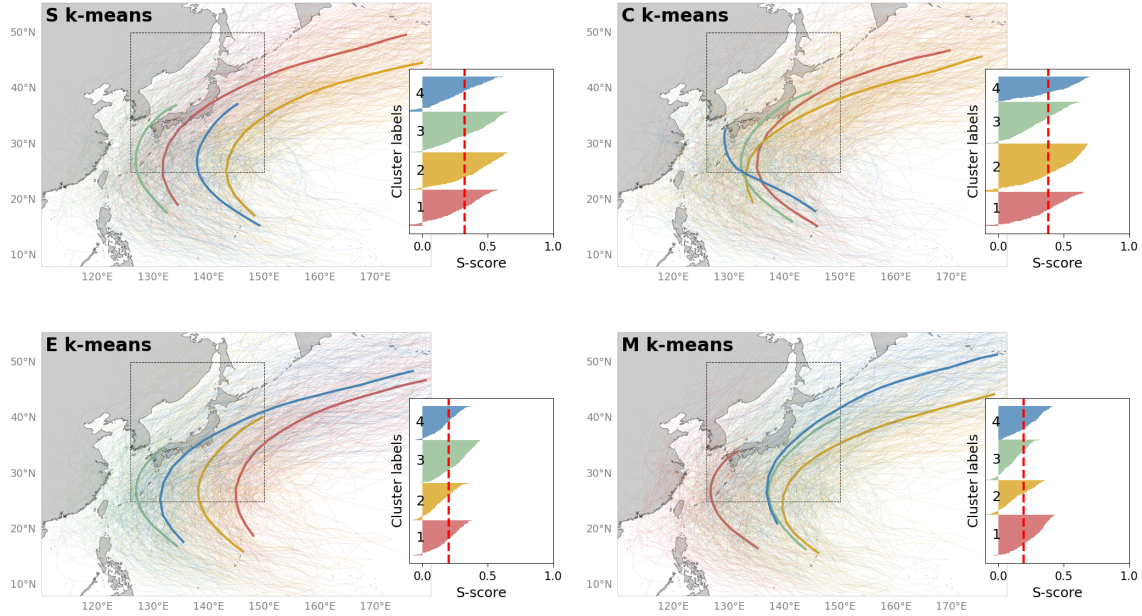


Fig. S24 Similar with Figure 5 in the maintext but for run 6. Results of the TC experiment for clustering tropical cyclone paths. The pattern was revealed by S, C, E, and M k-means, with $k = 4$. Input data are the best TC tracks obtained by the JMA from 1951 – 2020. Only TCs that passed the dashed box in the map are used to feed the k-means. Thus, a total of 863 TC tracking data points are used. The left side of each panel show the general silhouette analysis results, where the x-axis indicates the score (S-score) and y-axis presents the labels of clusters numbered 1 – 4. The centroid TC path is illustrated by the bold line, and the color is consistent with that in the silhouette diagram.

