

## Users Guide for Early-est



### EARthquake Rapid Location sYstem with ESTimation of Tsumamigenesis

[2015-04-24 - Early-est Version 1.1.9]

Anthony Lomax [anthony@alomax.net](mailto:anthony@alomax.net)  
*ALomax Scientific, Mouans-Sartoux, France.*

Alberto Michelini [alberto.michelini@ingv.it](mailto:alberto.michelini@ingv.it) Fabrizio Bernardi, Valentino Lauciani  
*Istituto Nazionale di Geofisica e Vulcanologia (INGV), Roma, Italy.*

- 
- [Introduction](#)
  - [Program Requirements](#)
  - [Installing Early-est](#)
  - [Running Early-est, Examples](#)
  - [Output Files](#)
  - [Algorithm Description](#)
  - [References](#)
  - [Revision History Notes](#)
- 

## Introduction

**Early-est** (Bernardi et al., 2015) is a software package for rapid location and seismic / tsunamigenic characterization of earthquakes. The **Early-est** software package operates on offline-event or continuous-realtime seismic waveform data to perform trace processing and picking, and, at a regular *report interval*, phase association, event detection, hypocenter location, and event characterization. This characterization (Table 1) includes  $m_b$  and  $M_{wp}$  magnitudes, the determination of apparent rupture duration,  $T_0$ , large earthquake magnitude,  $M_{wpd}$ , and assessment of tsunamigenic potential using  $T_d$  and  $T50Ex$ , as described in Lomax and Michelini (2009ab, 2011).

This distribution includes two **Early-est** time-domain processing programs that operate on mini-SEED data packets:

*miniseed\_process* - Runs **Early-est** time-domain processing on Mini-SEED event files, e.g. downloaded from IRIS-Wilbur [http://www.iris.edu/cgi-bin/wilberII\\_page1.pl](http://www.iris.edu/cgi-bin/wilberII_page1.pl)).

*seedlink\_monitor* - Runs **Early-est** time-domain processing on continuous data returned from one or more SeedLink servers.

The software and examples in this distribution are configured for global (teleseismic to regional scale) monitoring of events of about  $M_4$  and larger, using up to 200-500 stations and the ak135 model (Kennett et al., 1995). This configuration uses a sliding analysis window of 1 hour, an association/location/reporting interval of 1 min, and a maximum of 10 events located each report interval.

These settings may limit application to the local scale, though the software is stable and robust for processing events recorded only at nearby stations. Program application properties files allow changing the configuration settings for use on a regional or local scale.

This distribution also includes a set of bash, GMT and Python extension scripts to perform focal mechanism determinations and generate html and graphical output showing extensive system and event information.

The core Early-est software consists of a set of time-domain processing modules written in C and distributed under the GNU General Public License. The core software uses libslink (<http://www.iris.edu/pub/programs/SeedLink>) and libmseed (<http://www.iris.edu/software/libraries>) for managing SeedLink connections and reading mini-SEED data, and the libxml2 XML parser and toolkit.

The **Early-est** extension scripts make use of: the HASH program (Hardebeck and Shearer, 2002; <http://quake.wr.usgs.gov/research/software/#HASH>) for computing focal mechanisms; and bash, Python and the GMT package (<http://www.soest.hawaii.edu/GMT>) for plotting maps, histograms and alphanumeric information.

Early-est development has been supported by INGV, through "Centro Nazionale Terremoti" institutional funds and through grants from the Italian Dipartimento della Protezione Civile, and by the NERA project (Network of European Research Infrastructures for Earthquake Risk Assessment and Mitigation) under the European Community's Seventh Framework Programme [FP7/2007-2013] grant agreement n° 262330.

Early-est is distributed under the terms of the GNU General Public License:

*This program is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.*

*This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.*

*You should have received a copy of the GNU General Public License along with this program. If not, see <http://www.gnu.org/licenses>.*

---

## Program Requirements

**Early-est** has been developed and tested on Mac OS X 10.9.5 (BSD UNIX). **Early-est** should compile and run correctly in a command-line shell on all Linux/UNIX based systems with gcc 4.2 or higher, or equivalent C compiler and linking support.

Compilation of **Early-est** requires that libslink, libmseed and libxml2 are installed and available on the gcc library search path. Use of the HASH program requires gfortran or equivalent FORTRAN compiler and linking support. Use of the **Early-est** extension scripts requires that bash, ghostscript and Python (e.g. v2.7.5 or later) are installed. For graphical output, the executables from the plotting package GMT (e.g. v4.5.7 or later) must be installed and available on the system executable path.

The *seedlink\_monitor* program generates output files in a new directory for each report interval (e.g. 1 min). The number and total size on disk of these directories and files can become large if *seedlink\_monitor* is run for an extended period; periodic archiving and cleanup of these files is recommended.

---

## Installing Early-est

1. Download and install the following required libraries:

libslink - <http://www.iris.edu/pub/programs/SeedLink>

libmseed - <http://www.iris.edu/software/libraries>

libxml2 - <http://xmlsoft.org/> or <http://www.macports.org/>

2. Copy libslink.a and libmseed.a to your default library directory or a directory on your LD\_LIBRARY\_PATH, e.g. '/usr/local/lib'
3. Copy libslink.h, slplatform.h, libmseed.h and lmplatform.h to your default include directory, e.g. '/usr/local/include'
4. Set the MYBIN environment variable to a directory on your PATH (e.g. 'export MYBIN=/home/my\_home/bin')
5. Unpack early-est-<version>.tgz in a convenient directory.
6. A simple 'make clean' and then 'make' in the early-est-<version>/ directory should build the Early-est programs on most Unix-like systems.
7. To build HASH, cd to early-est-<version>/HASH\_v1.2/ and type 'make clean' and then 'make'. Type cd .. to return back to the early-est-<version>/ directory.
8. For report generation and graphical output using plot\_warning\_report\_seedlink\_runtime.bash: make sure that bash, ghostscript and Python are installed on your system and are on your PATH; install the GMT software package, as described at <http://gmt.soest.hawaii.edu>
9. For uploading report files to a remote site see Running *seedlink\_monitor* below.

### Important Notes:

1. Early-est needs to open simultaneously approximately 15 files for each associated/located event. This number (e.g. about 150 for 10 events) may exceed the limit on number of open file descriptors for each process on some operating systems; this will lead to an “too many open files” error. On Mac OS X 10.6 the default limit is too low; it can be increased by running 'ulimit -n 4096' in the terminal window from which miniseed\_process or seedlink\_monitor will be run, or the command 'ulimit -n 4096' can be placed in the ~/.profile file.
- 

## Running Early-est, Examples

The programs should be run from the early-est-<version>/work directory, which contains necessary helper files.

### *Run miniseed\_process:*

For more information on command line options, use 'miniseed\_process -h'

```
> miniseed_process msprocess_out/Honshu_2011_0_90deg.mseed -g
miniseed_process_station_coordinates.csv -a -pz gainfile.csv -pz-query-host
service.iris.edu -pz-query irisws/resp/1/query -pz-query-type IRIS_WS_RESP -sta-
query-host service.iris.edu -sta-query fdsnws/station/1/query -sta-query-type
```

```

FDSN_WS_STATION -timeseries-query-hosturl http://alomax.net/webtools/sgweb.html?
http://service.iris.edu -timeseries-query irisws/timeseries/1/query -timeseries-
query-type IRIS_WS_TIMESERIES -timeseries-query-sladdr
rtserve.iris.washington.edu:18000 -timeseries-query-hosturl
http://alomax.net/webtools/sgweb.html?http://service.iris.edu -timeseries-query
irisws/timeseries/1/query -timeseries-query-type IRIS_WS_TIMESERIES -timeseries-
query-sladdr 137.227.224.97:18000 -mwp -mb -mwpd -report-delay 1200 -v -alarm -i
AU_XMIS
> python python/processEvents.py MECHANISM msprocess_plots/
msprocess_out/Honshu_2011_0_90deg.mseed.out/ msprocess_out/
msprocess_out/Honshu_2011_0_90deg.mseed.out/ Honshu_2011_0_90deg
> python python/processEvents.py EVENTS msprocess_plots/
msprocess_out/Honshu_2011_0_90deg.mseed.out/ msprocess_out/
msprocess_out/Honshu_2011_0_90deg.mseed.out/ Honshu_2011_0_90deg
> ./plot_warning_report.gmt msprocess_plots/
msprocess_out/Honshu_2011_0_90deg.mseed.out/
msprocess_out/Honshu_2011_0_90deg.mseed.out/ Honshu_2011_0_90deg 0.2
> $PS_VIEWER msprocess_plots/Honshu_2011_0_90deg_Monitor.pdf

```

Open early-est-<version>/work/msprocess\_out/Honshu\_2011\_0\_90deg.mseed.out/events/hypo.1299822380908.html in your browser.

### ***Run seedlink\_monitor:***

For more information on command line options, use 'seedlink\_monitor -h'

Extra installation notes for ***seedlink\_monitor***:

1. Create the directory seedlink\_out/iris (or other path corresponding to the -o option argument)
2. To update http server directory with last real-time warning output files (html, jpg, csv, etc.):
  - a) Set the variable LOCAL\_DIR\_WARNING in plot\_warning\_report\_seedlink\_runtime.bash (or other file you use for the -c value) to an existing local directory or http server directory to use for last real-time warning output files (html, jpg, csv, etc.)
  - b) Copy '\*.html' to the LOCAL\_DIR\_WARNING directory on the http server
3. To use FTP to update an http server directory with last real-time warning output files (html, jpg, csv, etc.)
  - a) Set FTP\_USER\_PW\_HOST and FTP\_DIR\_WARNING environment variables, e.g.,
 

```
export FTP_USER_PW_HOST="ftp://<user>:<password>@<hostname>"
export FTP_DIR_WARNING="/projects/early-est"
```

 or modify the code in plot\_warning\_report\_seedlink\_runtime.bash to copy files directly to a www directory on an http server without using ftp.
  - b) In file plot\_warning\_report\_seedlink\_runtime.bash , comment out with '#' the line DO\_FTP\_TO\_REMOTE=NO.
  - c) Copy '\*.html' to the FTP\_DIR\_WARNING directory on the http server

Use a command of the following form:

```

> seedlink_monitor rtserve.iris.washington.edu:18000 -S
AD_*:00BHZ,IU_*:00BHZ,II_*:00BHZ,G_*:00BHZ,JP_*:00BHZ,FR_*:00BHZ,MN_*:00BHZ,TT_*:00BHZ,HT_*:00BHZ,HL_*:00BHZ,AT_*:00BHZ,PM_*:00BHZ,MY_*:00BHZ,GT_*:00BHZ,CZ_*:00BHZ,IC_*:00BHZ,BL_*:00BHZ
-t -3600 -nt 120 -locs --,00,10,01 -c
./plot_warning_report_seedlink_runtime.bash -g realtime_station_coordinates.csv
-o seedlink_out/iris -a -v -pz gainfile.csv -pz-query-host service.iris.edu -pz-
query irisws/resp/1/query -pz-query-type IRIS_WS_RESP -sta-query-host
service.iris.edu -sta-query fdsnws/station/1/query -sta-query-type
FDSN_WS_STATION -timeseries-query-hosturl http://alomax.net/webtools/sgweb.html?
http://service.iris.edu -timeseries-query irisws/timeseries/1/query -timeseries-
query-type IRIS_WS_TIMESERIES -timeseries-query-sladdr
rtserve.iris.washington.edu:18000 -timeseries-query-hosturl
http://alomax.net/webtools/sgweb.html?http://service.iris.edu -timeseries-query
irisws/timeseries/1/query -timeseries-query-type IRIS_WS_TIMESERIES -timeseries-

```

```
query-sladdr 137.227.224.97:18000 -mwp -mb -mwpd -si -alarm -agency-id  
my.domain.name
```

Note: Press Ctrl-C to stop program after several 'New report generated' messages.

```
> $PS_VIEWER seedlink_plots/t50.pdf
```

and view other files in seedlink\_plots/ (Note: these files can be viewed while the above command is running.)

