

SKHASH: A Python Package for Computing Earthquake Focal Mechanisms

v0.1 – manual last modified: 16 February 2024

Robert J. Skoumal, U.S. Geological Survey, Moffett Field, CA, rskoumal@usgs.gov

Jeanne L. Hardebeck, U.S. Geological Survey, Moffett Field, CA, jhardebeck@usgs.gov

Peter M. Shearer, Scripps Institution of Oceanography, La Jolla, CA, shearer@igpp.ucsd.edu

Overview

SKHASH has been tested on all versions of Python that are currently supported and released (as of January 2024: 3.8, 3.9, 3.10, & 3.11) on Linux, macOS, and Windows.

User inputs are provided in a control file that is passed as a command line argument:

```
python3 SKHASH.py path_to/control_file.txt
```

Optionally, user inputs can also be passed via the command line with (or without) the control file:

```
python3 SKHASH.py path_to/control_file.txt --num_cpus=8
```

Control File

The control file is an ordinary text file containing the user-defined variables. In the control file, variable names are preceded by a '\$' with the desired value on the following line, e.g.:

```
# Text that follows a '#' (like this line) are ignored.
```

```
$fpfile      # P-polarity input filepath
examples/hash1/IN/north1.phase.txt
```

```
$plfile      # station polarity reversal filepath
examples/hash1/IN/scsn.reverse.txt
```

The exception is for velocity model filepaths, which can be either space delimited or put on multiple lines, e.g.:

```
$vmodel_paths # velocity model filepaths on multiple lines
examples/velocity_models/socal.txt
examples/velocity_models/north.txt
examples/velocity_models/lab1.txt
```

Text on a line following a '#' is ignored. The input variables in the control file can be listed in any order. If a user input is not provided in the control file, the default value (described below) is used.

File Formats

This manual principally describes the flexible 'SKHASH' file formats, and the use of this format is the default. To maintain backwards compatibility, SKHASH also accepts file formats for all previous HASH v1.2 drivers. For file formats using these HASH 1-5 drivers, refer to the [HASH v1.2 manual](#).

When using the SKHASH formats, provide a comma separated file with column names. Any lines that start with a '#' are treated as comments and are ignored. The columns that are required depend on how you would like to compute the mechanisms, and these requirements are described below.

Examples

Included with the code are a handful of examples of how to use SKHASH. All examples are found in the SKHASH/examples/ folder. A brief description of each of these examples are provided here.

smile

A synthetic example in the shape of a friendly face. *P*-wave polarities and *S/P* ratios are provided in the SKHASH format. The source-receiver azimuths and takeoff angles are precomputed. Quality control output files are created along with an automatic beachball plot.

maacama

Composite mechanisms for two earthquake families near the Maacama Fault in Northern California. *P*-wave polarity weights were computed using correlation consensus. Earthquake hypocentral information is provided in the polarity file rather than a separate catalog file. Using the station metadata, source-receiver azimuths and takeoff angles are computed.

hash1

The remaining five examples come directly from HASH, demonstrating the backwards compatibility of our code. Each example contains 24 mechanisms from the Northridge, California, sequence. See the [HASH v1.2 manual](#) for information about the file formats.

P-polarities with precomputed source-receiver azimuths and takeoff angles. File formats follow the hash_driver1.f example.

hash2

P-polarities and station metadata are provided. The source-receiver azimuths and takeoff angles are computed. File formats follow the hash_driver2.f example.

hash3

P-polarities, *S/P* ratios, and station metadata are provided. The source-receiver azimuths and takeoff angles are computed. File formats follow the hash_driver3.f example.

hash4

P-polarities and station metadata are provided similar to hash2, but using an updated SCEDC file format that includes five character station codes. The source-receiver azimuths and takeoff angles are computed. File formats follow the hash_driver4.f example.

hash5

P-polarities are provided similar to hash2, but using precomputed source-receiver azimuths and takeoff angles determined with SIMULPS 3D ray-tracing format file. File formats follow the hash_driver5.f example.

