REAL

0. Author

Miao Zhang, Dalhousie University (miao.zhang@dal.ca)

1. References

Zhang, M., W.L. Ellsworth, and G.C. Beroza. Rapid Earthquake Association and Location, Seismol. Res. Lett., 90.6, 2276-2284, 2019.

Kissling, E., W.L. Ellsworth, D. Eberhart-Phillips, and U. Kradolfer: Initial reference models in local earthquake tomography, J. Geophys. Res., 99, 19635-19646, 1994.

Waldhauser F. and W.L. Ellsworth, A double-difference earthquake location algorithm: Method and application to the northern Hayward fault, Bull. Seismol. Soc. Am., 90, 1353-1368, 2000.

2. Introduction

REAL (Rapid Earthquake Association and Location) associates arrivals of different seismic phases and locates seismic events primarily through counting the number of P and S picks and secondarily from traveltime residuals. A group of picks are associated with a particular earthquake if there are enough picks within the theoretical traveltime windows. The location is determined to be at the grid point with most picks (Fig. 1a). If multiple locations have the same maximum number of picks, the grid point with smallest traveltime residual is selected (Fig. 1b). We refine seismic locations using a least-squares location method (VELEST) and a high-precision relative location method (HypoDD).

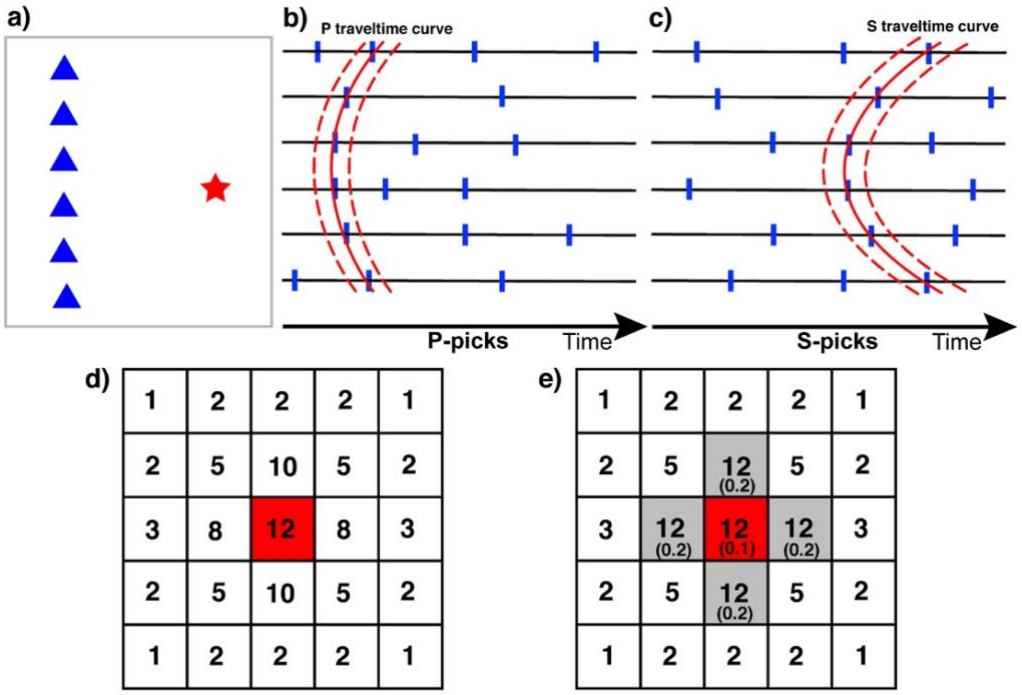

Figure 1. Cartoon illustrating the concept of REAL. a) The distributions of seismic event (red star) and seismic stations (blue triangles). b) *P* arrival time curve (red) with its uncertainty range (red dashed), associated *P* picks and other false *P* picks (blue bars). c) same as b) but for *S* picks. d) The optimal location is determined to be at the grid point with most picks or e) the grid point with smallest traveltime residual (shown in bottom parentheses, unit: second) among multiple locations with the same maximum number of picks.

3. Requirements
VELEST (http://www.seg.ethz.ch/software/velest.html, for further location, already provided)
hypoinverse (https://www.usgs.gov/software/hypoinverse-earthquake-location, optional)
PhaseNet (https://github.com/wayneweiqiang/PhaseNet.git, optional)
hypoDD (https://www.ldeo.columbia.edu/~felixw/hypoDD.html, for DD relocation)
ObsPy (https://github.com/obspy/obspy, for data downloading, processing and picking only)
SAC (http://ds.iris.edu/ds/nodes/dmc/software/downloads/sac, for data processing only)
GMT (https://www.soest.hawaii.edu/gmt/, for figure plotting only)
Matlab (https://www.mathworks.com/products/matlab.html, for figure plotting only)

4. Usage (Type "REAL")
-D (nyear/nmon/nday/lat_center)
-R (rx/rh/tdx/tdh/tint[/gap/GCarc0/latref0/lonref0]])
-V (vp0/vs0[/s_vp0/s_vs0/ielev])
-S (np0/ns0/nps0/npsboth0/std0/dtps/nrt[/drt/nxd//rsel/rsel])
[-G (trx/trh/tdx/tdh)]
station
pickdir
[ttime]

----------explanation----------
"[]" denotes optional parameters.

