File explanations

1. Source data for logistic regression analyses (rdml\_and\_55fkt.csv)

This file includes topography and geological variables of one million pseudo-absence sites and 55 wind-hole sites:

* fkt: dummy variable indicating whether each point is wind-hole site or not (1/0)
* sl0: mean and maximum slope angles at 0-m radius circular buffer
* ms’xx’: mean slope angle (‘xx’ means the radius of circular buffer)
* xs’xx’: maximum slope angle (‘xx’ means the radius of circular buffer)
* c’xx’: curvature: (‘xx’ means the radius of circular buffer)
* sg01: categorical geological variable whether volcanic rocks had larger proportions or not (1/0)
* sgm: continuous geological variable indicating the proportion of volcanic rocks

1. Source data for MAXLIKE

* fkt55points\_xy.csv: x-y coordinates of 55 wind-hole sites
* sl0\_rm.tif: raster data of mean and maximum slope angles at 0-m radius circular buffer
* ms’xx’\_rm.tif: raster data of mean slope angle (‘xx’ means the radius of circular buffer)
* xs’xx’\_rm.tif: raster data of maximum slope angle (‘xx’ means the radius of circular buffer)
* c’xx’\_rm.tif: raster data of curvature (‘xx’ means the radius of circular buffer)
* sg200\_01rm2.tif: raster data of categorical geological variable
* sg200m\_rm.tif: raster data of continuous geological variable

Variables used in the logistic regression analyses were generated from these raster data.

1. Predicted large-scale distribution of wind-hole sites (maxlike2013\_100\_50.tif)

A predicted occurrence probabilities generated from the fitted MAXLIKE model