SKHASH File Format

Polarity File(s) [\$fpfile, \$impfile, \$conpfile, \$dlpfile]

Files containing traditional (\$fpfile), imputed (\$impfile), consensus (\$conpfile), and deep-learning (\$dlpfile) *P*-polarity measurements.

Always required:

event_id: The earthquake event identifier. Measurements with the same *event_id* are used to produce a unique mechanism.

p_polarity: weight of the polarity measurement. +: up, -: down. Float within [-1, 1]

Required by default:

network: Network code. To ignore, set \$require_network_match=False

station: Station code. To ignore, set \$require_station_match=False

location: Location code. To ignore, set \$require_location_match=False

channel: Channel code. To ignore, set \$require_channel_match=False

If computing takeoff angles and source-receiver azimuths*:

origin_latitude: Latitude of earthquake.

origin_longitude: Longitude of earthquake.

origin_depth_km: Depth of earthquake in km.

horz_uncert_km: Optional. Horizontal hypocentral uncertainty in km.

vert_uncert_km: Optional. Vertical hypocentral uncertainty in km.

*This information can be given in a catalog file (\$catfile) instead.

If using precomputed takeoff and source-receiver azimuths:

takeoff: Takeoff angle, degrees from vertical. (<90: upgoing, >90: downgoing).

takeoff_uncertainty: Optional. Takeoff angle uncertainty in degrees.

azimuth: Source-receiver azimuth.

azimuth_uncertainty: Optional. Source-receiver azimuth uncertainty in degrees.

sr_dist_km: Optional. Source-receiver distance in km.

If computing composite mechanisms:

event_id2: Represents the earthquake event identifier that the measurement was made on.

The first few lines from an example polarity file:

| <i>event_id</i> | <i>station</i> | <i>network</i> | <i>location</i> | <i>channel</i> | <i>p_polarity</i> |
|-----------------|----------------|----------------|-----------------|----------------|-------------------|
| 123456, | ST000, | XX, | --, | EHZ, | 1.0 |
| 123456, | ST001, | XX, | --, | EHZ, | -1.0 |
| 123456, | ST002, | XX, | --, | EHZ, | 0.5 |

The first few lines from an example polarity file with precomputed takeoffs and azimuths:

| <i>event_id</i> | <i>station</i> | <i>network</i> | <i>location</i> | <i>channel</i> | <i>p_polarity</i> | <i>takeoff</i> | <i>azimuth</i> |
|-----------------|----------------|----------------|-----------------|----------------|-------------------|----------------|----------------|
| 123456, | ST000, | XX, | --, | EHZ, | 1.0, | 115.0, | 105.0 |
| 123456, | ST001, | XX, | --, | EHZ, | -1.0, | 115.0, | 110.3 |
| 123456, | ST002, | XX, | --, | EHZ, | 0.5, | 115.0, | 115.6 |

The first few lines from an example polarity file to compute composite mechanisms, in addition to providing hypocentral information:

| <i>event_id</i> | <i>event_id2</i> | <i>station</i> | <i>location</i> | <i>channel</i> | <i>p_polarity</i> | <i>origin_latitude</i> | <i>origin_longitude</i> | <i>origin_depth_km</i> |
|-----------------|------------------|----------------|-----------------|----------------|-------------------|------------------------|-------------------------|------------------------|
| 1, | 123456, | ST000, | --, | EHZ, | 0.02566, | 39.36743, | -123.25380, | 7.14200 |
| 1, | 123456, | ST001, | --, | DPZ, | -0.00193, | 39.36743, | -123.25380, | 7.14200 |
| 1, | 123456, | ST002, | 01, | EHZ, | 0.00002, | 39.36743, | -123.25380, | 7.14200 |

S/P File(s) [*\$ampfile*, *\$relampfile*]

Files containing traditional (*\$ampfile*) and relative (*\$relampfile*) *S/P* ratio measurements.