-D:
(nyear/nmon/nday/lat_center)
(year/month/day)
e.g., 2016/10/01
date of the day (yy/mm/dd) and and latitude center of the target region (deg., so that lat and lon have consistent distance in km)
[REAL can process seismic picks recorded in one day (or a few days but only up to 31 days, e.g., 2016/10/01 – 2016/10/31 but not eligible for 2016/10/02 – 2016/11/01). All picks are relative to ZERO of the day (e.g., 60.00 corresponds to 2016/10/14 00:01:00.00 and 86460 corresponds to 2016/10/15 00:01:00.00)]

-R:
(rx/rh/tdx/tdh/tint[/gap/GCarc0/latref0/lonref0])
(degree/km/degree/km/sec[degree/degree/degree/degree])
e.g., 0.1/20/0.02/2/5.0 or 0.1/20/0.02/2/5.0/360 or 0.1/20/0.02/2/5.0/360/180 or 0.1/20/0.02/2/5.0/360/180/42.75/13.25
rx: search range in horizontal centered at the station recorded the initiating phase (degree)
[e.g., > twice of average station interval; smaller than the travel time range]
rh: search range in depth (km)
[e.g., 30 km for crustal earthquakes, will search depth from 0 to 30 km]
tdx: search grid size for epicenter (degree)
tdh: search grid size for depth (km)

[time cost mostly depends on the total number of grids! please consider using sparse search grid and strict threshold for quick test. Your tdx and tdh don't have to be the same as that in the travel time table. REAL will automatically interpolate travel time.]
tint: two events cannot appear within tint sec, otherwise only keep the most reliable one (tint sec)
[REAL will use the S travel time within one grid instead if your tint is too small; if not provided, it will use S time window]
gap: only keep events with reasonable station gap (e.g., default is 360^o)
GCarc0: only keep picks with distance of < GCarc0 (e.g., default is 180^o)
[GCarc0 should be smaller than trx in -G, otherwise it will be replaced by trx-0.05]
latref0: reference latitude (degree)
lonref0: reference longitude (degree)
[latref0 and lonref0 are optional; they will be the search center for all picks. if not provided, the location of the station recording the initiating phases will be used]

-V:
(vp0/vs0/[s_vp0/s_vs0/ielev])
(km/s|km/s|[km/s|km/s|int])
e.g., 6.2/3.3 or 6.2/3.3/5.5/2.7/1
vp0: average P velocity (homogenous model) (km/s)
vs0: average S velocity (homogenous model) (km/s)
[vp0 and vs0 will be used to automatically calculate the default time window (time window between two dashed lines in Fig. 1), i.e., sqrt(tdx**2+ tdx**2+tdh**2)/vp0 for P time window, sqrt(tdx**2+ tdx**2+tdh**2)/vs0 for S time window, also see nrt in -S]
s_vp0: shallow P velocity (km/s)
s_vs0: shallow S velocity (km/s)
ielev: station elevation correction (optional; default is 0; if 1, then s_vp0 and s_vs0 must be given for elevation correction. Note this assumes vertical ray path and corrects by elevation/s_vp0 for P phase and elevation/s_vs0 for S phase. Usually, you don't need to use the elevation correction due to crustal compensation. In your velocity model, zero depth means the average elevation of your stations if ielev is 0. Please shift your model if it is relative to the sea level)

-S:
(np0/ns0/nps0/npsboth0/std0/dtps/nrt[drt/nxd/rsel/ires])
(int/int/int/int/double/double/double/[double/int])
e.g., 3/2/5/2/0.5/0.2/1.5/0.5, 3/2/5/2/0.5/0.2/1.5/0.5/0.5, 3/2/5/2/0.5/0.2/1.5/0.5/0.5/4, or 3/2/5/1/0.5/0.2/1.5/0.5/0.5/4/1
np0: threshold for number of P picks
ns0: threshold for number of S picks
nps0: threshold for total number of picks (P&S)
npsboth0: number of stations that recorded both P and S picks
[It can significantly improve your association reliability, especially when you have a poor station coverage. Along with np0, ns0, nps0, more picks/stations, more strict thresholds]
std0: standard deviation threshold (residual)
[Here residual is defined as the deviation of the origin times that were estimated from different picks (for the same event). It is not the traditional RMS residual]