## ***Run seedlink\_monitor with multiple SeedLink servers:***

Use a command of the following form:

```
> seedlink_monitor rtserve.iris.washington.edu:18000 -S  
AD_*:00BHZ,IU_*:00BHZ,II_*:00BHZ,G_*:00BHZ,GE_*:BHZ,JP_*:BHZ,FR_*:BHZ,MN_*:BHZ,T  
T_*:BHZ,HT_*:BHZ,HL_*:BHZ,AT_*:BHZ,PM_*:BHZ,MY_*:BHZ,GT_*:BHZ,CZ_*:BHZ,IC_*:BHZ,  
BL_*:BHZ -t -3600 -nt 120 -locs --,00,10,01 rtserve.iris.washington.edu:18000 -S  
DK_*:BHZ,GB_LRW:BHZ,GB_KPC:BHZ,GB_JSA:BHZ,GB_MCH1:BHZ,SS_SAIP:BHZ,AU_MORW:BHZ,AU  
_ARMA:BHZ,AU_MTSU:BHZ,AU_COEN:BHZ,AU_FITZ:BHZ,AU_GIRL:BHZ,AU_MILA:BHZ,AU_RIV:BHZ  
,AU_TOO:BHZ,AU_LHI:BHZ,AU_EIDS:BHZ,AU_MCQ:BHZ,AU_MTN:BHZ,AU_NFK:BHZ,AU_XMI:BHZ,A  
U_KDU:BHZ,AU_MOO:BHZ,AU_MANU:BHZ,AU_RABL:BHZ,AU_NIUE:BHZ,AU_MAW:BHZ,NZ_ODZ:BHZ,N  
Z_KHZ:BHZ,NZ_URZ:BHZ,NZ_OUZ:BHZ,BK_CMB:BHZ,CI_PASC:00BHZ,MS_BTDF:BHZ,IM_ATTUB:BH  
Z,PS_PSI:BHZ,AK_GAMB:BHZ,AK_TNA:BHZ -c  
./plot_warning_report_seedlink_runtime.bash -g realtime_station_coordinates.csv  
-o seedlink_out/iris -a -v -pz gainfile.csv -pz-query-host service.iris.edu -pz-  
query irisws/resp/1/query -pz-query-type IRIS_WS_RESP -sta-query-host  
service.iris.edu -sta-query fdsnws/station/1/query -sta-query-type  
FDSN_WS_STATION -timeseries-query-hosturl http://alomax.net/webtools/sgweb.html?  
http://service.iris.edu -timeseries-query irisws/timeseries/1/query -timeseries-  
query-type IRIS_WS_TIMESERIES -timeseries-query-sladdr  
rtserve.iris.washington.edu:18000 -timeseries-query-hosturl  
http://alomax.net/webtools/sgweb.html?http://service.iris.edu -timeseries-query  
irisws/timeseries/1/query -timeseries-query-type IRIS_WS_TIMESERIES -timeseries-  
query-sladdr 137.227.224.97:18000 -mwp -mb -mwpd -si -alarm -agency-id  
my.domain.name
```

Notes:

1. All SeedLink server options must follow immediately after the corresponding SeedLink server address, otherwise the last options set for a previous SeedLink server address will be propagated to subsequent server addresses.
2. With multiple SeedLink servers, data retrieval, latencies and processing at start-up may be asynchronous until all streams have been received up to current realtime.
3. To send an e-mail alert messages, append command-line options of the following form:  
-sendmail "<http://my.domain.name/early-est.me@my.domain.name>,[me@my.domain.name](mailto:me@my.domain.name),[someone.else@domain.name](mailto:someone.else@domain.name)"  
(<base web url>,<from>,<to>[,<to>]...)

---

## **Application properties file**

One startup, miniseed\_process and seedlink\_monitor read additional application options from a properties file in the current working directory named miniseed\_process.prop or seedlink\_monitor.prop, respectively. If the properties file is not present, all option values are set to defaults corresponding to the teleseismic monitor mode of Early-est and to the values set in ttimes.c through included header file ttimes/ttimes\_\*.h.

The format and contents of the application properties files is described below, see one of the application properties files for example property settings.

Syntax

Structure of a settings entity: A settings entity is a tree of depth 2. At the top level are the section nodes that represent isolated sections of application settings (e.g. general application properties, font attributes, color attributes, and so on). At the bottom level are the key-value pairs that represent set values for keys inside a particular section.

A settings entity is read from and written to disk using a textual format. The textual format is parsed by sequentially scanning each line. Blank lines are ignored. For all other lines, leading and trailing blanks are ignored. Sections are enclosed by square brackets ("[" and "]") on separate lines. Key-value pairs are written as the key followed by a single equal sign ("="), followed by the value, all of which must be together on one separate line. A key-value pair is associated to the last seen section. A key-value pair is not valid until at least one section has been seen. All lines that start with a hash("#") are considered comments and are ignored. Empty sections, that is, sections that do not contain any key-value pairs, are not written to disk. Empty sections can be present in the textual format. However, they are ignored when creating a settings entity in memory.

A value can be a single or multi-valued. A multi-valued value is a tuple. Tuples are separated by comma and are singly-typed. This means that a tuple can only contain objects of one kind (e.g. integer tuples, double tuples, etc.).

```
=====
[WaveformExport]
  properties file section with settings for waveform export (miniseed data files)
  Exported waveforms are written in a waveforms/<event_id> subdirectory of the -o output directory,
  e.g., seedlink_out/iris/waveforms/1389094395723/   The waveform files are deleted after the time
  specified by the parameter file_archive.age_max
-----
enable
  enable waveform export
  default: 0
-----
memory.sliding_window.length
  length in seconds of sliding window waveform data memory
  sliding window ends at last data packet start time for each channel
  default: 3600
-----
window.start.before.P
  earliest start time in seconds before P for waveform window
  start time may be later if all waveform data not available
  default: 300
-----
window.end.after.S
  latest end time in seconds after S for waveform window
  end time may be earlier if all waveform data not available
  default: 600
-----
file_archive.age_max
  maximum time period in seconds to keep waveform export files on disk
  default: 10 days
=====
```

```
[Report]
  properties file section with settings for report generation
  The Early-est reporting module calls the oct-tree associate/locate module with the current pick list,
```

and then the reporting module which determines event characterization results and generates graphical and alpha-numeric reporting output.

-----  
warning.colors.show  
enable or disable HTML and alert message display of warning level colors (grey, green, yellow, red)  
(1, 0)  
default: 0

-----  
tsunami.evaluation.write  
enable or disable output of "Tsunami evaluation" information in the e-mail and in the html event  
pages (1, 0)  
default: 0

-----  
report.sliding\_window.length  
length in seconds of the report window.  
default: 3600

-----  
report.interval  
interval in seconds between reports.  
default: 60

-----  
alert.mb\_min  
minimum mb magnitude to trigger an alert (e-mail, run\_action\_\_alert\_sent.bash, etc).  
default: 5.5

-----  
alert.mwp\_min  
minimum Mwp magnitude to trigger an alert (e-mail, run\_action\_\_alert\_sent.bash, etc).  
default: 5.8

-----  
alert.mwpd\_min  
minimum Mwpd magnitude to trigger an alert (e-mail, run\_action\_\_alert\_sent.bash, etc).  
default: 6.95 (rounds to 7.0)

=====  
[AssociateLocate]  
properties file section with settings for association and location

-----  
assoc\_loc.min\_weight\_sum\_associate  
assoc\_loc.min\_weight\_sum\_associate sets the minimum sum of total phase weights for a set of picks  
to be associated into an event  
default: 4.5 (for teleseismic monitor mode)

-----  
assoc\_loc.min\_time\_delay\_between\_picks\_for\_location  
default: 25.0 (same as aref, for teleseismic monitor mode)

-----  
assoc\_loc.gap\_primary\_critical  
default: 225 (for teleseismic monitor mode)

-----  
assoc\_loc.gap\_secondary\_critical  
default: 270 (for teleseismic monitor mode)

-----  
phase.names.use  
sets a list of phase names to use in location/association  
This includes all phases that will count in association/location with non-zero weight and other phases

that will

be passively associated.

Comma separated list with optional \$ separated list of required channels for phase  
default: all phases set in ttimes.c through included header file ttimes/ttimes\_\*.h

-----  
assoc\_loc.\*.\*

sets the bounds and step size for the regular grid search that initiates the oct-tree association/location algorithm

default: e.g. for teleseismic monitor mode:

depth: (DIST\_TIME\_DEPTH\_MAX, 0.0, DIST\_TIME\_DEPTH\_MAX) where  
DIST\_TIME\_DEPTH\_MAX is set in ttimes.c through included header file ttimes/ttimes\_\*.h

lat: (7.2, -90, 90)

lon: (7.2, -180, 180)

-----  
assoc\_loc.nominal\_critical\_oct\_tree\_node\_size

nominal (approximate) size of oct tree node that must be reached before accepting a location  
may be reduced automatically in location.c -> octtreeGlobalAssociationLocation()

default: 50.0 (for teleseismic monitor mode)

-----  
assoc\_loc.min\_critical\_oct\_tree\_node\_size

minimum size of nominal\_critical\_oct\_tree\_node\_size

default: 1.0 (for teleseismic monitor mode)

-----  
assoc\_loc.nominal\_oct\_tree\_min\_node\_size

nominal (approximate) minimum size of oct tree node to be used for association/location  
may be reduced automatically in location.c -> octtreeGlobalAssociationLocation()

default: 5.0 (for teleseismic monitor mode)

---

## Station/Channel Metadata Files

Early-est uses metadata files to read and write station/channel coordinates and gain metadata. These files are read at start-up when data for a station/channel first becomes available. Missing metadata can be added to the files while Early-est is running, and, if the appropriate FDSN or IRIS web-services are available, Early-est can automatically find meta data for new station/channel data (see below). In the future it is planned that Early-est will intermittently check for changes to station/channel metadata using these web-services.