Always required:

event_id: Earthquake event identifier. Measurements with the same *event_id* are used to produce a unique mechanism.

Required by default:

network: Network code. To ignore, set *\$require_network_match=False*

station: Station code. To ignore, set *\$require_station_match=False*

location: Location code. To ignore, set *\$require_location_match=False*

channel: Channel code. To ignore, set *\$require_channel_match=False*

If computing *S/P* ratios:

noise_p: Noise level prior to the *P*-wave window

noise_s: Noise level prior to the *S*-wave window

amp_p: *P*-amplitude

amp_s: *S*-amplitude

If using pre-computed *S/P* ratios:

sp_ratio: Your pre-computed *S/P* ratio.

If computing takeoff angles and source-receiver azimuths*:

origin_latitude: Latitude of earthquake.

origin_longitude: Longitude of earthquake.

origin_depth_km: Depth of earthquake in km.

horz_uncert_km: Optional. Horizontal hypocentral uncertainty in km.

vert_uncert_km: Optional. Vertical hypocentral uncertainty in km.

This information can be given in a catalog file (\$catfile*) instead.

If using precomputed takeoff and source-receiver azimuths:

takeoff: Takeoff angle, degrees from vertical. (<90: upgoing, >90, downgoing).

azimuth: Source-receiver azimuth.

takeoff_uncertainty: Optional. Takeoff angle uncertainty in degrees.

azimuth_uncertainty: Optional. Source-receiver azimuth uncertainty in degrees.

sr_dist_km: Optional. Source-receiver distance in km.

Station Corrections [\$corfile]

When using S/P ratios, you may wish to include station corrections.

Always required:

sta_correction: The station correction, $\log_{10}(S/P)$.

Required by default:

network: Network code. To ignore, set \$require_network_match=False

station: Station code. To ignore, set \$require_station_match=False

location: Location code. To ignore, set \$require_location_match=False

channel: Channel code. To ignore, set \$require_channel_match=False

Station File [\$stfile]

File containing the station locations. The station file is required whenever computing takeoff angles and source-receiver azimuths.

Always required:

latitude: Station latitude.

longitude: Station longitude.

elevation: Station elevation in km.

Required by default:

network: Network code. To ignore, set \$require_network_match=False

station: Station code. To ignore, set \$require_station_match=False

location: Location code. To ignore, set \$require_location_match=False

channel: Channel code. To ignore, set \$require_channel_match=False

Polarity Reversal File [\$plfile]

File containing information about stations with reversed polarities (e.g., up appears as down). If there are no reversed stations, this file is not needed.

Required:

start_time: Start of the reversal window at the given station, string (e.g., "1990-04-01")

end_time: End of the reversal window at the given station, string (e.g., "1991-11-01")

Required by default:

network: Network code. To ignore, set \$require_network_match=False

station: Station code. To ignore, set \$require_station_match=False

location: Location code. To ignore, set \$require_location_match=False

channel: Channel code. To ignore, set \$require_channel_match=False

Catalog File [`$catfile`]

Rather than providing information about hypocentral locations in the above *P*-polarity and *S/P* files, they can be specified in the catalog file.

Required:

event_id: Earthquake event identifier.

time: Origin time of the earthquake

latitude: Latitude of the hypocenter.

longitude: Longitude of the hypocenter.

depth: Depth of the hypocenter.

Optional:

horz_uncert_km: Horizontal hypocentral uncertainty in km.

vert_uncert_km: Vertical hypocentral uncertainty in km.

Velocity model file format

Each velocity model should be in a separate file. Each velocity model file is a comma delimited file following the format:

Depth (km) , Vp (km/s)

Any lines that begin with a '#' are treated as comment lines and are ignored. An example velocity model file:

```
# Comment line in the velocity model file
```

```
# Depth(km) , Vp(km/s)
```

```
0.0, 4.02
```

```
1.0, 4.28
```

```
3.0, 4.79
```

```
4.0, 5.04
```

```
10.0, 5.92
```

```
20.0, 6.78
```

Variables

Variables are declared in your control file and change how the program is run. While not recommended, default values can be overwritten by changing the values near the top of the SKHASH.py file.

