Automatic Earthquake Detection System at AEIC: Setting and Performance

by Natalia Ruppert, Seismologist

November, 2006

The mission of AEIC is to rapidly determine location and size of earthquakes in the State of Alaska and to disseminate this information to the State and Federal agencies, scientists, and the general public. The AEIC real-time and post-processing systems are based on the Antelope software package from BRTT, Inc. AEIC is using Datascope relational database platform for the data collection and archival. For earthquake parametric information and waveforms, CSS3.0 schema is used. Waveforms are stored in SEED format.

The purpose of this report is to describe settings of the real-time earthquake detection system at AEIC and to provide analysis of it's current performance.

1. Real-time system settings.

1.1. Real-time arrival detections.

Automatic arrivals are detected by the *orbdetect* module with multi-frequency STA/LTA algorithm on vertical data channels. A signal-to-noise ratio (SNR) of 4 is used for "detection on" condition, and SNR=2 used for "off" condition.

As of October, 2006, there were ~370 vertical channels available for real-time arrival detections. Roughly half are the AVO data channels, and the rest are the AEIC, ATWC, IRIS (both in Alaska and Russia), and Canadian. Of those, about 170 channels were actually used for the real-time detections as of October 16, 2006 (~70 AVO and ~100 of the rest). No strong motion channels are currently being used for the real-time detections. The goal is to define more or less uniform station distribution throughout the network. All regional stations are being used, when available. Exceptions are the stations that are not working or are causing too many false alarms due to high level of telemetery glitches or other noise. Since AVO networks are very dense concentrations of stations in the regional monitoring sense, only subsets of 3-4 stations from each AVO subnet are being used for detections. Also, volcano networks have higher level of telemetery glitches due to various reasons. For example, whole Peulik network experiences some kind of simultaneous glitch as often as several times per hour. This causes too many mislocations. Therefore, currently all Peulik stations are excluded from the real-time detections. List of good detection channels is defined in *orbdetect.pf*. This list is being updated when new information about station state of health becomes available (duty person reports, Seismolab meeting, good-size tele checks). See */iwrun/op/run/pf/orbdetect.pf* file for details.

1.2. Automatic locations.

If there are at least 6 con-curing detections within a specified time window (300 sec currently), then a location that fits the best this set of detections is searched over several pre-calculated three-dimensional regional grids and one global teleseismic grid. Correct identification of teleseismic events aids in avoiding mislocating them as bogus regional events. Lowering the number of candidate detections to 5 dramatically increases number of bogus events and mislocations. Increasing it cases too many smaller real events to be missed. So, 6 is an optimal number that compromised between the number of missed smaller events and the mislocations.

The candidate pick time window is set to 300 sec which roughly corresponds to the total phase moveout time difference between closest and furthest source-station distance across the network. Large events in the Aleutians are recorded by the whole network. For those events, station distance approaches 20 degrees for the most distant sites. In case when candidate pick time window is set too small, large events get split into two events, one based on the set of picks from the close-field and the other from the far-field stations. Larger window, however, require more CPU time, since more picks will be included into internal candidate pick list. Potentially, aftershock sequences could cause delay in the pick list processing.

For larger events with greater number arrivals, multiple subsequent locations may be determined as more detections become available. How often secondary locations are determined is control by the time passed from the last detection and/or number of additional picks available. For settings of these and other variables see *orbassoc.pf* parameter file and *orbassoc* man page.

1.3 Regional grids.

The regional grid geometry matches natural distribution of seismicity in the state. There are 7 regional grids:

- *central_aleut*
- *eastern_aleut_shal* and *eastern_aleut_deep*
- *northern_ak*
- *scak_shal* and *scak_deep*
- *southeast_*ak

Having multiple regional grids reduces number of bogus and mislocated events. Each grid is using only subset of stations that are located within the grid and in the neighboring region. A single regional grid for Alaska/Aleutian region would have to include the whole depth range of 0-200 km and all available stations. This would result in many mislocations and bogus events.

See Appendix A, Figures A1 through A6, for details of each grid layout. New travel time grid files (*/iwrun/op/run/ttgrid_reg* and *ttgrid_utele)* need to be calculated occasionally, such as when new stations come online, or the sites close off. See parameter file */iwrun/op/run/pf/ttgrid_reg.pf* and */iwrun/op/run/pf/ttgrid_utele.pf* for details of the grid settings. See man pages for *ttgrid* and *ttgrid_show* for even more details.

2. Performance analysis.

2.1. Analysis of automatic location post times.

The ANSS standards for hypocenter post time are 2 min for hi-risk urban areas and 4 min for mod-high hazard areas, respectively. I have done analysis for the month of August of 2006 for operational (earlybird) and migration (energy) systems (Figures 1 and 2). In Figure 1, horizontal axis represents an event ID with events sorted by the origin time and vertical axis is the difference between load time and origin time for each automatic origin in the database. Multiple origins for the same *evid* are represented by crosses along a single vertical line with red circle marking the earliest location. You can see a large delay (up to 100 min on *op* and up to 70 min on *mig*) caused by problems with the communications or something (I don't remember what).

Mean time delays are 5.6 min and 4.2 min for *op* and *mig* systems, respectively. Taking into account only delays less than 30 min, mean values are 4.3 and 3.9 min, respectively. On energy, 28% of events are detected within 2 min or sooner and 61% are within 4 min or less. On earlybird, 20% of events are detected within 2 min or less and 57% within 4 min or less. This quantitative analysis supports my earlier suspicion that earlybird performs worse than ice or energy. See the table below for more detailed statistics.

Table 1: Time delays of event detections.

	op	mig
mean delay of the earliest locations for all events	5.6 min	4.2 min
mean delay of the earliest locations for 30 min or less delays	4.3 min	3.9 min
first detections within 2 min or less	20%	28%
first detections within 4 min or less	57%	61%
75% of first detections within	5.6 min	5.1 min
90% of first detections within	7.8 min	6.7 min

Note, that how soon location is determined depends not only on how soon and how many picks become available for an association, but also how many previous internal pick candidate lists are being processed at any given time. The time delays can be reduced by reducing number of stations within each grid, and/or by reducing size of the grids, i.e. horizontal grid spacing. Depth intervals are already spaced out. One of the other major reasons for long delays in automatic locations will be in the case of large aftershock sequences. I've seen delays up to 3 hours on *op* system in June, 2006 when the M6.5 Rat Island earthquake struck.

Once the event is located, its magnitude (Ml) is calculated within less than 1 minute.

2.2. Automatic versus reviewed magnitudes.

The following analysis includes data for January-August of 2006 from operational system. Automatic event detections have been matched with the reviewed events. When available, *mb*'s were added from the PDE catalog. See Figures 3, 4, 5 and Table 2 below.

Table 2: Automatic vs reviewed magnitudes.

	All	Aleutians	Alaska
mean(Mauto-Ml)	0.10	0.17	0.07
Mauto underestimates Ml	26%	19%	30%
Mauto overestimates Ml	57%	67%	52%
Mauto=Ml	17%	14%	18%
Mauto within 0.2 of Ml	59%	46%	64%
Mauto within 0.5 of Ml	92%	91%	93%
Mauto differs by 1 or more from Ml	0.02%	0.01%	0.02%
mean(Mauto-mb)		-0.07	0.07
Mauto within 0.2 of mb	31%	32%	28%
Mauto within 0.5 of mb	83%	84%	78%

Overall, statistics is rather favorable to us for real events where we are within 0.5 of the reviewed magnitude for 92% of the earthquake detections.