### ***Station coordinates file (e.g. realtime\_station\_coordinates.csv)***

The station coordinates file can be created off-line, the format is:

MN VLC 44.159400 10.386400 555.000000 2014 3 13

<net> <sta> <lat> <lon> <elev (meters)> <year> <month> <day>

where

<year> <month> <day> are the date when the station coordinates file entry was set, e.g. the date for which the gain is valid. The idea is that in the future Early-est will re-check gains every now and then, if a web-service providing station coordinates is available.



## ***Channel gain file (e.g. gainfile.csv)***

The channel gain file can be created off-line, the format is:

```
AK ATKA -- BHZ 2014 14 4.887423e+08 0.150000 3
```

```
<net> <sta> <loc> <chan> <year> <day of year> <gain> <freq> <type>
```

where

<year> <day of year> are the date when the gain file entry was set, e.g. the date for which the gain is valid. The idea is that in the future Early-est will re-check gains every now and then, if a web-service providing gains is available.

<gain> is gain in <data count units> / <physical ground motion units>

<freq> is frequency for gain (should be 0.15Hz for Early-est, but can differ if response is flat around 0.15Hz)

<type> is type of physical ground motion represented by data relative to ground displacement:

```
#define ERROR_TYPE -1
#define UNKNOWN_TYPE 0
#define SCALING_TYPE 1
#define INTEGRAL_TYPE 2
#define DERIVATIVE_TYPE 3
#define DOUBLE_DERIVATIVE_TYPE 4
#define OTHER_CONVERSION_TYPE 5
#define NO_CONVERSION_TYPE 6
```

so, for example, velocity data has type DERIVATIVE\_TYPE 3

Early-est currently only accepts velocity data with DERIVATIVE\_TYPE 3 for all magnitude and other processing that requires true ground motion scaling.

## ***Station/Channel Metadata from web-services***

Station meta-data for coordinates and gain can be read at start-up and dynamically from FDSN Station and IRIS Station and Response web-services. The parameters for accessing these web-services are specified on the command-line:

```
-g geogfile      file with geographic coordinates of stations (Network Station
  Latitude Longitude Elevation(m))
-sta-query-type type of query station web service (e.g. FDSN_WS_STATION,
  IRIS_WS_STATION)
-sta-query-host  host name for station web service (e.g. www.iris.edu)
-sta-query       query root for station web service (e.g. ws/station/query)
  -sta-query-type, -sta-query-host and -sta-query can be repeated in
  corresponding triplets

-pz gainfile     file with gain info for channels (Network Station Location Channel
  Year, DayOfYear, Gain, Freq, Type)
```

- pz-query-type type of pole-zero response (SEED-RESP format) web service (e.g. FDSN\_WS\_STATION, IRIS\_WS\_RESP)
  - pz-query-host host name for pole-zero response (SEED-RESP format) web service (e.g. www.iris.edu)
  - pz-query query root for pole-zero response (SEED-RESP format) web service (e.g. ws/resp/query)
- pz-query-type, -pz-query-host and -pz-query can be repeated in corresponding triplets

## Output Files

All output for *miniseed\_process* is written in sub-directories under the base directory <file>.out/, where <file> is the Mini-SEED file argument to *miniseed\_process*.

All output for *seedlink\_monitor* is written in sub-directories under the base directory

<outpath>/<year>/<month>/<day>/<year>.<month>.<day>.<hour>.<min>.<sec>/

where <outpath> is specified by the -o argument to *seedlink\_monitor* and <year>, <month>, etc. are the time of generation of a report (generated every 1 min).

The directory <base\_directory>/ contains csv, html, xml and other files describing the picks, hypocenters, magnitudes and tsunamigenic indicators, station health, and miscellaneous plotting information.

The subdirectory <base\_directory>/plots contains GMT format data files for psimage, psxy, pshistogram and other plotting.

If plot\_warning\_report\_seedlink\_runtime.bash is run unmodified, the directories specified by SAVE\_DIR and LOCAL\_DIR\_WARNING will each contain many of the files from <base\_directory>/ along with jpg and/or pdf files generated by GMT; the directory specified by SAVE\_DIR/events and LOCAL\_DIR\_WARNING/events will each contain, for each associated event, data files for mechanism determination, event html an with jpg and/or pdf files generated by GMT.

**IMPORTANT NOTE: the *seedlink\_monitor* output can total around 1Gb/day – you are responsible for regular monitoring and managing (archived, moved, deleted, ...) of the output directories to avoid using excessive space on the output file system or device.**

## Algorithm Description

The Early-est programs *miniseed\_process* and *seedlink\_monitor* read mini-seed data packets from file or a SeedLink server, respectively, and pass each packet to a *trace-processing* module. These programs also call an *associate/locate - reporting* module at regular *reporting intervals* (e.g. after all data is read for *miniseed\_process*; every 1 min for *seedlink\_monitor*), and perform general initialization, quality control and cleanup operations. Figures 1 and 2 show schematic diagrams of for the two programs and the main modules.

The Early-est software maintains a persistent *pick list* for the current *reporting window* (e.g. a specified time window for for *miniseed\_process*; the last hour before real-time for *seedlink\_monitor*) and an *event list* for a specified archive interval (e.g. the last 10 days). The *pick list* is updated continuously as picking and trace processing are applied to new data packets. The *event list* is updated at each *reporting interval* as new event locations are found or previous locations are deleted. At each *reporting interval* the *associate/locate* module processes the current *pick list* from scratch, without making use of previous associations or location information from the *event list*; this memory-less procedure simplifies the *associate/locate* module and makes it very robust with respect to changes in the *pick list*, but increases the computational load.

## ***Trace-processing module***

The ***trace-processing*** module (Figure 1) processes each new data packet passed by ***miniseed\_process*** or ***seedlink\_monitor***. This processing includes channel identification, quality control, filtering for picking, picking using ***FilterPicker*** and further filtering and pre-processing as required for seismic and tsunamigenic event characterization (Table 1).

***FilterPicker*** (Lomax et al., 2011; Vassallo et al., 2011) is a general purpose, broad-band, phase detector and picker which is applicable to real-time seismic monitoring and earthquake early-warning. ***FilterPicker*** uses an efficient algorithm which operates stably on continuous, real-time, broadband signals, avoids excessive picking during large events, and produces onset timing, realistic timing uncertainty, onset polarity and amplitude information. In practice, it operates on a pre-defined number of frequency bands by generating a set of band-passed time-series with different center frequencies. Characteristic functions are determined for each frequency band and a pick is declared if and when, within a window of predefined time width, the integral of the maximum of the characteristic functions exceeds a pre-defined threshold.

After picking on each new data packet, for each pick in the ***pick list*** for the current packet channel, the ***trace-processing*** module applies various analyses on the channel data and updates values needed for event characterization. Recursive, time-domain algorithms are used for all filtering and other time-series processing.

## ***Associate/locate - reporting module***

The Early-est ***associate/locate - reporting*** module (Figure 1) calls the ***oct-tree associate/locate*** module with the current ***pick list***, and then the ***reporting*** module which determines event characterization results and generates graphical and alpha-numeric reporting output.

## ***Oct-tree associate/locate module***

The ***oct-tree associate/locate*** module (Figure 2) efficiently and robustly associates picks, and detects and locates seismic events over the whole Earth from 0 to 700km depth using the efficient, non-linearized, probabilistic and global, oct-tree importance-sampling search (Lomax and Curtis, 2001; Lomax et al., 2009). The objective function for the oct-tree search is a probability function,  $P(\mathbf{x})$ , based on stacking of implicit origin-times for each pick for each potential source  $\mathbf{x}_{test}$ : given a seismic wave velocity model (currently ak135 (Kennett et al., 1995)), a pick time  $t_p$  at a seismic station, and assuming a seismic phase type that may have produced the pick, the phase travel-time from source  $\mathbf{x}_{test}$  to station  $T_x$  can be calculated and thus the implicit origin-time  $T_0$  for the source and phase can be determined by back projection (e.g.,  $T_0 = t_p - T_x$ ). The set of stacks of  $T_0$  for all picks forms a histogram over potential origin-times for a source at  $\mathbf{x}_{test}$ . If the maximum histogram value exceeds a specified threshold, and if the associated picks for the maximum pass tests on amplitudes and station distributions, then  $P(\mathbf{x}_{test})$  is retained to drive further the oct-tree search to find a maximum  $\mathbf{x}_{max} = \max[P(\mathbf{x})]$  and define a seismic event at  $\mathbf{x}_{max}$  and associated picks.

The oct-tree search is *direct* and non-linearized – it does not involve linearization of the equations relating the pick times to the source location, and is *global* and *probabilistic* – it samples throughout the prior probability density function (PDF) for the seismic location problem. The search uses an initial, coarse, regular grid-search followed by recursive, octal sub-division and sampling of cells in three-dimensional, latitude/longitude/depth space to generate a cascaded, oct-tree structure of sampled cells. The oct-tree search produces approximate *importance-sampling* - the spatial density of sampled cells follows the objective function  $P$ .

For each latitude/longitude/depth cell of volume  $v$  visited by the oct-tree search, a histogram-like stack over implicit origin-times for first-arrival, P phases (currently Pg, P, Pdiff, PKPdf) for all picks in the

*pick list* is constructed. Each origin-time value  $T_0$  is assigned a distance and pick-quality weighted amplitude  $A$  between 0 and 1.0, and an uncertainty  $\sigma$  determined by the sum of half the maximum travel-time range across the cell volume with the travel-time and pick uncertainties. Each implicit origin-time is included in the origin-time stack with amplitude  $A$  using two step-function time-limits at  $T_0 \pm \sigma$  inserted in time order. After all picks have been processed, the maximum of the origin-time stack is found by a systematic scan over the available time-limits; the use of step-function time-limits and time ordering makes this scan very fast. All picks whose origin time-limits overlap the stack maximum time are flagged as associated. The stack value, combined with the variance of the implicit origin-times from all associated picks, is converted to a probability,  $P(\mathbf{x}, v)$ . If the maximum stack value exceeds a specified threshold (currently 4.5), and if the associated picks for the maximum pass tests on amplitude attenuation, and station distance and azimuth distributions, then  $P(\mathbf{x}, v)$  is stored for use in the progression of the oct-tree search. If any of these conditions are not met, then the *oct-tree associate/locate* module returns, with a flag that no event has been associated.  $P(\mathbf{x}, v)$  represents the relative probability that an event is located within a cell of volume  $v$  at position  $\mathbf{x}$ .