File Formats

Tells the program how to read your input files. If you are using the 'SKHASH' format, none of these file format variables are needed.

`$input_format` (default: `skhash`)

The general format of your input files. All other files will be read using this format unless you specify a different file type for the individual file (e.g., `input_format_fpf`) that are listed below.

Possible values:

`skhash`, `ncsn`, `hash1`, `hash2`, `hash3`, `hash4`, `hash5`.

The following variables are only needed if they differ from your `$input_format`.

`$input_format_fpf`: Traditional *P*-polarity file format.

`$input_format_impfile`: Imputed *P*-polarity file format.

`$input_format_conpf`: Consensus *P*-polarity file format.

`$input_format_dlpfile`: Deep-learning *P*-polarity file format.

`$input_format_plfile`: *P*-polarity reversal file format.

`$input_format_ampfile`: *S*- and *P*-wave amplitude input file format.

`$input_format_relampfile`: Relative *S/P* ratio file format.

`$input_format_stfile`: Station list file format.

| <i>Acceptable input formats for the different file types</i> | | File Format | | | | | | |
|--|--------------------------------------|-------------|------|-------|-------|-------|-------|-------|
| | | skhash | ncsn | hash1 | hash2 | hash3 | hash4 | hash5 |
| Variable Name | <code>input_format_fpf</code> | X | X | X | X | X | X | X |
| | <code>input_format_impfile</code> | X | X | X | X | X | X | X |
| | <code>input_format_conpf</code> | X | X | X | X | X | X | X |
| | <code>input_format_dlpfile</code> | X | X | X | X | X | X | X |
| | <code>input_format_plfile</code> | X | | X | X | X | X | X |
| | <code>input_format_ampfile</code> | X | | | | X | | |
| | <code>input_format_relampfile</code> | X | | | | X | | |
| | <code>input_format_stfile</code> | X | | | X | X | X | |

Input File Paths

The following variables will point SKHASH to the files that it needs to run. All values should be a filepath to the corresponding file. For example:

`$catfile`
`PATH_TO/catfile.txt`

Metadata:

`$catfile` : Earthquake catalog information file

`$stfile` : Station list file

`$plfile` : Station polarity reversal file

P-polarities:

`$fpfile` : *P*-polarity input file
`$impfile` : Imputed *P*-polarity input file
`$conpfile` : Consensus *P*-polarity input file
`$dlpfile` : Deep-learning *P*-polarity input file

S/P Ratios:

`$ampfile` : Amplitude input file
`$corfile` : Station correction file
`$relampfile` : Relative *S/P* ratios input file

SIMULPS File (for HASH5 format):

`$simulpsfile` : SIMULPS file containing azimuth and takeoff angles.

Output File Paths

The following variables will determine the type of output files that are produced and where to save them. All values should be a filepath to the corresponding file. For example:

`$outfile1`
PATH_TO/outfile1.txt

`$outfile1` : Path to focal mechanisms output file
`$outfile2` : Path to acceptable plane output file
`$outfile_pol_agree` : Path to record of polarity (dis)agreement output file
`$outfile_sp_agree` : Path to record of *S/P* difference output file
`$outfile_pol_info` : Path to record of all polarities considered in the mechanisms along with information (station lat/lon, avg takeoff angles, avg source-receiver azimuths, etc.) file
`$outfolder_plots` : Path to folder where simple focal mechanism plots will be created (outfolder_plots/event_id.png). To ignore, leave blank (no plots will be generated)

Direct Program Controls

The following variables control specific numerical aspects of how SKHASH will generate focal mechanisms.

`$npolmin`
Minimum number of polarity data (integer value)
Default value:
8

`$nmc`
Number of trials (integer value)
Default value:
30

`$maxout`
Maximum number of focal mechanisms outputted (integer value)
Default value:
500

`$ratmin`

Minimum allowed signal-to-noise ratio (float).