The incidents with overestimated magnitude for M6+ Aleutian events in June and July were traced to a typo in *orbmag.pf* file for station SMY which caused erroneous magnitude values for this station. Because of the remoteness and large size of these events, SMY was the only station available for magnitude calculation, for the rest of the broadbands were too far and all short-periods were clipped. This was an unfortunate occurrence.

Next, there were 11 events with automatic magnitudes in M4-5 range that ended up in M1-3 range after review (Figure 3, upper left plot). All events except for 1 had fewer than 10 associated arrivals and therefore these events were filtered out from our QDDS submissions and from *dbrecenteqs* page. In other words, no damage has been done by these worst offenders.

There is not much to say about *mb*'s (Figures 4 and 5). There were 224 events with the *mb* values in the Aleutians and 65 events outside the Aleutians. Both Mauto and Ml are not too far off from the corresponding *mb*'s. For Aleutian events linear regression between available *mb* and Ml values yields this formula:

$$mb=0.8082*Ml+1.0659$$

Outside of the Aleutians, the formula is:

$$mb=0.6446*Ml+1.5839$$

Overall, *mb*'s for Aleutians are consistently higher than Ml's. And since automatic Ml's are often a bit higher than the reviewed Ml's, they are a closer match to *mb*'s.

There are, however, 8% of mislocated events, which were not included into the above analysis. In this analysis "mislocated" means no corresponding reviewed location or reviewed location differs by more than 200 km from the automatic one. This will be discussed later.

2.3. Location errors.

The following analysis includes data for January-August of 2006 from operational system, the same as for the above magnitude analysis. Automatic event detections have been matched with the reviewed events. See Figure 6 and Table 3 below.

The ANSS performance standards set epicenter uncertainty of 2 km for the high-risk urban areas and 5 km for mod-high hazard areas (20 km for Alaska). Depth uncertainties are 4 km for high and 10 km for mod-high hazard areas (20 km for Alaska). Since our automatic locations are produced by a grid search, there are no error estimates currently available for our automatic hypocenters. The following compares automatic epicenters and depth estimates with the corresponding reviewed values.

Table 3: Automatic vs reviewed hypocenters.

	All	Aleutians	Alaska
automatic epicenters within 10 km of reviewed	59%	29%	74%
automatic epicenter within 20 km of reviewed	77%	55%	89%
75% of automatic and reviewed epicenters are within	18 km	36 km	10-11 km
90% of automatic and reviewed epicenters are within	43 km	71 km	21 km
automatic and reviewed depths are within 5 km	71%	65%	74%
automatic and reviewed depths within 10 km	81%	75%	84%
75% of automatic and reviewed depths are within	6-7 km	10 km	5-6 km
90% of automatic and reviewed depths are within	17-18 km	24-25 km	14-15 km

Not bad overall. There is some room for improvement. We can make grids denser, both in horizontal and depth. This move, however, will increase the grid size and lead to longer time delays in hypocenter associations.

2.4. Mislocated events.

In this analysis "mislocated" means no corresponding reviewed location or a reviewed location differs by more than 200 km from the automatic one. There were 7.6% of mislocated events for the time period between January and August of 2006, a total of 452 events, or roughly two mislocated events per day. Of those, 158 events (or 35% of all mislocated and 2.6% of all detections) had no magnitudes. This leaves 4.9% of truly misidentified events. Most of these events (83%) had fewer than 10 associated arrivals. Events without magnitudes do not get submitted to QDDS system, as well as events with automatic magnitudes of 4 and greater and fewer than 10 arrivals. Automatic events with magnitudes 3.5 or greater and number of arrivals less than 10 do not get posted on *dbrecenteqs* page. So, if we count mislocated events with magnitudes greater than 4 that were submitted to QDDS, there were 10 events total that could have caused false alarms down the QDDS system. This is 0.002% of total event detections by AEIC in first 8 months of 2006. Most of mislocated events are deleted or relocated by the seismologist-on-duty within 24 hours of the occurrence. See Figure 7 for details.

3. Conclusions.

In terms of the post time of automatic locations, energy and ice outperforms earlybird (Figures 1 and 2). On energy, 28% of events are located within 2 min of the occurrence, and 61% are within 4 min. Longer delays may be caused by vigorous aftershock sequences.

For the first 8 months of 2006, 92% of all automatic magnitudes were within 0.5 of the reviewed values (Figure3). Event filters that are based on magnitude and number of associated phases helped to filter out bogus events with larger magnitudes (M>3.5-4.0). *mb*'s are in good agreement with automatic and reviewed magnitudes.

There were 7.6% of mislocated and/or bogus events. Mislocations are caused by con-curing small events without enough arrivals being combined by associator into bogus events or by noise and glitches in the waveform data.

The greatest challenge in ensuring reliability of automatic locations is the size of the region and abundance of various tectonic regimes, such as active crustal activity across the whole region, subduction zone seismicity down to 250 km depth in southern Alaska and Aleutian arc, volcano-tectonic events, glacial quakes and quarry blasts. AEIC constantly monitors and evaluates performance of the real-time earthquake detection system and makes necessary adjustments and improvements, such as removing noisy stations form autodetection list, tuning autodetection parameters, improving regional grids. This ongoing effort will continue in the future. Additional challenge is older stations with analog telemetery that are prone to spurious spikes and high levels of noise. AEIC engineers and field technicians continually work on improving quality of recorded seismic signals. Updating some of these stations with digital telemetery systems and broad-band sensors would also help. In my opinion, however, there will be always a certain percentage of mislocated events. Our goal would be to keep this number low.