The oct-tree search to associate / locate is paused when the subdivided cells reach an adaptively determined, minimum size (e.g.  $\leq 5$ km for a location constrained by regional to globally distributed stations,  $\leq 1$ km for a location constrained by locally distributed stations); at this pause uncertainty measures (e.g. PDF scatter samples) on the association stage are generated. The oct-tree search and cell subdivision is then continued for a fixed number of samples (currently about 4600) to obtain a refined, precise location by fixing the associated phases to those corresponding to the maximum of the  $P(\mathbf{x}, v)$  found in the association stage. The fixing of the associated phases is necessary for small cell sizes since a decreasing cell volume combined with the step-function limits on origin-time leads to a continuous reduction in  $P(\mathbf{x}, v)$  values and eventual instability and non-convergence of the oct-tree search near and at the optimal source location. The precise oct-tree results provide uncertainty measures (e.g. PDF scatter samples, uncertainty ellipsoid) on the location.

When the *oct-tree associate/locate* module returns an event, the associated picks for this event are masked in the *pick list* and the *oct-tree associate/locate* module is called again using the remaining, non-associated picks, until no further events are returned. Thus multiple events can be associated and located within a report interval, and, in general, the events are identified in order of the number of associated picks and better location constraint.

Early-est runs the *oct-tree associate/locate* module every 1 minute using all picks from the past hour, without knowledge of or preserving information from previously associations and event locations. This procedure makes Early-est relatively simple algorithmically and robust with regards to changes in the set of available picks and the number of associated picks defining locations. In particular, this procedure allows early stage locations with few associated picks to easily move in space or origin time, or to split into multiple events, or to be absorbed into other events, or to disappear as more pick data becomes available. However, this procedure is inefficient for later stage event locations which are defined by a larger number of associated picks, e.g. more than 10-20 picks, since such locations are very unlikely to change; much processing effort is repeated each minute to re-obtain a previous result. This inefficiency can be problematic after large earthquakes, when the repeated re-processing of hundreds of picks from a mainshock and large aftershock can cause Early-est to fall behind real-time.

## **Reporting module**

The *reporting* module (Figure 1) processes the current *pick list* and *event list* to determine event characterization results (Table 1) and generate graphical, alpha-numeric, XML, HTML and other reporting output for events, picks, stations, etc. An e-mail alert message (Figure 3) is generated (see `-sendmail` command line option) for each event with magnitudes or tsunamigenic potential exceeding pre-set thresholds (Table 2); the email message body is in file `<ALERT_HYPO_FILE>` in the output directory. A system call to a script in the working directory `'./run_action_alert_sent.bash <ALERT_HYPO_FILE> <ALERT_EVENT_PAGE_URL>'` (e.g. for Twitter messaging) is also called

whenever an e-mail alert is sent.

The module then runs an external script specified by the `-c` command line option (currently `-c ./plot_warning_report_seedlink_runtime.bash`) to run HASH, fmamp, generate GMT graphics and final, alpha-numeric, XML, HTML and other reporting output, and to manage output file conversions, copying and cleanup. Figure 4 shows the main graphical display of Early-est, which summarizes the evolving trace-processing, associate/locate and event characterization results in real-time. Figure 5 shows some of the browser-based HTML output of Early-est, which details the evolving Early-est processing and results in real-time.

### ***Computational efficiency***

On a 3.33GHz Intel Core 2 Duo iMac with OS X 10.6, the *oct-tree associate/locate* module for a single event of around M5-6 (20-100 associated picks) takes about 1sec and the *reporting* module about 5-10sec (primarily the GMT graphics and HASH search). For *seedlink\_monitor*, collecting data packets from the SeedLink servers for about 200 stations and processing the data in the *trace-processing* module takes very little CPU time, 2-3% CPU on average.

To investigate efficiency for large event sequences, we consider two, hour-long windows of data at stations at 0-180° distance for the 2011.03.11 Mw9 Tohoku, Japan earthquake, one with about 200 stations (run\_MED) and a second with about 400 stations (run\_BIG). Running *miniseed\_process* off-line with run\_MED or run\_BIG associates and locates the mainshock and 9 aftershocks (the association/location is currently limited to 10 events maximum in a report window). On the above iMac the *oct-tree associate/locate* module takes about 30 sec for run\_MED and about 1.5 min for run\_BIG. Additionally, the report and event processing for both runs for 10 events (HASH, event graphics, event HTML, ...) takes about 30 sec (though this runs in a separate process from the data processing and associate/locate).

Thus for run\_MED (similar to the current Early-est running at INGV) the processing will approximately keep up with real-time for a mega-event, teleseismic sequence like 2011 Tohoku, Japan, with the latency for display of the results increasing from the usual ~10s to ~1 min. However, the *associate/locate - reporting* for run\_BIG takes over 2 min, much longer than the reporting interval of 1 min currently used in *seedlink\_monitor*; to avoid that the processing and reporting fall behind realtime in this situation, Early-est will increase the reporting interval. In summary, Early-est in its current form can be used on a standard PC/Mac in real-time for up to around 200-300 channels and up to about 10 events in a report window with report interval 1 min.

**Table 1 - Event characterization measures used in Early-est**

<i>Measure</i>	<i>References</i>	<i>Description, modifications</i>
Td	Lomax and Michelini (2011)	Maximum dominant period (smoothed over 5s) in window from P arrival time (Tp) to Tp + 55s.
T50Ex	Lomax and Michelini (2011)	T50 Exceedance, modified as follows: 1) Reduced T50Ex minimum distance to 5deg.
Td*T50Ex	Lomax and Michelini (2011)	Period-duration discriminant for tsunami potential, modified as follows:  1) Reduced Td*T50Ex minimum distance to 5deg.
To	Lomax and Michelini (2009ab, 2011)	High-frequency, apparent source-duration, modified as follows: 1) Removed smoothing window width of 10s from To for short durations; applied with a linear ramp from 10 → 0s for initial durations of 20 → 60s, minimum duration is highest frequency in HF stream (0.2s). 2) Reduced To minimum distance to 5deg. 3) Added reference of To duration to S arrival time (Ts) if raw duration end time (To_end) is after Ts (e.g. if To_end > Ts+(Ts-Tp)/3 then To=To_end-Ts).
mb(Vmax)	Bormann and Saul (2008, 2009)	mb body wave magnitude using Vmax formulation:  1) Apply to BRB velocity a recursive, time-domain filter that implements the WWSSN-SP displacement response:  * WWSSN-SP displacement response from Working Group on Magnitudes (Magnitude WG) of the International Association of Seismology and Physics of the Earth's Interior (IASPEI) Commission on Seismological Observation and Interpretation (CoSOI) 2011 ZEROS 3 POLES 5 -3.72500 -6.22000 -3.72500 6.22000 -5.61200 0.00000 -13.2400 0.00000 -21.0800 0.00000 CONSTANT 1.0  This filter is applied to the BRB velocity, so effectively gives: integrate → simulate the WWSSN-SP response → differentiate, without doing the integration and differentiation.  2) Measure Vmax - the peak from Tp to the lesser of Tp+To or Tp+30s,  3) Apply: $mb(V_{max}) = \log_{10} (V_{max}/2\pi) + Q(\Delta, h)$

<i>Measure</i>	<i>References</i>	<i>Description, modifications</i>
Mwp	Tusboi et al. (1995, 1999)	Mwp magnitude, modified as follows: 1) Applied from Tp to the lesser of Tp+To or Tp+120s.
Mwpd(RT)	Lomax and Michelini (2009a)	Mwpd duration-amplitude, large earthquake magnitude, modified as follows to allow simple and robust real-time application without event type determination:  1) Use constant from Tusboi et al. (1995, 1999) = 4.213e19. Use PREM depth correction; no geometrical spreading or attenuation corrections.  2) Moment correction of Lomax and Michelini (2009a) applied to all event types if To>80s.  3) Reduced Mwpd minimum distance to 5deg (Stable since added reference of To duration to Ts).
Focal mechanism	(Hardebeck and Shearer, 2002)	P-arrival, first-motion focal mechanism using the HASH program.
Focal mechanism	(Lomax, 20XX)	Probabilistic, P-arrival, first-motion and amplitude focal mechanism algorithm (fmamp). Uses oct-tree search; solution quality based on wighted distribution (quasi-pdf) of P and T axes. (Note: Under development; included for information in Early-est distribution.)