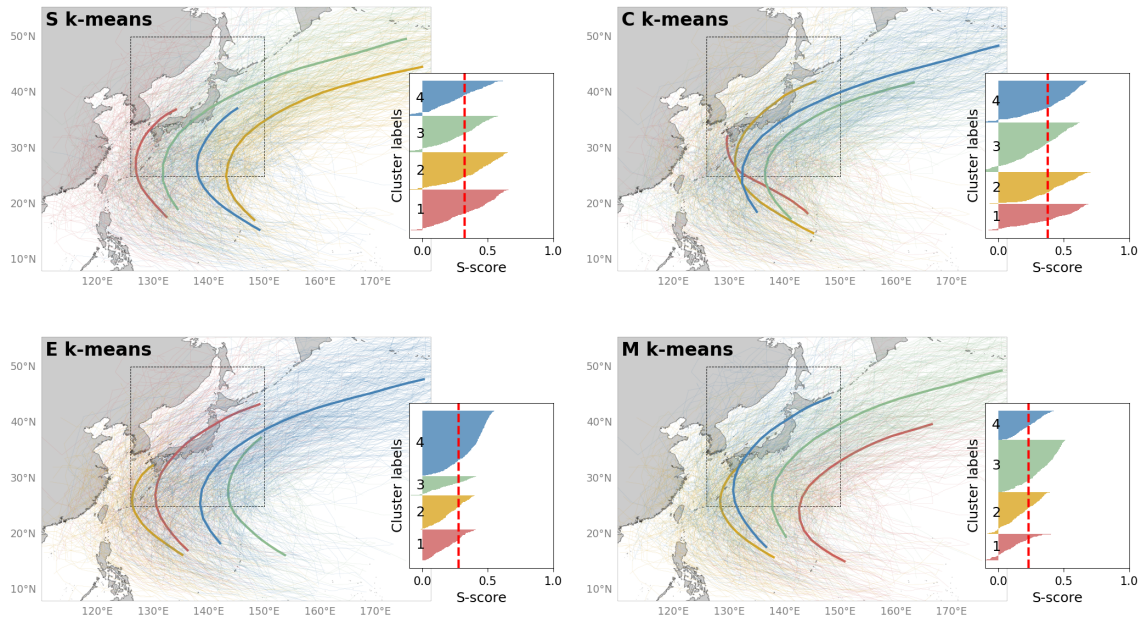


Fig. S25 Similar with Figure 5 in the maintext but for run 7. Results of the TC experiment for clustering tropical cyclone paths. The pattern was revealed by S, C, E, and M k-means, with $k = 4$. Input data are the best TC tracks obtained by the JMA from 1951 – 2020. Only TCs that passed the dashed box in the map are used to feed the k-means. Thus, a total of 863 TC tracking data points are used. The left side of each panel show the general silhouette analysis results, where the x-axis indicates the score (S-score) and y-axis presents the labels of clusters numbered 1 – 4. The centroid TC path is illustrated by the bold line, and the color is consistent with that in the silhouette diagram.