4. Figures

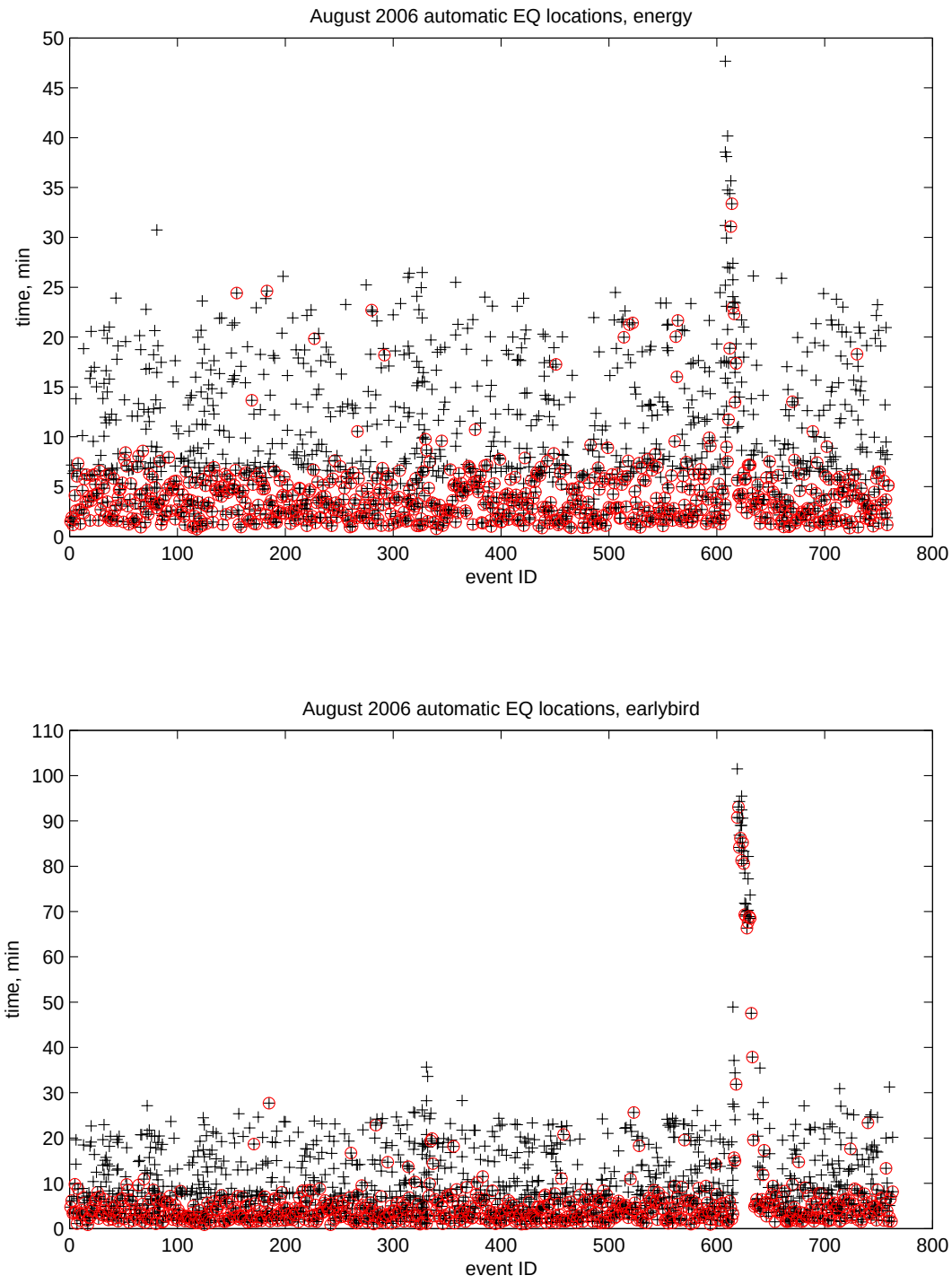

Figure 1. Time-plot for August, 2006 automatic locations from operational system earlybird (bottom) and migration system energy (top). Events along X-axis are sorted by time and all event ID's are renumbered sequentially from 1. Time on Y-axis is the difference between load time and origin time in minutes. Red circles denote the first detection for each event. Black crosses are the secondary detections. See Table 1 for statistics.

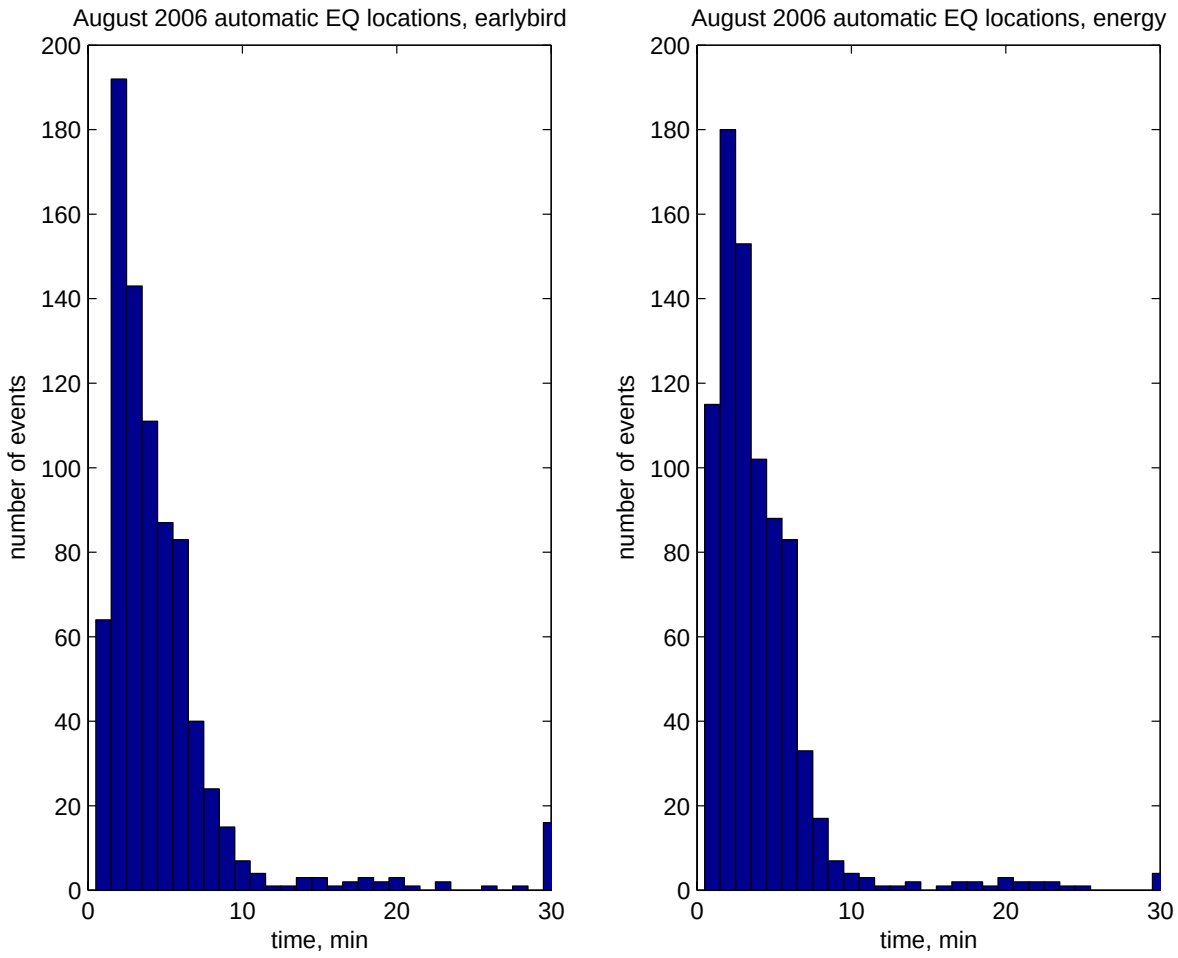

Figure 2. Time histogram for August, 2006 automatic locations from operational system early-bird (left) and migration system energy (right). Time along X-axis is the difference between load time and origin time in minutes. Events are grouped into 1 min intervals. Mean time delay is 5.6 min and 4.2 min for op and mig systems, respectively. Considering only delays that are less that 30 min, mean values are 4.3 and 3.9 min, respectively. See Table 1 for more statistics.