**Table 2 – Default alert message thresholds**

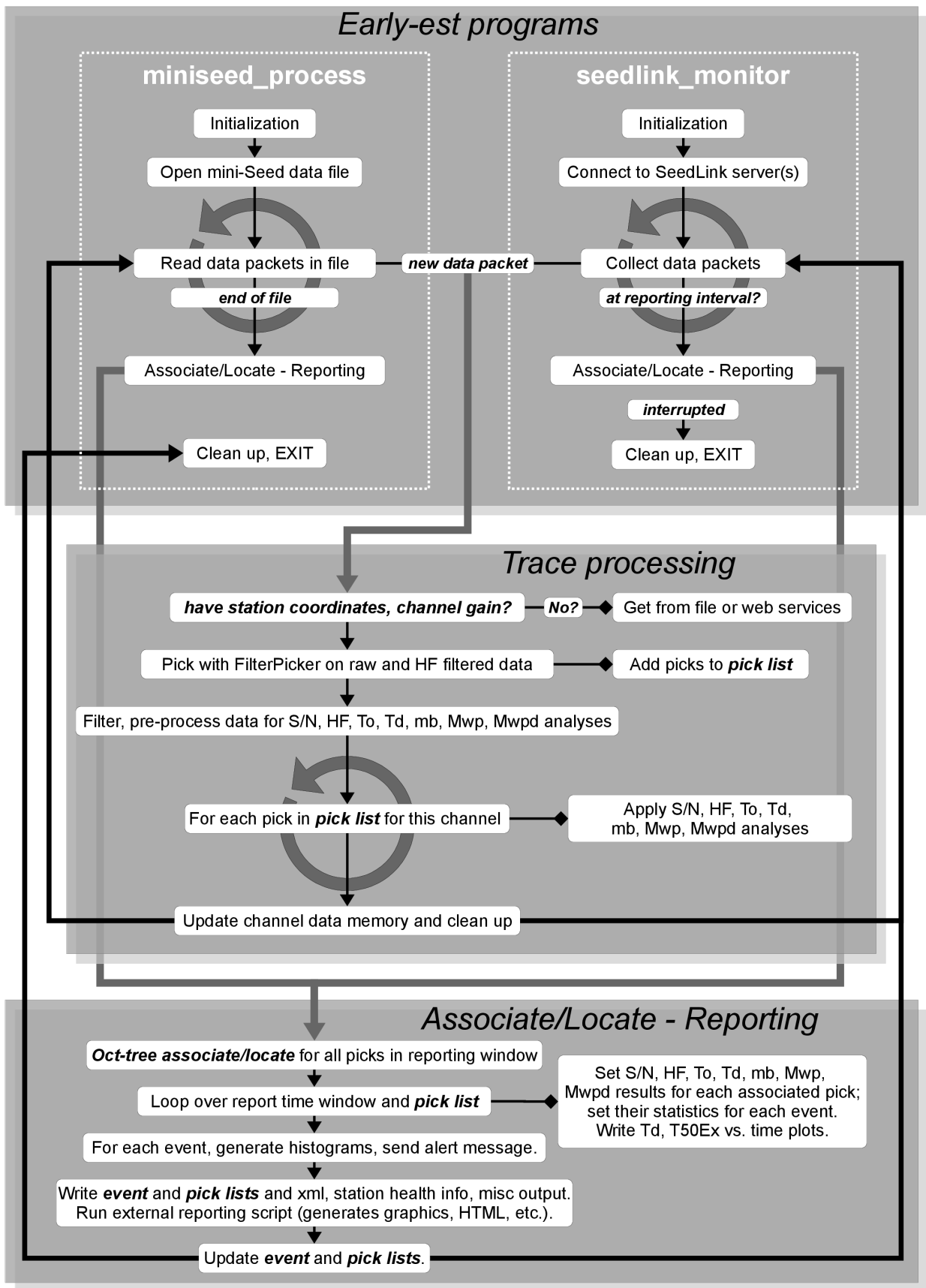
<i>Measure</i>	<i>Min num readings for alert</i>	<i>Min value for alert*</i>	<i>Min change in value for new alert**</i>
mb(Vmax)	4	5.5	+0.5
Mwp	4	5.8	+0.5
Mwpd(RT)	4	6.95 and Mwp≥6.95	+0.5
Td*T50Ex	4	7.2	+8.0 or change from >8.8 (RED) to <7.2 (GREEN)

\* can be set in application properties file

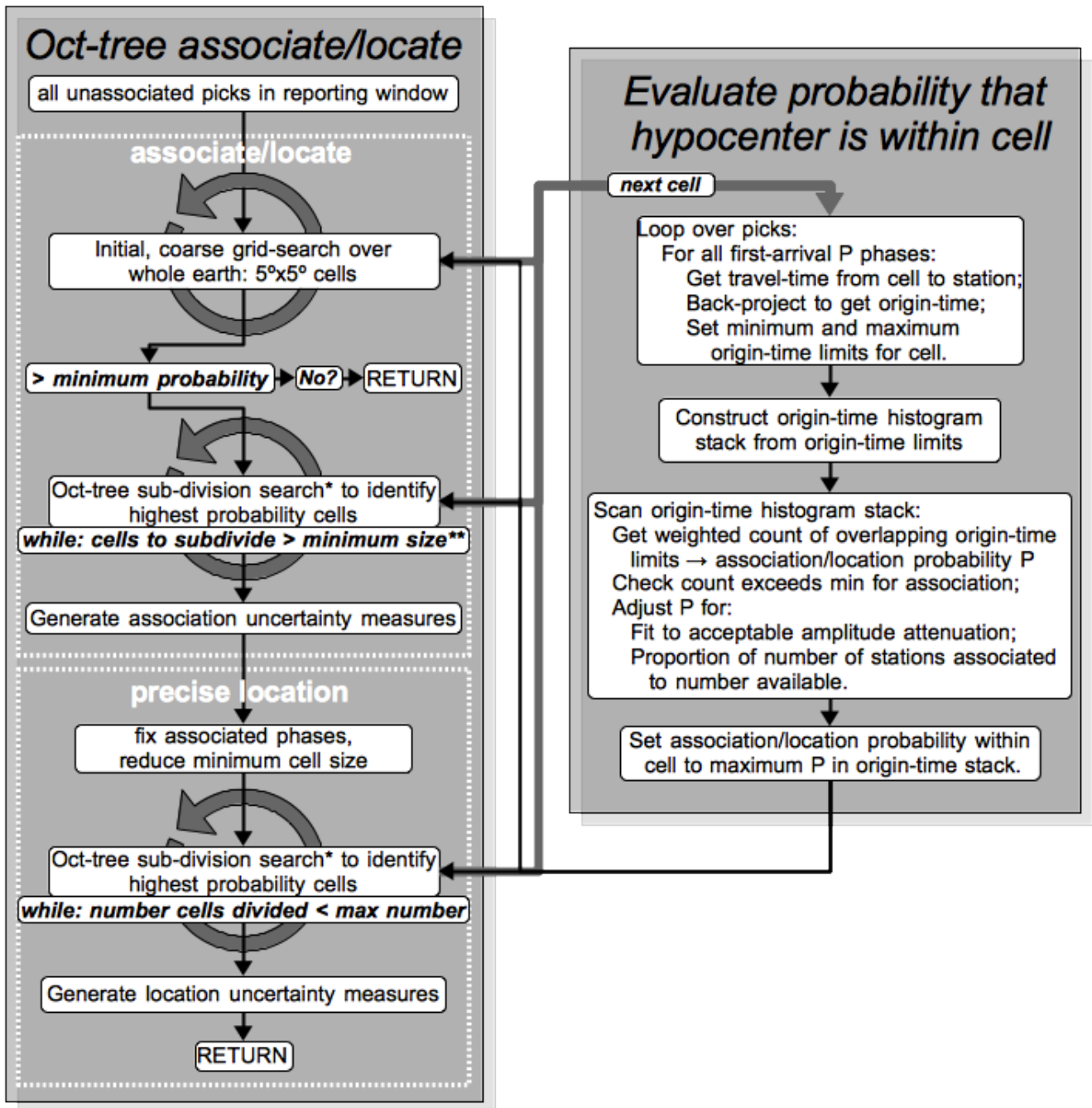
\*\* if an alert has been sent, a new alert is also sent 10min after event origin-time



**Figure 1. Early-est Schematic Flow-Diagram**



**Figure 2. Early-est Associate/Locate Flow-Diagram**



\* cell division is performed at a fixed cell size for a specified number of cells or until no cell available to divide; the fixed cell size is then reduced and cell division continued

\*\* minimum size is adaptively reduced in proportion to number of associated stations near epicenter