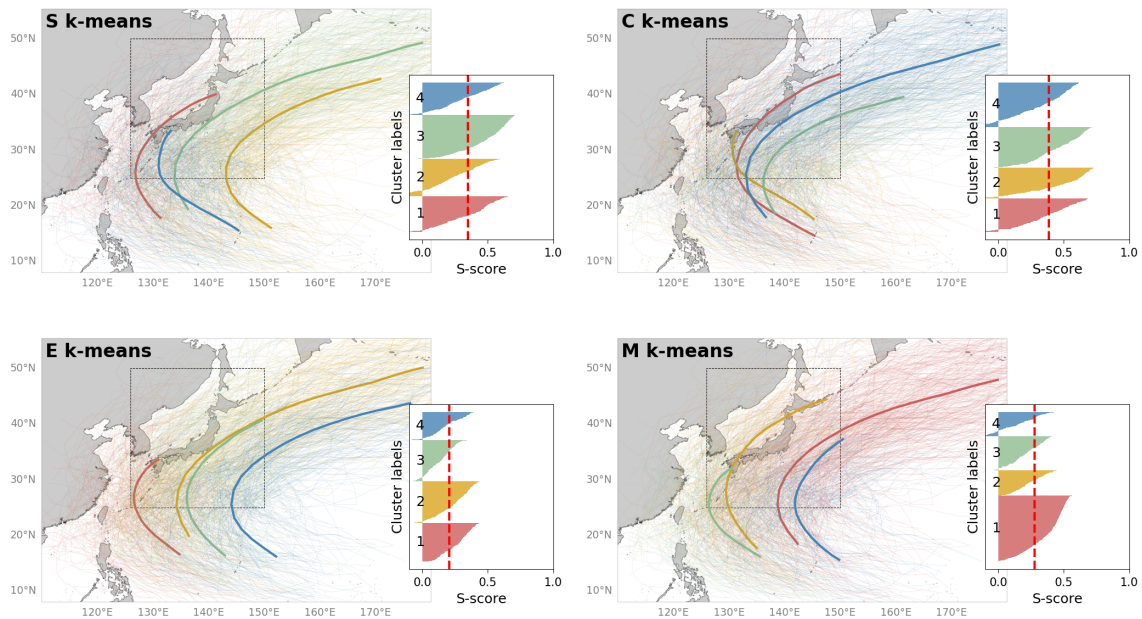


Fig. S26 Similar with Figure 5 in the maintext but for run 8. Results of the TC experiment for clustering tropical cyclone paths. The pattern was revealed by S, C, E, and M k-means, with $k = 4$. Input data are the best TC tracks obtained by the JMA from 1951 – 2020. Only TCs that passed the dashed box in the map are used to feed the k-means. Thus, a total of 863 TC tracking data points are used. The left side of each

panel show the general silhouette analysis results, where the x-axis indicates the score (S-score) and y-axis presents the labels of clusters numbered 1 – 4. The centroid TC path is illustrated by the bold line, and the color is consistent with that in the silhouette diagram.

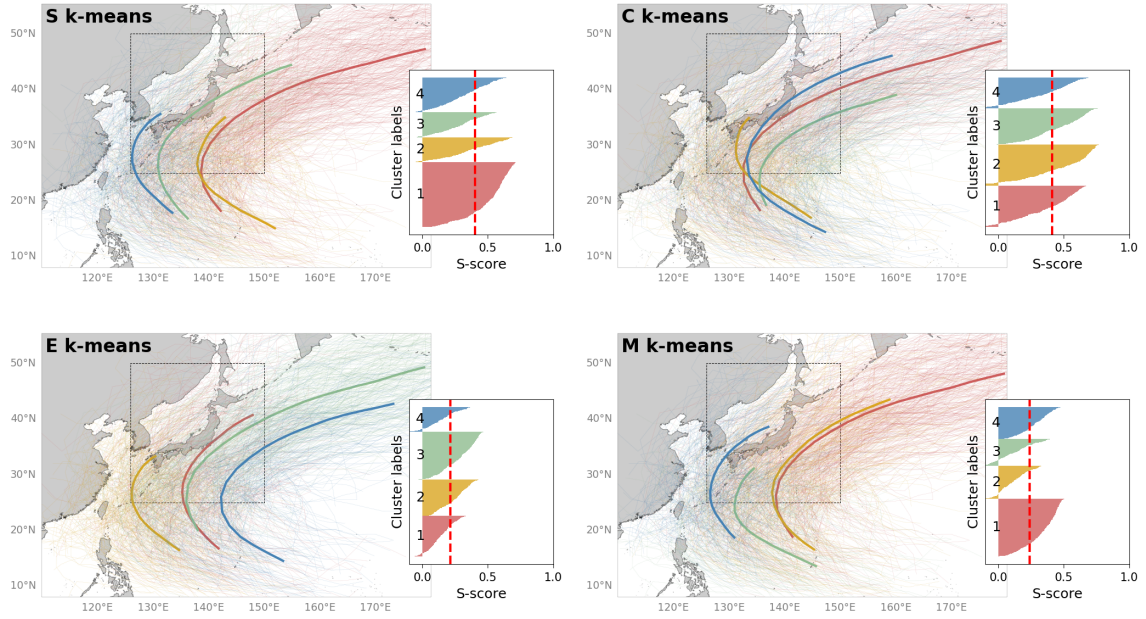


Fig. S27 Similar with Figure 5 in the maintext but for run 9. Results of the TC experiment for clustering tropical cyclone paths. The pattern was revealed by S, C, E, and M k-means, with $k = 4$. Input data are the best TC tracks obtained by the JMA from 1951 – 2020. Only TCs that passed the dashed box in the map are used to feed the k-means. Thus, a total of 863 TC tracking data points are used. The left side of each panel show the general silhouette analysis results, where the x-axis indicates the score (S-score) and y-axis presents the labels of clusters numbered 1 – 4. The centroid TC path is illustrated by the bold line, and the color is consistent with that in the silhouette diagram.