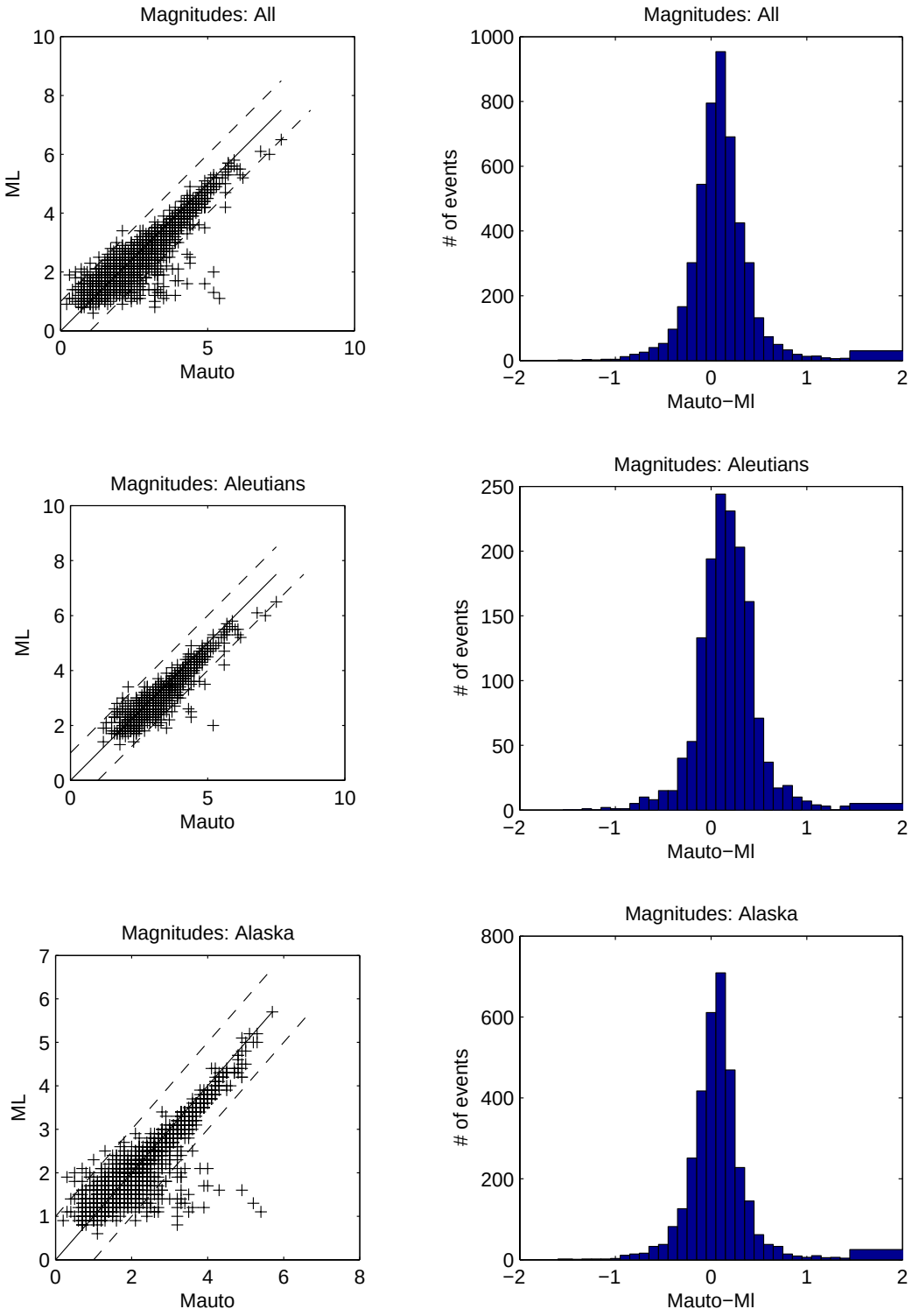

Figure 3. Left column: Automatic versus reviewed magnitude Ml plots for January-August 2006 data: top - all events, middle - Aleutian events, and bottom - mainland Alaska events. Solid line is Mauto=Ml, dashed lines mark a 1 unit difference between Mauto and Ml. Right column - same data presented in histogram form; each bin is 0.1 step. See Table 2 for more statistics.

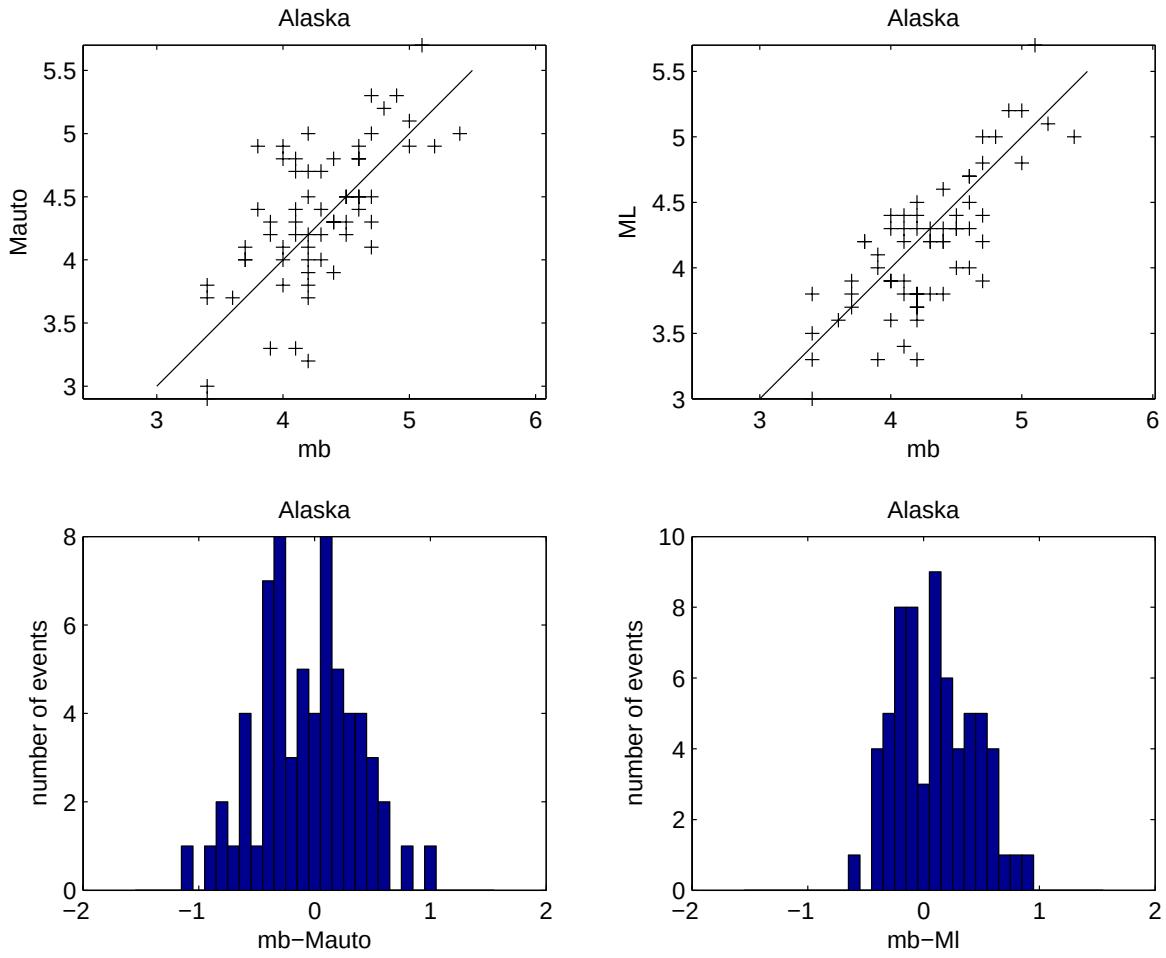

Figure 4. Comparison plots for mb versus Mauto and Ml for events outside of Aleutians. Left column: mb versus automatic magnitude. Right column: mb versus reviewed Ml. See Table 2 for more statistics.

