1 Extended Method

K-means Clustering

K-Means clustering is an unsupervised machine learning algorithm used to partition a dataset into distinct clusters based on feature similarity (MacQueen, 1967). In this study, K-Means is applied to climate data, specifically analyzing selected climate variables: Rx5day relative anomaly and anomalies of consecutive dry days (CDD). These variables represent projected changes in extreme rainfall and drought conditions across Indonesia. The analysis is based on CMIP6 model outputs, comparing data from 2081–2100 under SSP5-8.5 to the 1981–2010 baseline.

The selected variables are standardized using the scikit-learn library (Pedregosa et al., 2011) to ensure that all features contribute equally to the clustering process, preventing biases due to differences in magnitude (Han et al., 2012). The preprocessed dataset is then clustered using the K-Means algorithm with k=3, meaning the data is categorized into three distinct groups. The chosen number of clusters represents relative changes in climate variables: negative (decreasing), near neutral, and positive (increasing) compared to the baseline period. Then, the gridded raster files are used to obtain smoothed contour delineation, simplifying the three zoning classifications.

As shown in Figure 8 in the main manuscript, this classification provides a simplified yet insightful way to interpret future climate projections. One important consideration is that this clustering method does not account for the statistical significance of relative changes. Unlike traditional hypothesis testing or trend significance assessments, K-Means simply groups data points based on similarity without evaluating whether the identified changes are statistically robust. Despite this limitation, the method remains valuable for detecting spatial patterns in climate projections, offering an efficient way to identify regions with similar projected characteristics.

This approach is particularly useful for regional climate assessments, as it helps identify areas with similar projected climate characteristics. By providing clearer spatial patterns, the clustering results support more accessible climate information, aiding in the development of adaptation strategies and informed policy-making.

References

Han, J., Kamber, M., and Pei, J.: Data mining concepts and techniques, third edition, 2012.

MacQueen, J.: Some Methods for Classification and Analysis of Multivariate Observations, in: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability - Vol. 1, 281–297, 1967.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., and others: Scikit-learn: Machine learning in Python, the Journal of machine Learning research, 12, 2825–2830, 2011.

2 Supplementary Tables

Table S1 GCMs used in ensemble obtained from https://interactive-atlas.ipcc.ch/

No	Model	Native Resolution	
		X	Y
1	ACCESS-CM2	1.88°	1.25°
2	ACCESS-ESM1-5	1.88°	1.25°
3	BCC-CSM2-MR	1.12°	1.11°
4	CAMS-CSM1-0	1°	1°
5	CANESM5	2.81°	2.77°
6	CESM2	1.25°	0.9°
7	CESM2-WACCM	1.25°	0.94°
8	CMCC-CM2-SR5	1°	1°
9	CNRM-CM6-1	1.41°	1.39°
10	CNRM-CM6-1-HR	0.5°	0.5°
11	CNRM-ESM2-1	1.41°	1.39°
12	EC-EARTH3	0.7°	0.7°
13	EC-EARTH3-Veg	0.7°	0.7°
14	FGOALS-g3	2°	5.18°
15	GFDL-CM4	1.25°	1°
16	GFDL-ESM4	1.25°	1°
17	HADGEM3-GC31-LL	1.88°	1.25°
18	IITM-ESM	2°	2°
19	INM-CM4-8	2°	1.5°
20	INM-CM5-0	2°	1.5°
21	IPSL-CM6A-LR	2.5°	1.27°
22	KACE-1-0-G	1.88°	1.25°
23	MIROC-ES2L	2.81°	2.77°
24	MIROC6	1.41°	1.39°
25	MPI-ESM1-2-HR	0.93°	0.93°
26	MPI-ESM1-2-LR	1.88°	1.85°
27	MRI-ESM2-0	1.12°	1.11°
28	NESM3	1.88°	1.85°
29	NorESM2-LM	2.5°	1.89°
30	NorESM2-MM	1.25°	0.94°
31	TaiESM1	1.25°	0.9°
32	UKESM1-0-LL	1.88°	1.25°