### Figure 3

Figure 4

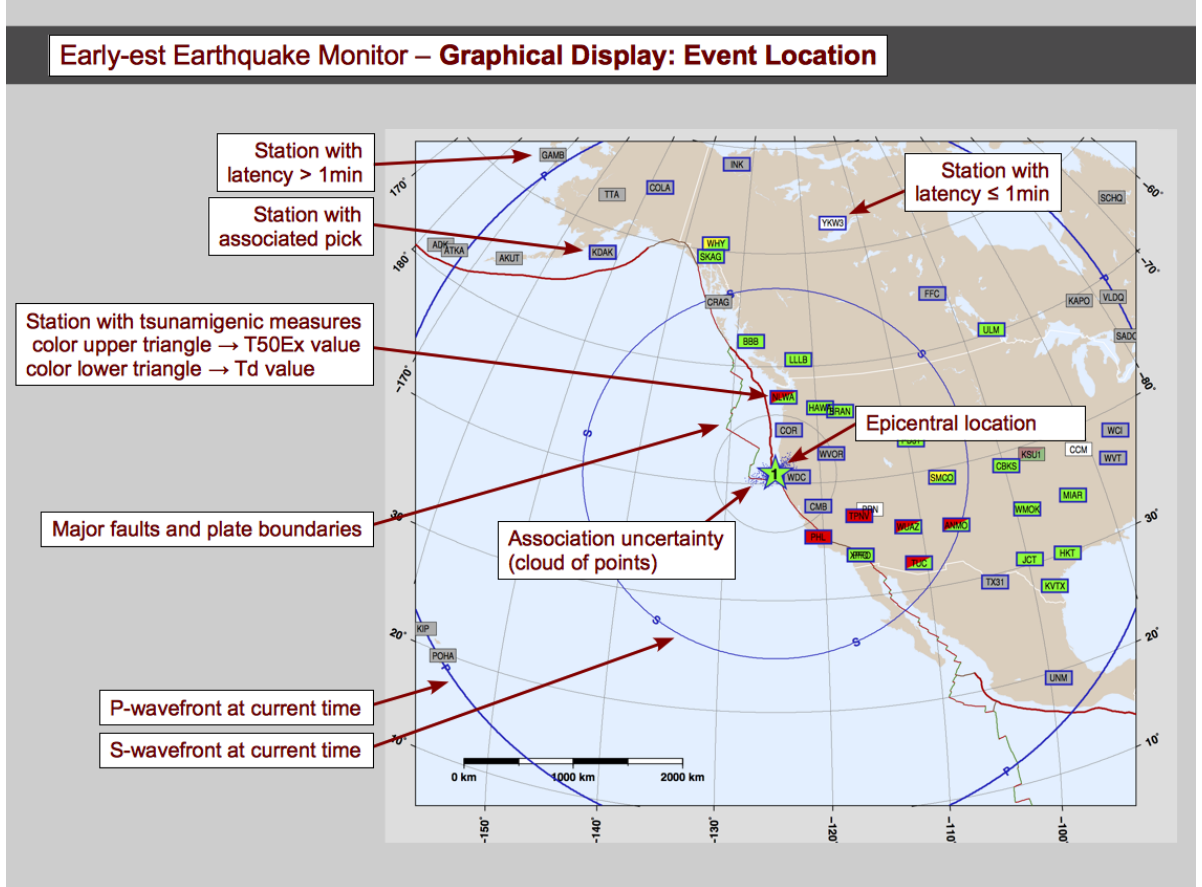
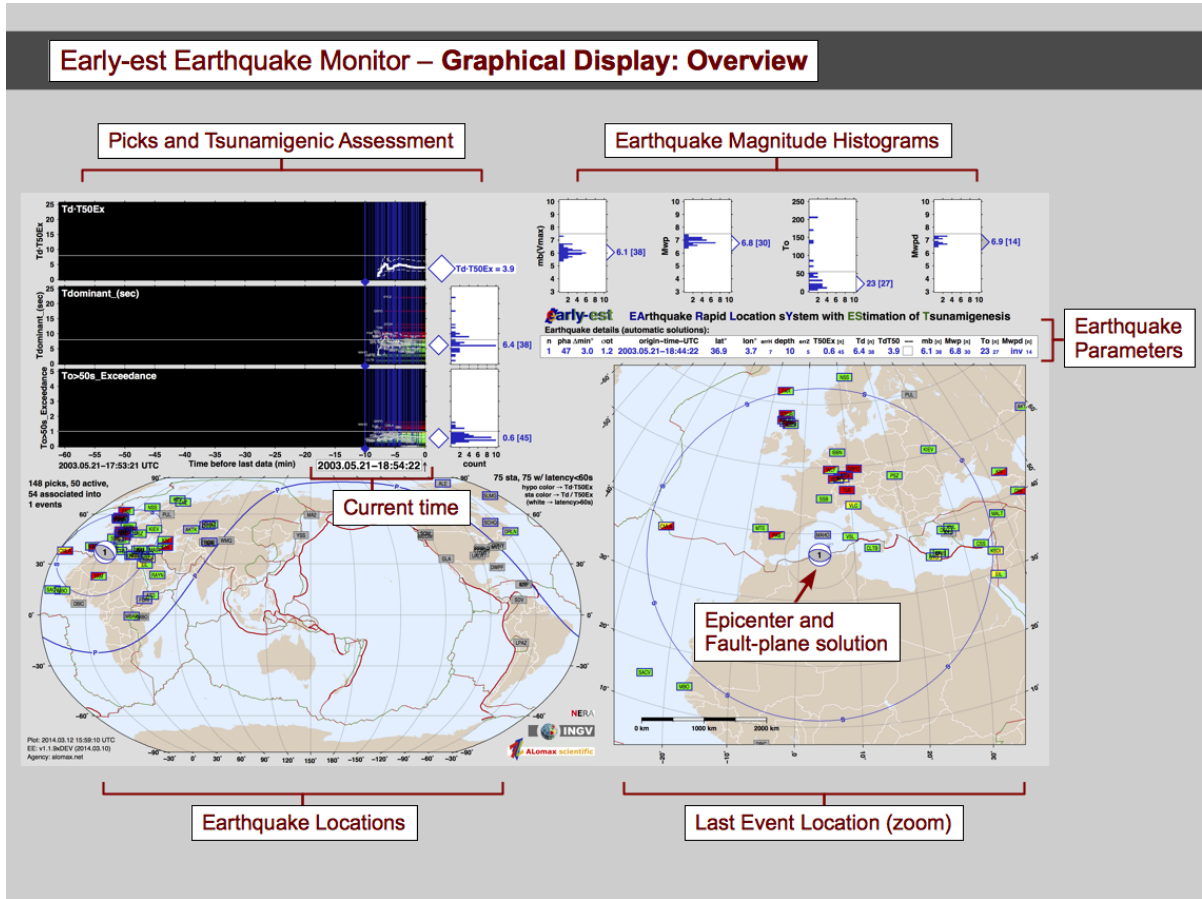
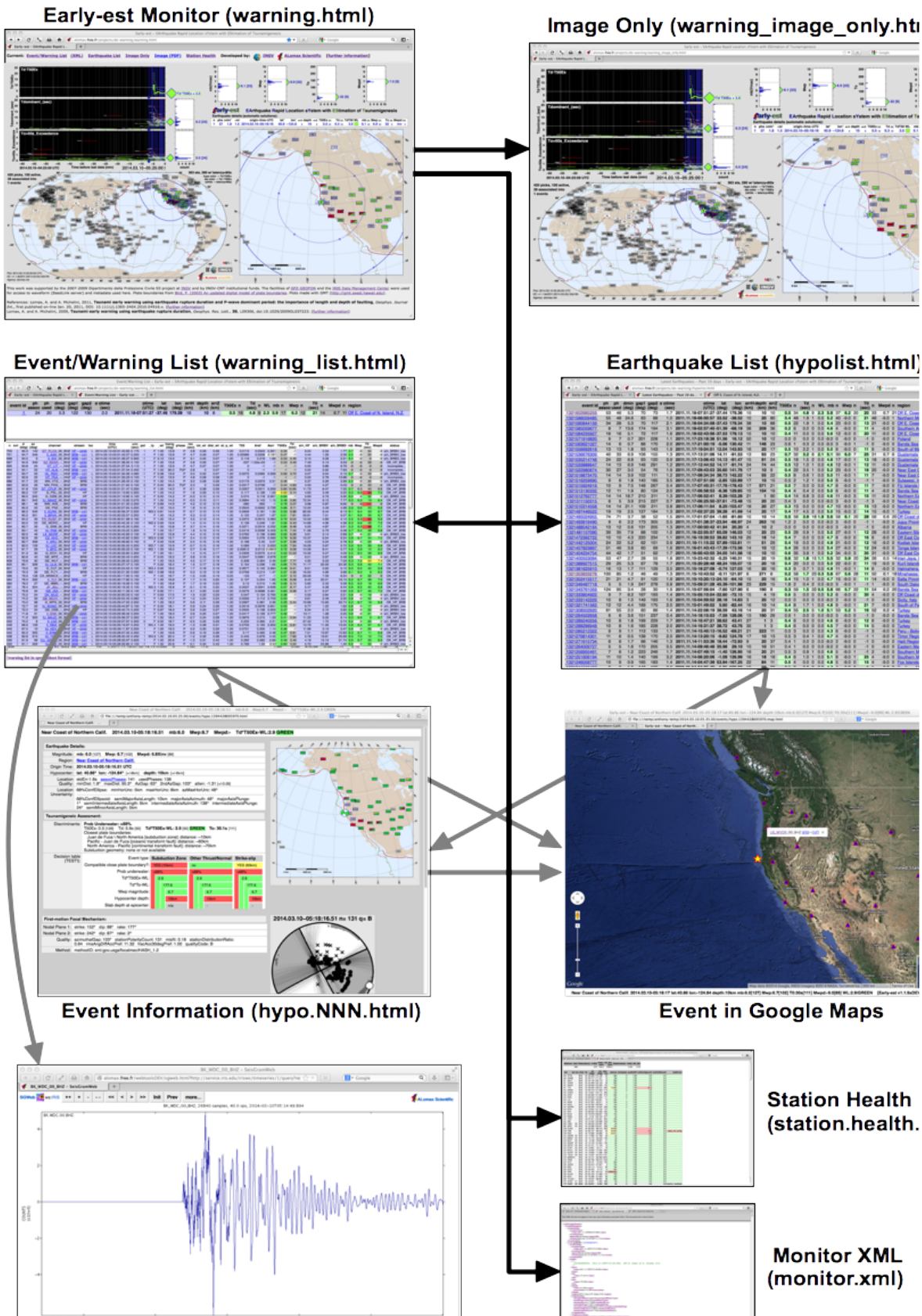


Figure 5



## References

- Bernardi, F., A. Lomax, A. Michelini, V. Lauciani, A. Piatanesi, and S. Lorito (2015). Appraising the Early-est earthquake monitoring system for tsunami alerting at the Italian candidate Tsunami Service Provider, *Nat. Hazards Earth Syst. Sci.*, **15**, 1–40, doi:10.5194/nhessd-15-1-2015
- Bormann, P. and J. Saul (2009). Earthquake magnitude, in *Encyclopedia of Complexity and Systems Science*, p. 10370, ed. Meyers, A., Springer, New York, doi:10.1007/978-0-387-30440-3\_151.
- Bormann, P., and J. Saul (2008). The new IASPEI standard broadband magnitude mB. *Seismological Research Letters* **79** (5), 698–705; doi: 10.1785/gssrl.80.5.698
- Hardebeck, Jeanne L. and Peter M. Shearer, (2002). A new method for determining first-motion focal mechanisms, *Bulletin of the Seismological Society of America*, **92**, 2264-2276.
- Kennett, B. L. N., E. R. Engdahl, and R. Buland (1995). Constraints on seismic velocities in the Earth from travel times, *Geophys. J. Int.* **122**, 108–124.
- Lomax, A., and A. Curtis (2001). Fast, probabilistic earthquake location in 3D models using oct-tree importance sampling, *Geophys. Res. Abstr.* **3**, 955.
- Lomax, A. and A. Michelini (2012). Tsunami early warning within 5 minutes, *Pure and Applied Geophysics*, **170**, 1385-1395, doi: 10.1007/s00024-012-0512-6.
- Lomax, A. and A. Michelini (2011)., Tsunami early warning using earthquake rupture duration and P-wave dominant period: the importance of length and depth of faulting, *Geophys. J. Int.*, **185**, 283–291, doi: 10.1111/j.1365-246X.2010.04916.x.
- Lomax, A. and A. Michelini (2009b). Tsunami early warning using earthquake rupture duration, *Geophys. Res. Lett.*, **36**, L09306, doi:10.1029/2009GL037223.
- Lomax, A. and A. Michelini (2009a). Mwpd: A Duration-Amplitude Procedure for Rapid Determination of Earthquake Magnitude and Tsunamigenic Potential from P Waveforms, *Geophys. J. Int.*, **176**, 200-214, doi:10.1111/j.1365-246X.2008.03974.x
- Lomax, A., A. Michelini, A. Curtis (2009). Earthquake Location, Direct, Global-SearchMethods, in *Encyclopedia of Complexity and Systems Science*, Part 5, 2449-2473, ed. Meyers, A., Springer, New York, doi: 10.1007/978-0-387-30440-3\_150.
- Lomax, A., C. Satriano and M. Vassallo (2012). Automatic picker developments and optimization: FilterPicker - a robust, broadband picker for real-time seismic monitoring and earthquake early-warning, *Seism. Res. Lett.* , **83**, 531-540, doi: 10.1785/gssrl.83.3.531.
- Tsuboi, S., Abe, K., Takano and Y. Yamanaka, (1995). Rapid determination of Mw from broadband P waveforms, *Bull. seism. Soc. Am.*, **85**, 606–613.
- Tsuboi, S., P.M. Whitmore and T.J. Sokolowski, (1999). Application of Mwp to deep and teleseismic earthquakes, *Bull. seism. Soc. Am.*, **89**, 1345–1351.
- Vassallo, M., C. Satriano and A. Lomax, (2012). Automatic picker developments and optimization: A strategy for improving the performances of automatic phase pickers, *Seism. Res. Lett.* , **83**, 541-554, doi: 10.1785/gssrl.83.3.541.