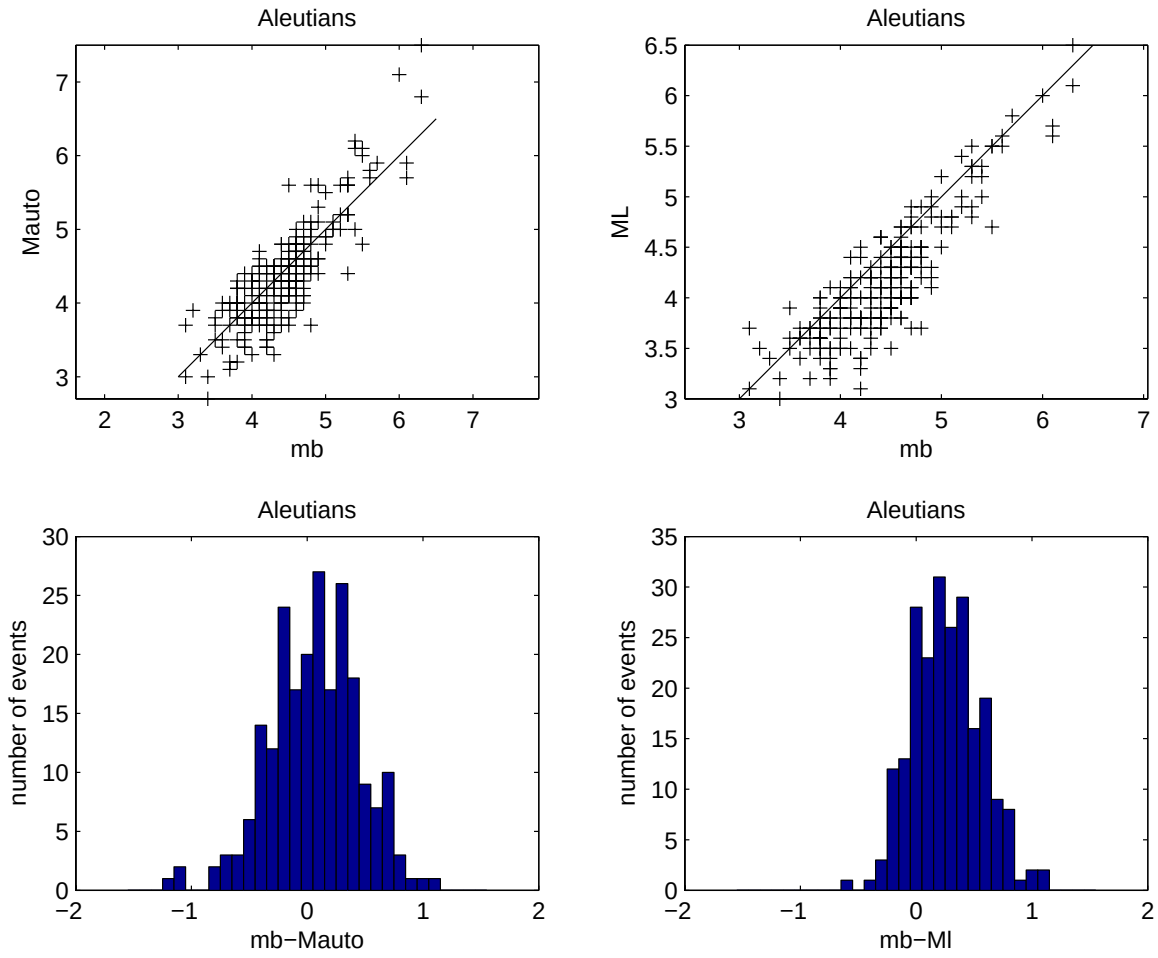

Figure 5. Comparing mb with Mauto and Ml for the Aleutian events. Left column: mb versus automatic magnitude. Right column: mb versus reviewed Ml. See Table 2 for more statistics.

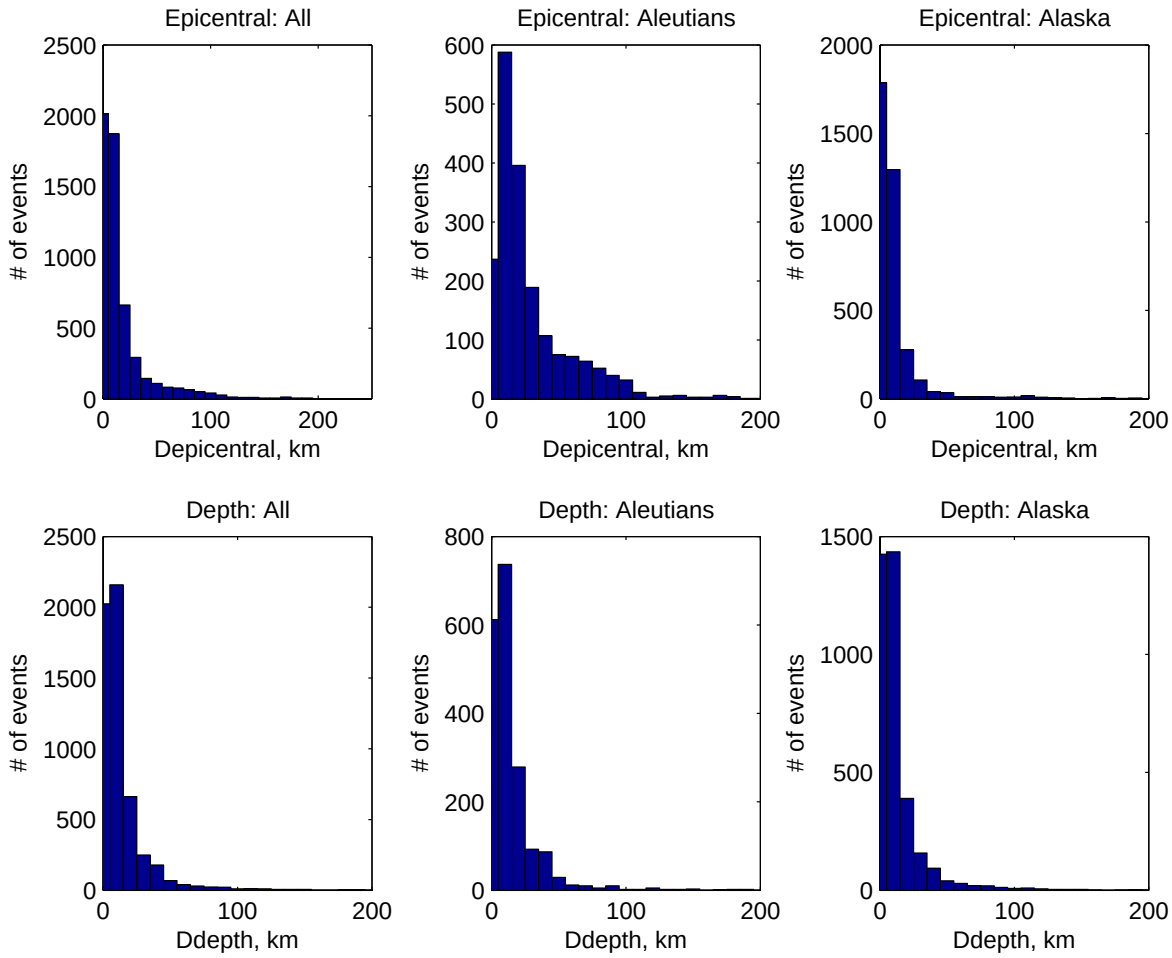

Figure 6. Comparing automatic and reviewed locations. Each bin is 10 km wide. Upper row: epicenter difference; bottom row - depth difference. Left column: all events, middle column: Aleutian events; right column: events outside of Aleutians. See Table 3 for more statistics.