Table S2 RCMs from CORDEX SEA used in ensemble obtained from https://interactive-atlas.ipcc.ch/

No	**Institution**	GCM	Version
1	MOHC-HadGEM2-ES	GERICS-REMO2015	v1
2	MPI-M-MPI-ESM-LR	GERICS-REMO2015	v1
3	NCC-NorESM1-M	GERICS-REMO2015	v1
4	ICHEC-EC-EARTH	ICTP-RegCM4-3	v4
5	IPSL-IPSL-CM5A-LR	ICTP-RegCM4-3	v4
6	MOHC-HadGEM2-ES	ICTP-RegCM4-3	v4
7	MPI-M-MPI-ESM-MR	ICTP-RegCM4-3	v4
8	MOHC-HadGEM2-ES	ICTP-RegCM4-7	v0
9	MPI-M-MPI-ESM-MR	ICTP-RegCM4-7	v0
10	NCC-NorESM1-M	ICTP-RegCM4-7	v0
11	MOHC-HadGEM2-ES	SMHI-RCA4	v1

3 Supplementary Figures

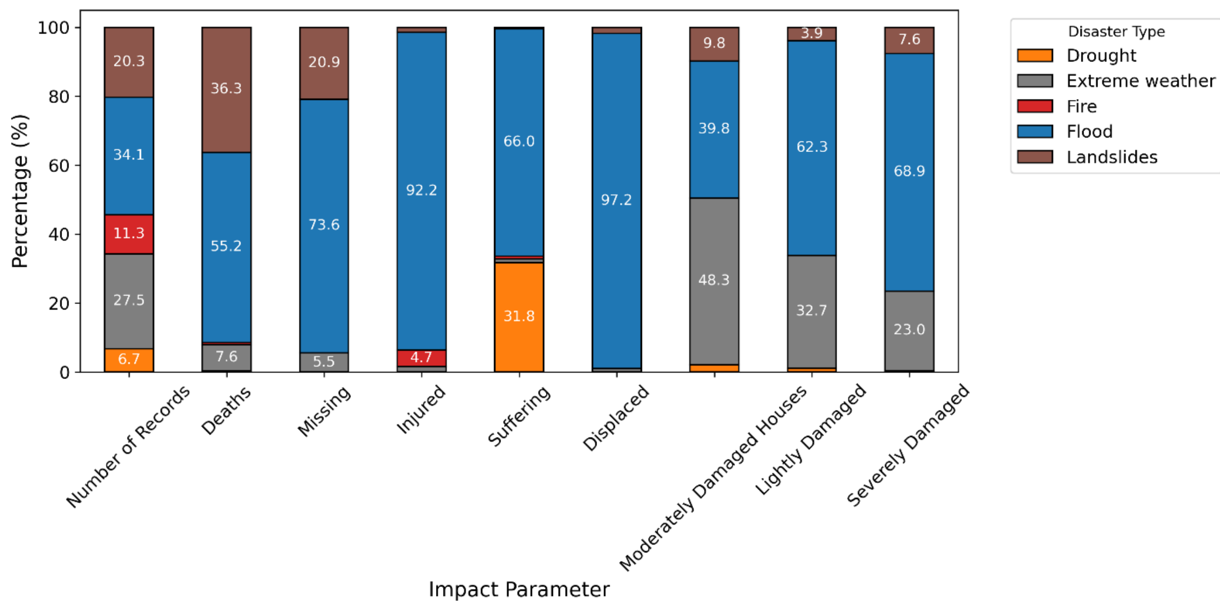

Figure S1 Proportion of recorded natural hazard events (2000–2023) categorized by hazard type (drought, extreme weather, wildfire, flood, and landslides) obtained from the national disaster database DIBI.