---

## Revision History Notes

- v0.1.0 2009-09-24 First stable version
- v0.1.1 2010-01-14 Replaced TestPicker4\_1 with FilterPicker5
- v0.2.1 2010-03-30 Added Td (dominant period) and Td\*T50Ex warning, broadband picking, multi-phase association, event persistence, Mwp, higher precision association/location, and much more.
- v0.2.2 2010-06-16 Added to event association checks on gap, secondary gap and percentage of available stations associated within distance of farthest associated station, support for multiple SeedLink servers, and much more.
- v0.2.3 2010-06-20 Bug fixes and improvements to association/location reporting and weight count

- grid.
- v0.2.4 2010-06-21 Bug fixes and improvements to association/location reporting and weight count grid.
  - v0.2.5 2010-07-05 Changed picking from BRB to HF and BRB, with BRB picks used for association/location only if there is an HF pick within the following 30 sec; improvements to association/location reporting and weight count grid.
  - v0.2.6 2010-11-09 Bug fixes. Changed picking to new version of FP, with changes/improvements to pick uncertainty settings, this may give earlier pick times.
  - v0.2.7 2010-12-06 Changed -c report command call to background to avoid delays to main trace processing. Changed HF picking tUpEvent from 4 to 2sec.
  - v0.3.0 2011-03-04 Added oct-tree association/location, mB, T0, Mwpd; extend output and graphics to support previous. Added simple e-mail alert messages, automatic retrieval of station coordinates from web services, XML format output, NLL format phase output for associated phases for each active location. Various bug fixes, improvements and modifications to association/location algorithm and general program functionality.
  - v0.3.1 2011-03-09 Fixed bug of writing to previously closed file.
  - v0.3.2 2011-03-31 Modified T0 algorithm to avoid erroneously long T0 durations (and large Mwpd) when P waves from a later event arrive soon after a pick. Changed Mwpd algorithm to use moment correction as in Lomax & Michelini 2008. Modified wavefront sampling algorithm to try to avoid GMT/Postscript plotting failures. Remove plotting of wavefronts and miscellaneous minor changes to plot\_warning scripts. Fixed bug in setting start index before P for mB, Mwp, Mwpd calculations. Modified SL collect loop in seedlink\_monitor.c to prevent cumulating large delays (multiple sleeps of 0.5s) in sl\_collect\_nb if a SL conn is down.
  - v0.3.3 2011-04-01 Fixed bug in setting memory array size. Fixed bug in storing trace data before P for BRB rms and mB, Mwp, Mwpd calculations.
  - v0.3.4 2011-04-11 Fixed bug in catching SL packet header error: "record header/quality indicator unrecognized". Added check to avoid processing of same event twice and possible loss of instance with greater associated P phases.
  - v1.0.0 2011-05-13 Changed check for same event (location.c:isSameEvent()) to use 4\*sigma\_otime difference and 4\*errH epicenter difference, more accurate epicentral distance difference calculation, with smaller otime difference allowing large epicentral difference and vice-versa; merged pickdata from same event with fewer associated P phases into other event, merged phases indicated by new 'M' merged flag to 'loc' column in pick tables. New check to ignore packets that have data end earlier than beginning of report window + report interval; prevents creation of duplicate events for events whose OT has passed before report window and whose associated phase data has been removed.
  - v1.0.1 2011-11-18 Replaced approximate travel-time estimation with accurate travel-time interpolated from table using 2D Lagrange interpolation. Corrected potential problems with mB and Mwp peak amplitude determination, depending on termination packet and timing of To estimation. Fixed minor bugs in T0, mB, Mwp, Mwpd realtime indexing. Added ray take-off angles. Added P first-motion mechanism calculation using USGS-HASHv1.2. Added bash and Python codes to generate event and mechanism information pages, and to generate FocalMechanism QuakeML for mechanisms. Replaced mB magnitude determination with mb(Vmax). Changed Mwpd calculation to report evolving Mwpd magnitude before final To available; limited Mwpd integration window

width to S-P time; changed  $T_0$  and  $M_{wpd}$  minimum distance to 15deg. Improved efficiency of inner-loop distance and azimuth checks. Changes to assoc/loc critical node size test algorithm - now compares node size to epicentral distance containing a specified number of associated stations. Changes to oct-tree probability factors (PROB\_FACTOR\_CONVERGENT, PROB\_FACTOR\_SAMPLE) for minimum probability to divide; changes to calculation of pdf scatter sample for association and for location.

- v1.0.2 2011-11-24 Added audio alert option for warning image-only HTML page.
- v1.0.3 2011-12-06 Increases phase uncertainty limits by x2 to avoid missed associations for PKP and other phases. Added check to sendmail parameters to avoid Segmentation Fault crashes. Minor additions and modification to main monitor graphic.
- v1.0.4 2011-12-12 Split event processing with processEvents.py into mechanism and events stages; allowed re-ordering of processing in plot\_warning\_report\_seedlink\_runtime.bash so that main monitor display with focal mechanism is completed first and rapidly, while slower event processing is done last.
- V1.1.0 2011-12-19 Changed PKIKP phase (not supported in iaspei-tau) to PKiKP, added new secondary phases and reduced P error from 4.0s to 1.0s to help avoid false events on PKP, PKiKP, etc. Current supported phase set: P, Pdiff, Pg, PKiKP, PKPab, PKPbc, PKPdf, PKKPab, PKKPbc, PKKPdf, PKPPKPab, PKPPKPbc, PKPPKPdf, PP, S, Sdiff, Sg, SKSac, SKSdf, SKPdf, pP, sP, PcP, ScP, ScS, PKKSab, PKKSbc, PKKSdf, SKKSac, SKKSdf.
- V1.1.1 2011-12-21 Fixed bug in S time calculation for limiting  $M_{wpd}$  integration window to S-P time. Added check if total weight of unique associated stations/phases is greater than min\_weight\_sum\_assoc to avoid association on multiple picks from same station.
- V1.1.2 2012-03-07 Removed smoothing window width from  $T_0$  for short durations (<60s). Reduced  $T_{50Ex}$ ,  $T_{50Ex} * T_d$ ,  $T_0$  and  $M_{wpd}$  minimum distance to 5deg; added reference of  $T_0$  duration to S arrival time if duration overlaps S time (e.g.  $T_0 = HF\_duration\_end - T_s$  if  $T_p + T_{0\_est} > T_s + (T_s - T_p) / 4$ ). Added COUNT\_IN\_LOCATION | FIRST\_ARRIVAL\_P for PKPab. Improved checking processing for clip and non-contiguous (gap or overlap) data. Truncated  $T_d$  estimation to maximum time-window of  $P + T_0$ . Increased P error from 1.0s to 2.0s; set SMALL, MED, LARGE phase errors to 4, 6, 6 sec, respectively (see ttimes.c). Modified monitor.xml formatting to better correspond to QuakeML standard. Added channel links to SeisGramWeb and IRIS Metadata on Event/Pick List HTML page (warning\_list.html); the activation of these links depends on the SeedLink server of the channel and is controlled by a new command-line options -timeseries-query\*.
- V1.1.3 2012-04-16 Added simple broadband, waveform first-motion estimation using sign of  $M_{wpd}$  integral at time after  $P = 1/8$   $M_{wpd}$  duration, if  $M_{wpd}$  duration > 0.5s; use FilterPicker first-motion otherwise. Fixed bug in setting station quality weight to prevent reduction in quality due to picks that will later be ignored. At  $dist < 20deg$ , use  $T_0$  and  $M_{wpd}$  for report only if  $T_{50Ex} > 1.0$  and  $\tau_{auc} > 5.0s$ ; this helps avoid unreasonably large  $T_0$  and  $M_{wpd}$  for small events.
- V1.1.4 2012-07-17 Changed duration cutoff for  $M_{wpd}$  correction from starting at unique event  $T_0$  value (100s) to ramp correction factor from 0 at 90s to 1.0 at 110s; prevents discontinuous change of  $M_{wpd}$  with respect to event  $T_0$ . Added BRB displacement (BRBD) S/N calculation using 200s RMS window; added BRBV columns in picks.csv and pickmessage.html, added count of picks with low S/N BRBD in picks.log. Changed S/N check for  $M_{wpd}$  from BRB velocity to BRB displacement; avoids overestimate of  $M_{wpd}$  when P waves are in surface wave



coda of preceding large event. S/N check for Mwp unchanged on BRB velocity.

V1.1.5 2013-04-16 Fixed bug in calculation of HASH focal mechanism: Nodal Plane 2 strike, dip and rake; were incorrect. Modified HASH inversion parameters (see `hash_driver_early_est.bash`). Modified HASH solution quality rating algorithm (see `hash_driver_early_est.f`). Added to pick list fields for: pick type (F – FilterPicker first-motion; W - broadband, waveform first-motion; numeric value is pulse width for W first-motion estimation), and pick weight.

Added additional event evaluation: Probability under-water (considers coastlines and epicentral uncertainty); Identify closest plate boundaries and their tectonic type (using Bird, 2003); Subduction zone analysis and slab depth (using Hayes and Wald, 2009; Hayes et al., 2012).

Added in event page a graphical representation of earthquake parameters, discriminants and derived quantities in a real-time, hybrid decision table/tree.

Added persistent list of associated phases for an event linked from event page.

Fixed problems in GMT plotting (corruption of plots) by setting different `GMT_TMPDIR` for each main GMT module to prevent overwriting of GMT script data for each module.

Added quality measures to HF FilterPicker and broadband waveform first-motion calculations (Either the broadband, waveform polarity or the FilterPicker polarity is used for mechanism determination, depending on the quality of the polarity measures and signal-to-noise ratios.), modified broadband waveform first-motion parameters.

Changed amplitude at pick (reference for integration) for Mwp and Mwpd integration from trace amplitude to zero; this gives more stable integration and reduces scatter and (upward) bias of station Mwp and Mwpd magnitudes.

Changed BRBV and BRBD S/N calculation to use signal level in window of length  $T_0$  after P instead of fixed window length (was 50s for BRBV S/N and 200s for BRBD S/N); this generally gives higher S/N for short duration P waves and lower S/N for long duration and large event P waves. Modified S/N calculations to use maximum window available after P for provisional S/N estimate until definitive S/N available; this avoids long delay before declaring S/N too low, this delay can allow reporting of temporary, too-high magnitudes which later will be ignored.

Added check to un-associate outlier associated stations (e.g. isolated, very distant stations or late associated with small, local event.); reduces greatly the number of false associations of later arrivals.

Removed `Pdiff` as count in location.

Fixed bug in ignoring non-contiguous data for location.

Improved and fixed minor bugs in comparing times of BRB and HF picks to set `use_for_location` for BRB picks.

Added support for a properties file to allow setting of some program parameters. This change is made primarily to enable microseismic/local/regional scale association, the default values of all properties are appropriate for the global scale. The properties file must be in the working directory with the name `seedlink_monitor.prop` or `miniseed_process.prop` for the `seedlink_monitor` and `miniseed_process` programs, respectively.

Added `-no_aref_level_check` command-line flag. Disables check of Aref level a means to ignore picks close in time and insure quality (non-overlap) of tsunami

discriminants). Instead checks for use\_for\_location picks less than assoc\_loc.min\_time\_delay\_between\_picks\_for\_location seconds (set in properties file, default = 25.0 sec) after another use\_for\_location pick for same source\_id Use this flag to support microseismic/local/regional scale monitoring where tsunami discriminants are not used.