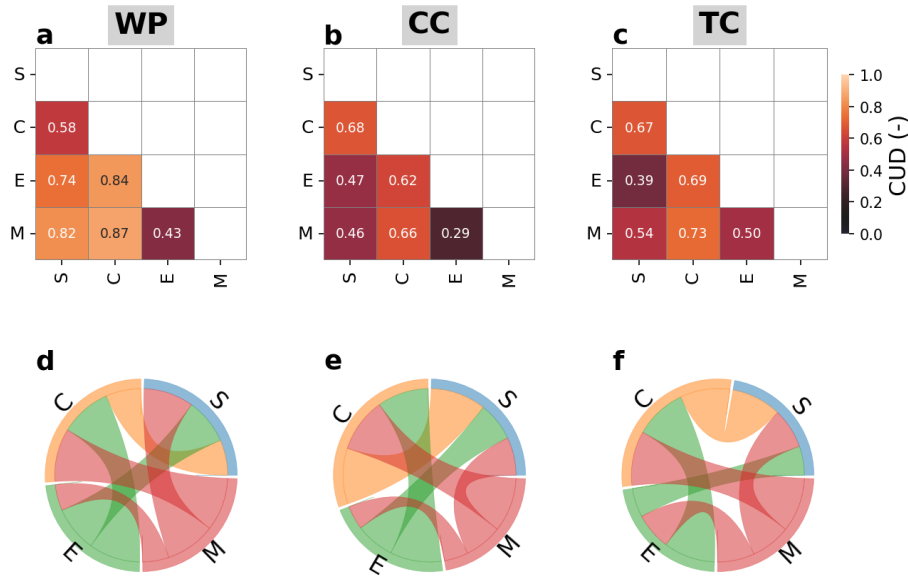


Fig. S28 Clustering uncertainty degree (CUD) based on adjusted mutual information (AMI) between clustering results from different k-means algorithms, i.e., S, C, E, and M k-means, for different experiments: WP, CC, and TC. (a, b, c) CUD in heatmaps, and (d, e, f) visualization of the interconnection using the chord diagrams. Note that the results are from the configuration with $k=4$ and the four k-means variants use the same starting centroids.

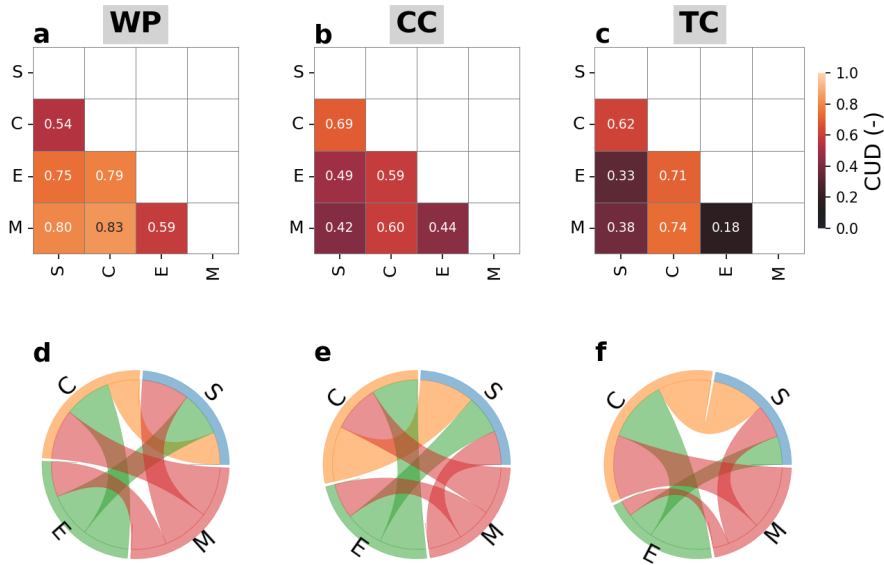


Fig. S29 Clustering uncertainty degree (CUD) based on adjusted mutual information (AMI) between clustering results from different k-means algorithms, i.e., S, C, E, and M k-means++, for different demo experiments: WP, CC, and TC. (a, b, c) CUD in heatmaps, and (d, e, f) visualization of the interconnection using the chord diagrams. Note that the results are from the configuration with $k=4$ and the first initialization run.