Default value:

3

`$badfrac`

Fraction of *P*-polarities assumed to be incorrect (float).

Default value:

0.1

`$badmin`

Minimum number of *P*-polarities per event assumed to be incorrect. If `$badmin` > [(sum of polarity weights)*`$badfrac`], then `$badmin` is used instead of `$badfrac` (float).

Default value:

2.0

`$qbadfrac`

Assumed noise in *S/P* amplitude ratios, log₁₀ (e.g. 0.3 for a factor of 2) (float).

Default value:

0.3

`$qbadmin`

Minimum *S/P* amplitude misfit allowed per event. If `$qbadmin` > [(sum of number of *S/P* measurements)*`$qbadfrac`], then `$qbadmin` is used instead of `$qbadfrac` (float).

Default value:

2.0

`$min_polarity_weight`

Any polarities with abs(weight) < min_polarity_weight will be ignored (float).

Default value:

0.1

`$delmin`

Minimum allowed source-station distance in km; if delmin ≤ 0, all results ≥ look_del[0] distance are used (float).

Default value:

0

`$delmax`

Maximum allowed source-station distance in km; if delmax ≥ 0, all results ≤ look_del[1] distance are used (float).

Default value:

0

`$azmax`

Maximum allowed source-station azimuth uncertainty in degrees. To use all results regardless of azimuthal uncertainty, set azmax to 0 (float).

Default value:

\$pmax

Maximum allowed takeoff (“plungal”) uncertainty in degrees. To use all results regardless of takeoff uncertainty, set pmax to 0 (float).

Default value:

0

\$cangle

Angle in degrees for computing focal mechanism probability (float).

Default value:

45

\$prob_max

Probability threshold for multiples (float).

Default value:

0.2

\$max_agap

Maximum azimuthal gap between stations in degrees (float).

Default value:

90

\$max_pgap

Maximum “plungal” gap between stations in degrees (float).

Default value:

60

\$vmodel_paths

List of paths to velocity models, delimited by whitespace or line breaks (string).

PATH_TO/vmodel1.txt PATH_TO/vmodel2.txt

Plotting Parameters

The following variables control how plots are made.

\$outfolder_plots : Path to folder where simple focal mechanism plots will be created (outfolder_plots/event_id.png). To ignore, leave blank (no plots will be generated)

\$plot_station_names : If True, station names will be included in the beachball plot

\$plot_acceptable_solutions: If True, all acceptable solutions will be plotted in the beachball plot

Additional Parameters

The following variables control error flags, quality-of-life features, and other functions.

\$write_lookup_table

If set to True, the lookup table will be saved so that it can be used in future runs. A “.npz” file extension will be added to the vmodel_paths (boolean).

Default value:

False

`$recompute_lookup_table`

If a lookup table with an identical filename already exists, SKHASH will remake it.

Default value:

`True`

`$require_temporal_match`

When determining the station information for a given pick, SKHASH requires the reported earthquake to occur between the station metadata start-end time

Default value:

`False`

`$require_network_match`

When determining the station information for a given pick, SKHASH requires the reported pick network to match the metadata

Default value:

`True`

`$require_station_match`

When determining the station information for a given pick, SKHASH requires the reported pick station to match the metadata

Default value:

`True`

`$require_channel_match`

When determining the station information for a given pick, SKHASH requires the reported pick channel to match the metadata

Default value:

`True`

`$allow_duplicate_stations`

Allows an event_id to have multiple polarities and *S/P* ratios from the same receiver

Default value:

`False`

`$iterative_avg`

If set to True, SKHASH will iteratively compute the average mech by removing the solution furthest from the average following HASH. This is significantly slower than the SKHASH approach.

Default value:

`False`

`$min_quality_report`

Only mechanism qualities that are better than or equal to min_quality_report will be reported. To return all accepted solutions regardless of quality, leave blank.