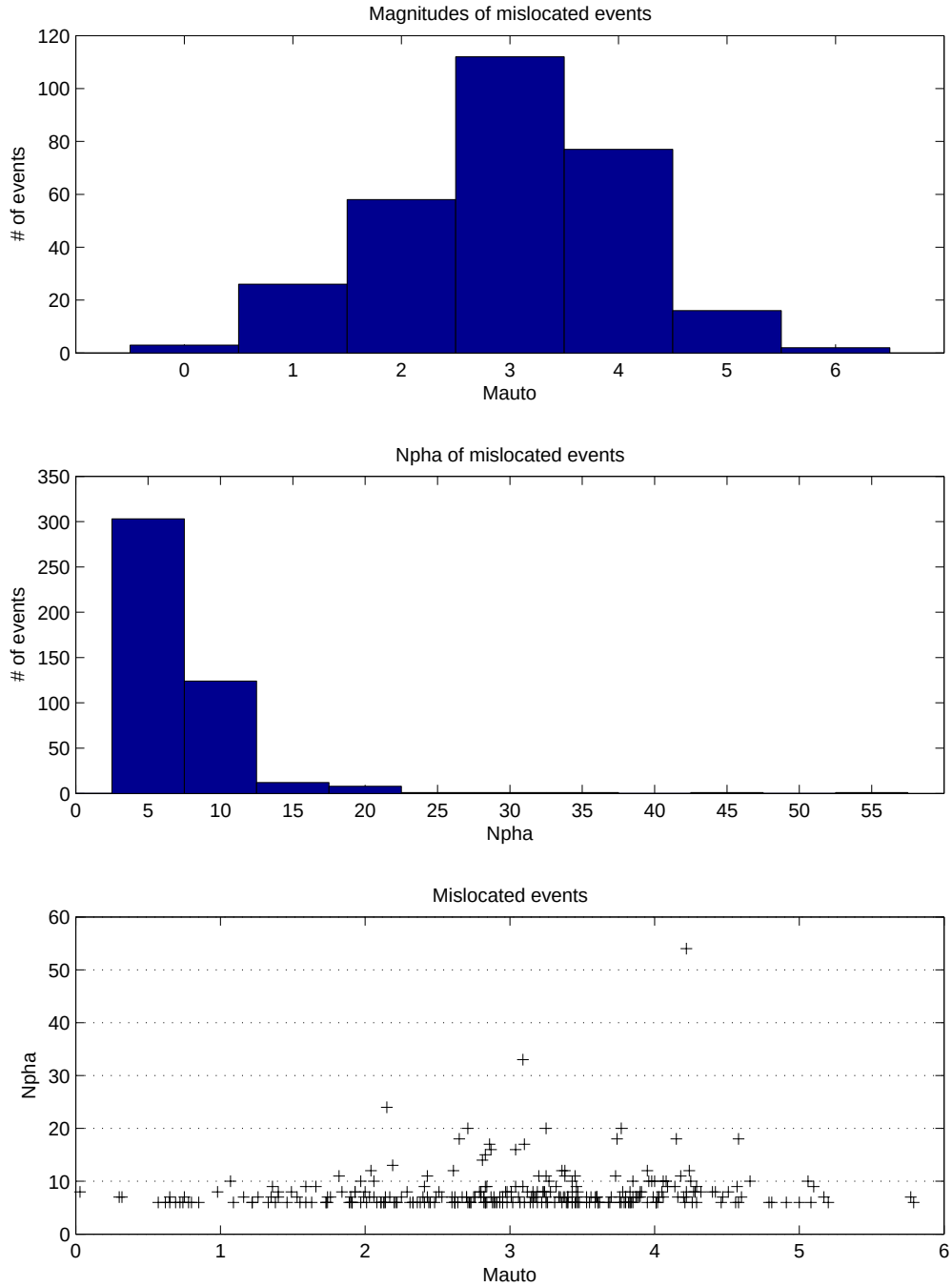

Figure 7. Analysis of mislocated events (automatic epicenter differs from the reviewed by more than 200 km). There were 7.6% of mislocated events in the first 8 month of 2006. Of those, 35% had no magnitudes. Upper plot is a magnitude histogram of mislocated events (bin centered on 0 includes events with magnitudes less than 1, bin centered on 1 includes events between 1 and 2, etc.). Middle plot is the histogram of associated arrivals for all mislocated events, with or without magnitudes (bin centered on 5 includes events with number of phases between 5 and 10, etc.). Bottom plot shows magnitude versus number of associated arrivals. Events with M>=4 and 10 or fewer associations do not get submitted into QDDS system. *dbrecenteqs* map does not include events with M>=3.5 and 10 arrivals or less.

Appendix A. Regional grids.

Figure A1. Central Aleutians grid (1-200 km depth).

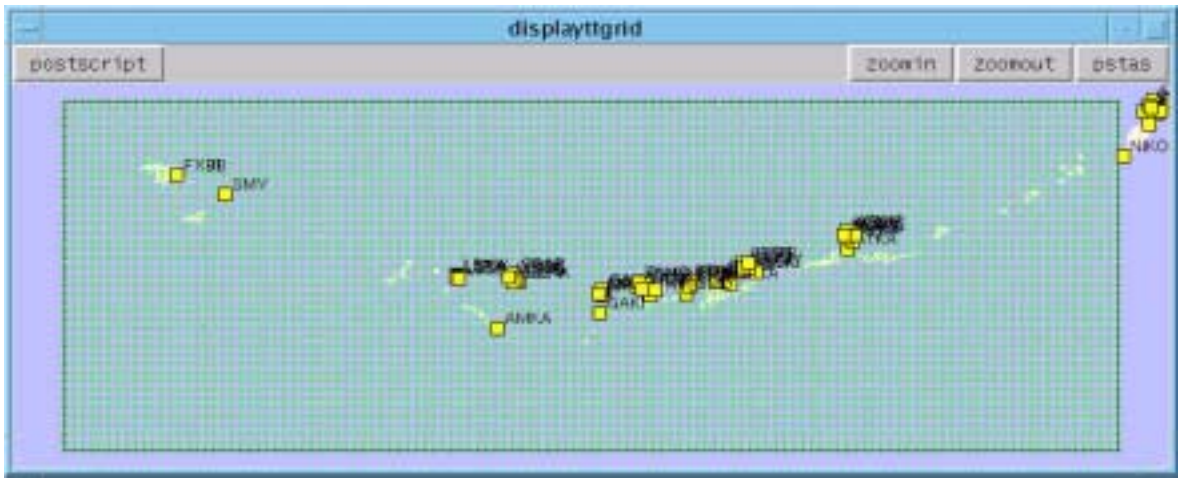

Figure A2. Eastern Aleutian shallow grid (1-50 km).