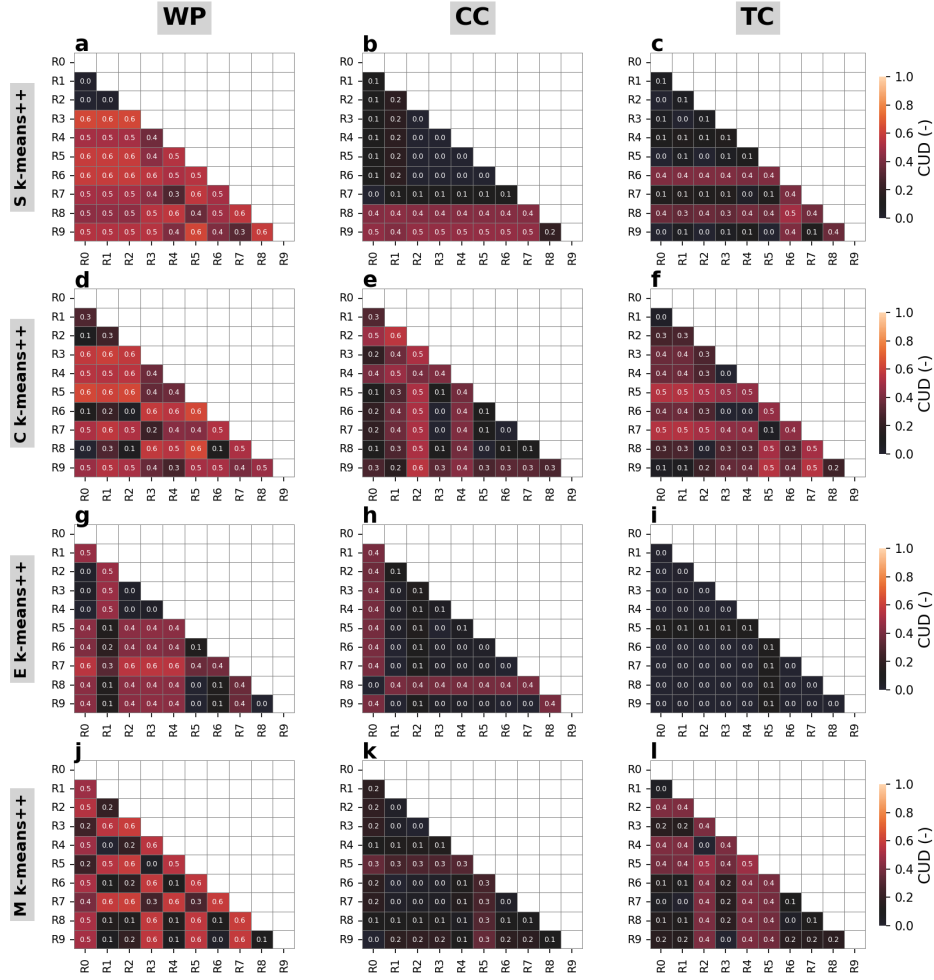


Fig. S30 Clustering uncertainty degree (CUD) based on adjusted mutual information (AMI) between the clustering results from different runs (10 runs indicated by R0, R1, ..., R9) of different k -means++ algorithms, i.e., S, C, E, and M k -means++ (rows), for different demo experiments: WP, CC, and TC (columns). Note that the results are from the configuration with $k=4$ and the first initialization run.