Possible values:

`A,B,C,D`

`$overwrite_output_file`

If set to True, SKHASH will overwrite any preexisting output files

Default value:

True

`$perturb_epicentral_location`

If True, SKHASH will randomly perturb the horizontal location of the earthquakes

Default value:

False

`$ignore_missing_metadata`

If set to True, any measurements with missing metadata will be ignored. If set to False, an error will be raised.

Default value:

False

`$use_fortran`

If set to True, SKHASH will attempt to load the Fortran grid search subroutine. If it cannot load or is False, the Python routine will be used instead.

Default value:

True

`$num_cpus`

Number of cores to run things in parallel. Set to 0 to use all available, or 1 to run in serial.

Default value:

1

`$dang`

Minimum grid spacing in degrees

Default value:

5

`$nx0`

Maximum source-station distance bins for lookup tables

Default value:

101

`$nd0`

Maximum source depth bins for lookup tables

Default value:

14

`$look_dep`

Minimum source depth, maximum, and interval for the lookup table

Default value:

[0, 39, 3]

`$look_del`

Minimum source-station distance, maximum, and interval for the lookup table

Default value:

[0, 200, 2]

`$allow_hypocenters_outside_table`

If set to True, then the minimum/maximum lookup table depth is used for hypocenters that fall outside of the lookup table range. If set to False, an error will be produced.

Default value:

False

`$nump`

Number of rays traced

Default value:

9000

`$min_sp`

If any user-provided S/P ratios are less than min_sp after applying station corrections, those values are replaced with min_sp. To ignore, use a value ≤ 0 .

Default value:

0.0005

`$max_sp`

If any user-provided S/P ratios are greater than max_sp after applying station corrections, those values are replaced with max_sp. To ignore, use a value ≤ 0 .

Default value:

2000

`$min_amp`

Minimum amplitude [0-1] considered when calculating expected *P* and *S* amplitudes. Values < min_amp are treated as 0. To ignore, use a value ≤ 0 .

Default value:

0.0005

`$epicenter_degree_precision`

Number of decimal places to round lat, lon epicenter coordinates.

Default value:

6

`$epicenter_degree_precision`

Number of decimal places to round lat, lon epicenter coordinates.

Default value:

6

`$output_angle_precision`

Number of decimal places to output for angles (e.g., strike, dip, rake, takeoff)

Default value:

1

`$output_quality_precision`

Number of decimal places to output for mechanism qualities (e.g., prob, mfrac, mavg, stdr)

Default value:

1

`$output_km_distance_precision`

Number of decimal places to output for source-receiver distances (reported in km)

Default value:

3

`$output_vector_precision`

Number of decimal places to output for normal vectors

Default value:

4

`$default_vert_uncert_km`

If the vertical uncertainty for an earthquake is missing or negative, this value is used instead when perturbing the source depth.

Default value:

1

`$default_horz_uncert_km`

If the horizontal uncertainty for an earthquake is missing or negative, this value is used instead when perturbing the source depth.

Default value:

1

‘NCSN’ File Format

A Northern California Seismic Network phase/polarity file in Hypoinverse format obtained directly from the [Northern California Earthquake Data Center](#) can also be read. For information on this file format, refer to the [Shadow 2000](#) document.

The assigned weight code for a phase arrival is used to determine the weight of the polarity measurement (column 17). The weights are identical to the station arrival time weighting. The weight code : absolute value of the polarity weight is:

| Weight Code | abs(weight) |
|-------------|-------------|
| 0 | 1.0 |
| 1 | 0.5 |
| 2 | 0.2 |
| 3 | 0.1 |
| ≥4 | 0.0 |

When using the ‘NCSN’ format for *P*-polarities, all other files formats will default to the ‘SKHASH’ format described above.

If you are reading in NCSN polarities, be sure to specify to add the following lines somewhere in your control file:

```
$input_format_fpfiler
ncsn
```