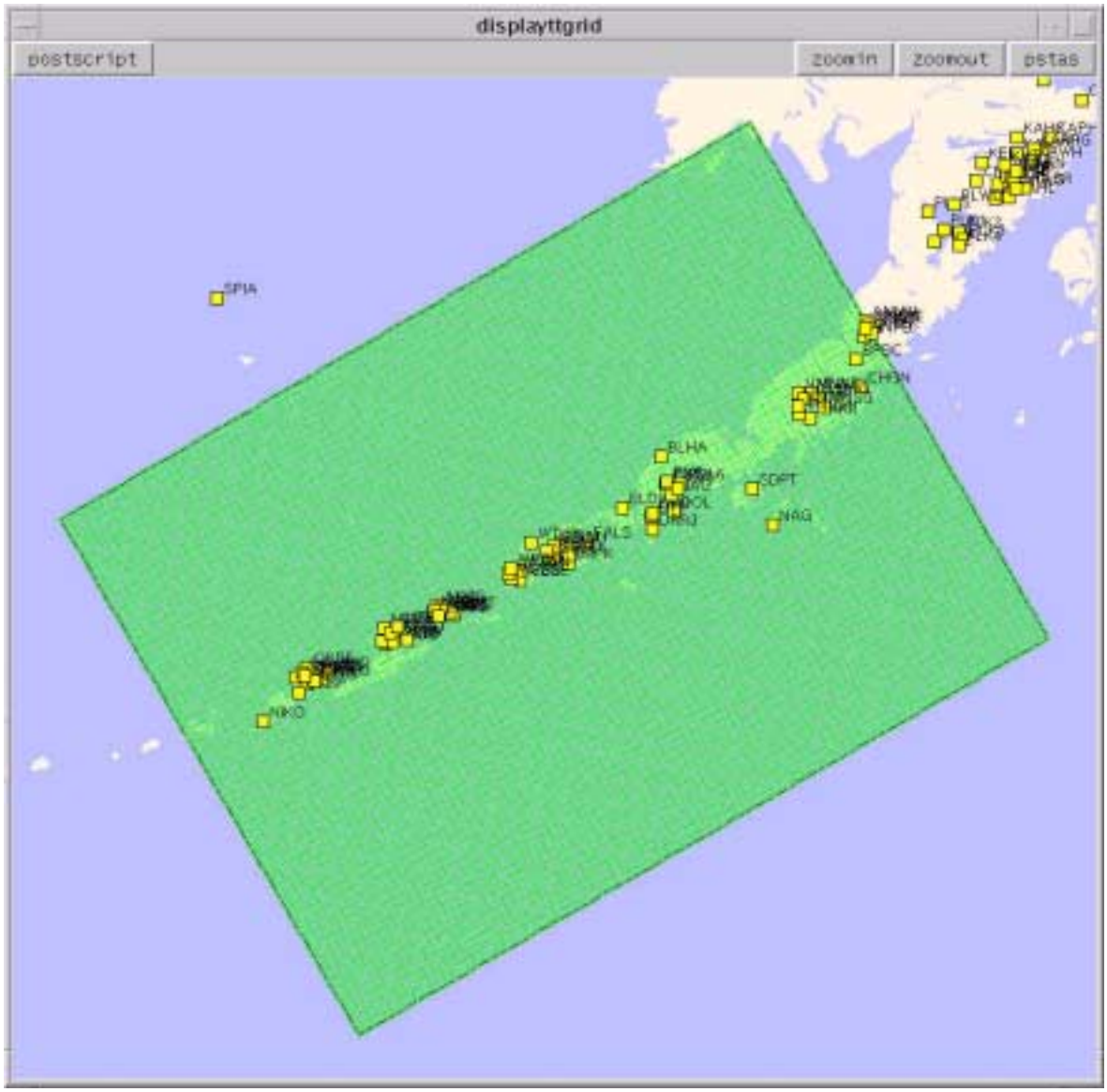

Figure A2. Eastern Aleutian shallow grid (1-50 km). Continued.

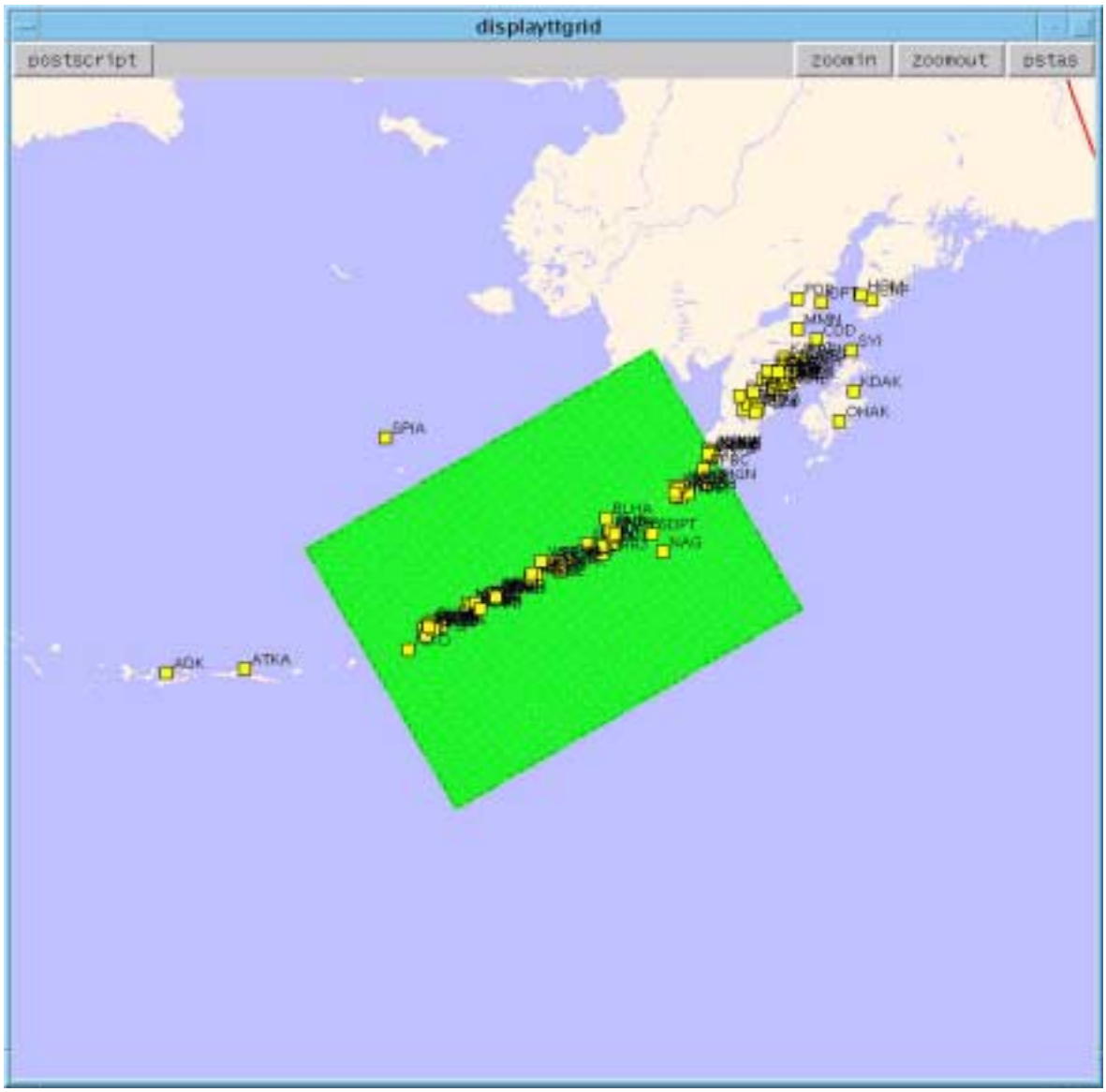

Figure A3. Eastern Aleutian deep grid (60-200 km).