Added phase error factor to pick weight to use for association/location. This factor = (reference\_phase\_ttime\_error / time\_errors), where reference\_phase\_ttime\_error is the minimum value of the nominal maximum acceptable travel-time error for each phase type used for location/association, and time\_errors is the travel-time plus pick error for a pick with specified phase. This change allows more balanced use of non-P phases (e.g. S) with non-zero weight for association/location. This change is made primarily to enable microseismic/local/regional scale association, but it does slightly change the pick weighting for the global scale when only P phases have non-zero weight for association/location.

Fixed bug: Calculated hypocenter to station azimuth was incorrect for some azimuth quadrants.

V1.1.6 2013-05-28 Possible fix of segmentation-fault bug (see use\_for\_location\_twin\_data in timedomain\_processing\_data.c->free\_TimedomainProcessingData() and timedomain\_processing.c->td\_process\_timedomain\_processing()).

Changed phase distance weighting (dist\_wt) from linear to geometrical decay by using square of previous dist\_wt value.

Begin modifications to code to support parameter setting with a properties file, robust association/location on a local scale, and easier use of different velocity models. Work in progress to be documented later.

V1.1.7 2014-01.08 Add support for changes to web-services at IRIS: response ws moved to <http://service.iris.edu/irisws/resp/1/>; station ws (Early-est query type IRIS\_WS\_STATION) replaced by <http://service.iris.edu/fdsnws/station/1/query> (Early-est query type FDSN\_WS\_STATION).

Added date stamp to geographic station coordinates file; date stamp is time corresponding to coordinate validity.

Added system call to a script run\_action\_\_alert\_sent.bash (e.g. Twitter messaging) called whenever an e-mail alert is sent.

BRB S/N: modified to use maximum window available after P for provisional s/n estimate; bug fix to Mwpd display in phase list.

Added properties files (seedlink\_monitor.prop for seedlink\_monitor, miniseed\_process.prop for miniseed\_process, process\_events.prop for processEvents.py. See these properties files for further documentation.

Added properties file options to enable/disable alerting and dissemination of tsunamigenic potential warning information and warning color display, and various event reporting options. See [Report] section of seedlink\_monitor.prop or miniseed\_process.prop and [EventInfo] section of process\_events.prop for options.

Added export for external processing (e.g. CMT calculation) of waveform segments in miniseed format and waveform header information in SeisGram2K ASCII format. See [WaveformExport] section of seedlink\_monitor.prop or miniseed\_process.prop for waveform export options. Waveforms are written under the waveform/ sub-directory of the program output directory.

- Modifications and bug fixes to date-stamped reading of station meta-data (station coordinates and gain) from file and web-services.
- V1.1.8 2014-01-20 Additional bug fix to date-stamped reading of station meta-data (station coordinates and gain) from file and web-services.
- Bug fix in HASH: gfortran compiler: uncert\_subs.f:130.54:  
Warning: Rank mismatch in argument 'rota' at (1) (scalar and rank-1):  
call MECH\_ROT(norm1\_avg,temp1,norm2\_avg,temp2,rota)
- V1.1.9 2015-04-24 Bug fix: matrix\_statistics/matrix\_statistics.c: CalcCovarianceSamples\*(): new algorithm which subtracts the expectation from each data value before squaring and summing, instead of correcting for expectation after summing and dividing by nSamples. Should prevent precision errors in covariances when expectation is far from coordinates origin. Bug identified by Tobias Megies. Errors in covariance will affect ellipsoid and ellipse, standard-errors (erh, erz, etc.). For teleseismic locations, errors with old algorithm could be large towards longitude +/-180deg.
- Amplitude attenuation check: Added calculation and weighting during association if the fit of the high-frequency (1-4Hz) P amplitude attenuation to a power law decay follows a theoretical decay power. The weight is related to the cumulative probability of the estimated attenuation power being between -0.5 and -3.5.
- Magnitude validation based on amplitude attenuation: Added check of channel amplitude error (ratio of observed amplitude to amplitude predicted by attenuation power law); if amplitude error > 5 or < 0.2, do not use channel magnitude estimates.
- Added persistent storage of (fixed) distance, azimuth and travel-times from each station to each cell of (fixed) initial search grid. This increases slightly program efficiency after first few association-locations.
- During oct-tree subdivision-search to divide highest probability cells: added subdivision of same-size neighbors of each cell identified to sub-divide. This addition helps to generate a smoother pdf for location and avoids missing optimal maximum likelihood locations for complex pdf's. The subdivision of neighbors has always been used in NonLinLoc.
- Excluded Pdiff phases from mb, Mwp and Mwpd measures.
- Changed selection of export waveform to be less strict so as to include all first arrival or direct P phases which can be counted in location.
- Developed prototype of new probabilistic, grid-search (oct-tree), mechanism determination code, fmamp, using first-motions and, optionally, relative P amplitudes. fmamp can be run and its results displayed along with the HASH mechanism results in the Early-est HTML event pages. Added fmamp mechanisms: FMAMP\_POLARITY, FMAMP\_AMP\_AREF (high-frequency) and FMAMP\_AMP\_MWP (broadband) to Early-est HTML event pages, HASH remains definitive (plotted) mechanism.
- Changed broadband, waveform first-motion estimation to sign of peak amplitude of Mwp integral or (moment, "WMO") if S/N of Mwp integral high enough; use FilterPicker first-motion otherwise.
- Bug fixes to make requests to multiple response or station web services.
- Updated support for displaying event information and epicenter in Google map (launched from Early-est hypocenter region-name links and in alert e-mail) - Google's map service was changed to "new Google Maps" in February 2014, this new service did not display the Early-est epicenter marker. Changed the Google map launching from a URL string (http://...) to Javascript in an event HTML web

page (hypo.<event\_unique\_id>.map.html) to correctly display an epicenter marker. The use of Javascript allows extensive and sophisticated Google map display through the Google Maps JavaScript API (<https://developers.google.com/maps/documentation/javascript/basics>). To test and demonstrate these display features, added display of epicenter uncertainty ellipsoid and pdf scatter sample, station markers with links to SeisGramWeb waveforms for event.

Added properties file options to set minimum magnitude thresholds (e.g. property [Report] alert.mb\_min) for sending alert messages. See [Report] section of seedlink\_monitor.prop or miniseed\_process.prop.

Removed references to “Warning Level” and “WL” and display warning level colors from all HTML and graphics output and alert messages.

Changed oct-tree target values (node value) from probability in cell to probability density in cell. This makes oct-tree search “standard” importance sampling, as in oct-tree for NonLinLoc, and should give more meaningful scatter samples (location uncertainties). TODO: This change may allow replacing the hierarchical searching by constant levels (oct-tree node size) with “standard”, multi-scale oct-tree search at all levels simultaneously.

Improvements and bug fixes in scatter samples (location uncertainties).

Added persistent list of all hypocenters located (hypolist\_persistent.csv). Hypocenters are written once to this list when beginning of report interval slides past hypocenter origin-time (hypocenter origin-time first falls before beginning of report interval). NOTE: this file will grow without limit! Should be archived and deleted each month(?), year(?).

Event persistence: Added check of number of possible new defining picks,  $N_p$ , (picks with non-zero association weight) for existing events. If  $N_p < N_a/20$ , where  $N_a$  is the number of existing defining picks, then preserve previous location/hypocenter results for event and only check for association of new picks.

Added support for automatic retrieval of channel gain from FDSN station web-services: FDSNStationXML InstrumentSensitivity Value

Added Mwp distance correction: calculate\_Mwp\_correction\_INGV\_EE() from Bernardi et al. (2015) Appraising the Early-est earthquake monitoring system for tsunami alerting at the Italian candidate Tsunami Service Provider.

Regression:  $f(d) = -1.32e-6*d^3 + 2.40e-4*d^2 - 0.0146*d + 0.314$

Changed minimum Mwp distance from 5deg to 1deg since Bernardi et al. (2015) correction based on readings down to about 0deg. Provides more and earlier Mwp readings.

Added mb magnitude correction: calculate\_corrected\_mB\_INGV\_EE() from Bernardi et al. (2015) Appraising the Early-est earthquake monitoring system for tsunami alerting at the Italian candidate Tsunami Service Provider.

Regression:

```
if (mb_uncorr >= 5.125)
    mb_neic = 0.52 * mb_uncorr + 2.46;
```

Standardized and modernized writing of QuakeML XML in C and python:

monitor.xml:

1. Added xmlns:ee="http://net.alomax/earlyest/xmlns/ee" for Early-est specific tags.

2. Added publicID attribute to: <arrival>, <pick>
3. Moved event mb, Mwp, Mwpc <magnitude> from children of <origin> to children of <event>
4. Moved event Td, T50Ex, TdT50Ex, To from <magnitude> children of <origin> to new tag <ee:discriminant> children of <event>
5. Changed <confidenceEllipsoidNLL> to <ee:confidenceEllipsoidNLL>
6. Added <ee:amplitudeAttenuationPower>, <ee:amplitudeAttenuationIntercept> to <origin><quality>
7. Moved unassociated picks to within new <qrt:eventParameters> element under root element <q:quakeml>, to conform to quakeml-rt specifications.
8. <pick>: changed fields to: <ee:stationQualityWeight>, <ee:T50>, <ee:Aref>, <ee:T50Ex>, <ee:Td>, <ee:snHF>, <ee:snBRB>; added <ee:T0>
9. <arrival>: added <timeWeight>, changed fields to: <ee:totalWeight>, <ee:distanceWeight>; removed <T0> (moved to <pick>)
10. TimeQuantity: added <uncertainty> and <confidenceLevel>

hypo.<event\_id>.mech.hash|fmamp.xml:

1. Moved <FocalMechanism> element to within <q:quakeml><EventParameters ><Event > elements.
2. Added xmlns:ee="http://net.alomax/earlyest/xmlns/ee" xmlns:hash="http://net.alomax/earlyest/xmlns/hash" for Early-est and method-specific FocalMechanism elements.
3. Changed tags to Early-est and method-specific FocalMechanism elements.
4. Modified publicID's to be a complete smi's.

Format changes:

1. picks.csv, pickmessage.html: added Aerr – ratio of observed Aref amplitude to predicted Aref amplitude when P amplitude attenuation check enabled.
2. hypo\*.csv, added:
  - n\_sta\_tot – number of stations for which data has been received in past.
  - n\_sta\_active – number of stations for which data has been received and data\_latency < report\_interval (e.g. 1min).
  - assoc\_latency – latency in seconds of first association after origin-time of event. May be erroneous (too large) if event de-associated and later re-associated.