dtps: time threshold for S and P separation
[dtps is used to remove some false S picks. P picks may appear in S pick pool in real data when applying STA/LTA pickers. S picks will not be used if ts-tp is less than dtps. Set dtps = 0 to turn this constraint off]
nrt: nrt*default time window (usually between 1 and 2)
[i.e., nrt*sqrt(tdx**2+ tdx**2+tdh**2)/vp0 for P time window or nrt*sqrt(tdx**2+ tdx**2+tdh**2)/vs0 for S time window. It accommodates the inaccuracy of velocity model (as well as pick uncertainty). Use larger nrt if velocity model is insufficient. Note: nrt and tdx/tdh trade off! If you use a large grid size, please use small nrt, vice versa. If your P and S t_dist curves are too narrow (clearly clipped), maybe your velocity model is too bad or the nrt is too small.]
drt: drt*default time window (usually < 1). Remove associated picks < drt*P_window from the initiating pick pool. Use it as small as possible if time affordable. Default is 0.5. Like the nrt, drt and tdx/tdh trade off. Please use a small drt if you use a large grid size.
nxd: suspicious events with the nearest station > nxd*GCarc0 will be discarded.
rsel: tolerance multiplier; keep picks with residuals less than rsel*STD, and remove suspicious picks in large distance (i.e., tpick should be smaller than tpick_median + 0.75*rsel*tpick_STD, 0.75 is fixed in code).
[rsel is used to remove picks with large residuals and large distance. Default rsel is 4.0 (a large value, i.e., approximately turn this constraint off). For example, if your final standard deviation is 0.1 sec, REAL will automatically remove those suspicious picks with residuals > 0.4 sec; If you have large topography (i.e., station corr.), inaccurate velocity or picking uncertainty, picks with large residual may be correct. please increase rsel.]
ires: resolution_or_not [optional; 1-output resolution file; 0-don't output. Note: only works for the first associated event when the first initiating pick is true. Thus, this is only recommended for synthetic resolution analysis]

[-G:]
[optional, no need for homogenous model case]
(trx/trh/tdx/tdh)
(degree/km/degree/km)
e.g., 1.4/20/0.01/1
trx: horizontal range in traveltime table (degree)
trh: vertical range in traveltime table (km)
tdx: grid size in traveltime table (degree)
tdh: grid size in traveltime table (km)

Formats for input files:
pick file:
File name – net.station.phase.txt (e.g., YR.ED01.P.txt)
Phase pick (e.g., 13.977 1.0 0.0)
arrivaltime (sec), stalta_ratio or phase_probability, amplitude_in_millimeter
[If you don't have amplitude information, set amplitude_in_millimeter as ZERO and REAL will not estimate the magnitude. Using weighting factors is highly recommended, which can be STA/LTA ratios (from STA/LTA picker) or phase probabilities (from machine-learning picker)]

station file:
e.g., 13.0929 42.5264 YR ED01 HHZ 0.739
lon., lat., network, station, component, elevation (km)

traveltime table:
e.g., 1.25 2 22.797 41.701 17.920 24.639 -0.0622 -0.2564 P S
distance (degree), depth (km), P traveltime (sec), S traveltime (sec), p_ray_parameter (sec/degree), s_ray parameter (sec/degree), p_vertical_slowness (sec/km), s_vertical_slowness (sec/km), P_or_p, S_or_s
[REAL accepts different search grids within the traveltime table. It will automatically look for the table and interpolate traveltimes using ray parameters. This file is not needed for homogenous velocity model case. Please make sure your distance and depth match your study region. REAL cannot use event-station distance larger than your largest distance in the table]

Formats for output files:
catalog_sel.txt:
e.g., 1 2016 10 14 00:00:08.175 8.175 0.0748 42.8167 13.2282 10.00 -inf nan 60 60 120 60 29.38
num, year, mon, day, time (hh:mm:ss), origin time (relative to ZERO, sec), residual (sec), lat., lon., dep., mag., mag var (uncertainty), number of P picks, number of S picks, total number of picks, number of stations with both P and S, station gap
[If you don't calculate magnitude, mag and mag_var would be "-inf" and "nan", respectively]

phase_sel.txt:
e.g., 1 2016 10 14 00:00:08.175 8.175 0.0748 42.8167 13.2282 10.00 -inf nan 60 60 120 60 29.38
YR ED06 P 14.7946 6.6192 0.00e+00 -0.0544 1.0000 0.0654
Event line: same as catalog.txt
Phase line: network, station, phase name, absolute travetime (relative to ZERO, sec), traveltime relative to event origin time (sec), phase amplitude in millimeter, individual phase residual (sec), weight, azimuth

hypolocSA.dat:
e.g., 2016 10 14 00 00 8.106 42.8158 13.2226 10.845 0.00 120 29.377 0.004 1
refined location using a simulated annealing method (two steps: 1) fix depth, search epicenter; 2) fix epicenter, search depth. Distance weighed, ranging from 1 to 0.5 (use a cos function)).
year, mon, day, hh, mm, ss, lat., lon., dep., mag., total number of picks, station gap (deg.), residual (sec), num

hypophase.dat:
e.g., # 2016 10 14 00 00 8.106 42.8158 13.2226 10.845 0.00 0 0 0 1
ED06 6.688 1 P
hypoDD's phase format. It can be used as hypoDD's input directly.
Flag, year, mon, day, hh, mm, sec, lat, lon, dep, mag., errz, errh, err_o, num
Station, traveltime, weighting, phase

resolution.txt [optional for resolution test]:
e.g., 29.9994 42.7533 13.2472 10.0000 60 60 120 0.0722
origin time (sec), lat., lon., dep., number of P picks, number of S picks, number of total picks, residual (sec)