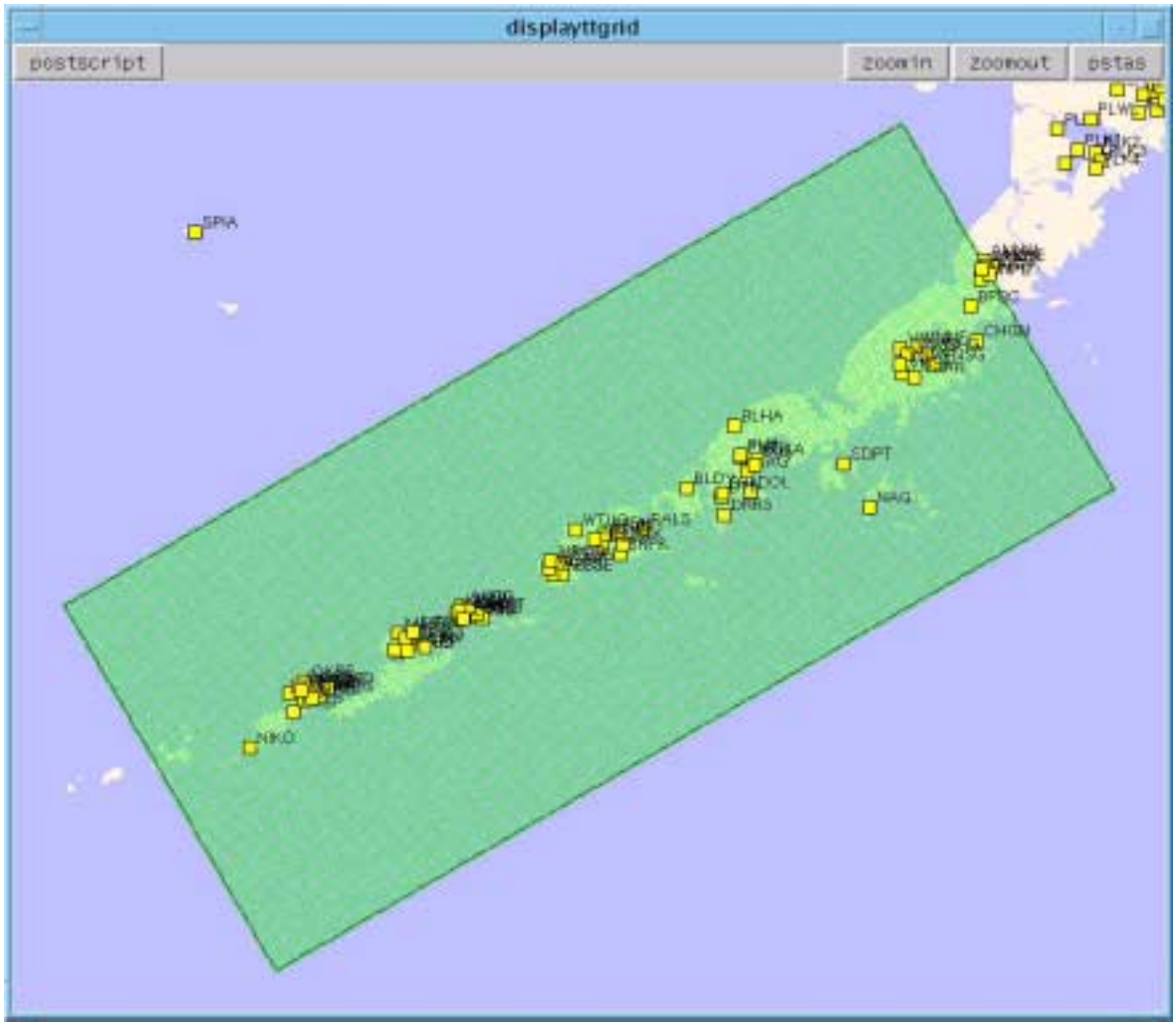

Figure A3. Eastern Aleutian deep grid (60-200 km). Continued.

Figure A4. South-central shallow grid (1-50 km).

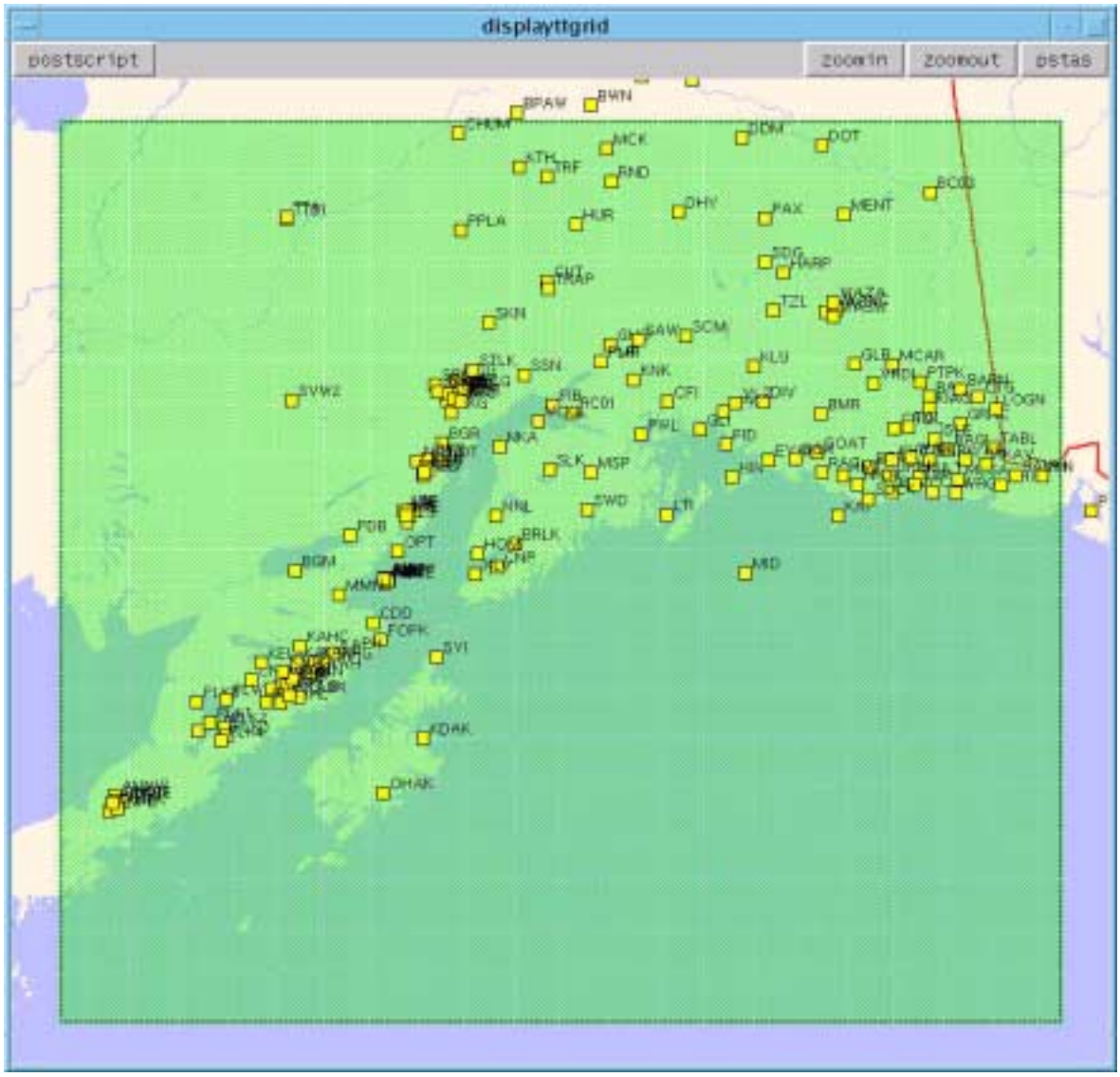

Figure A4. South-central shallow grid (1-50 km). Continued.

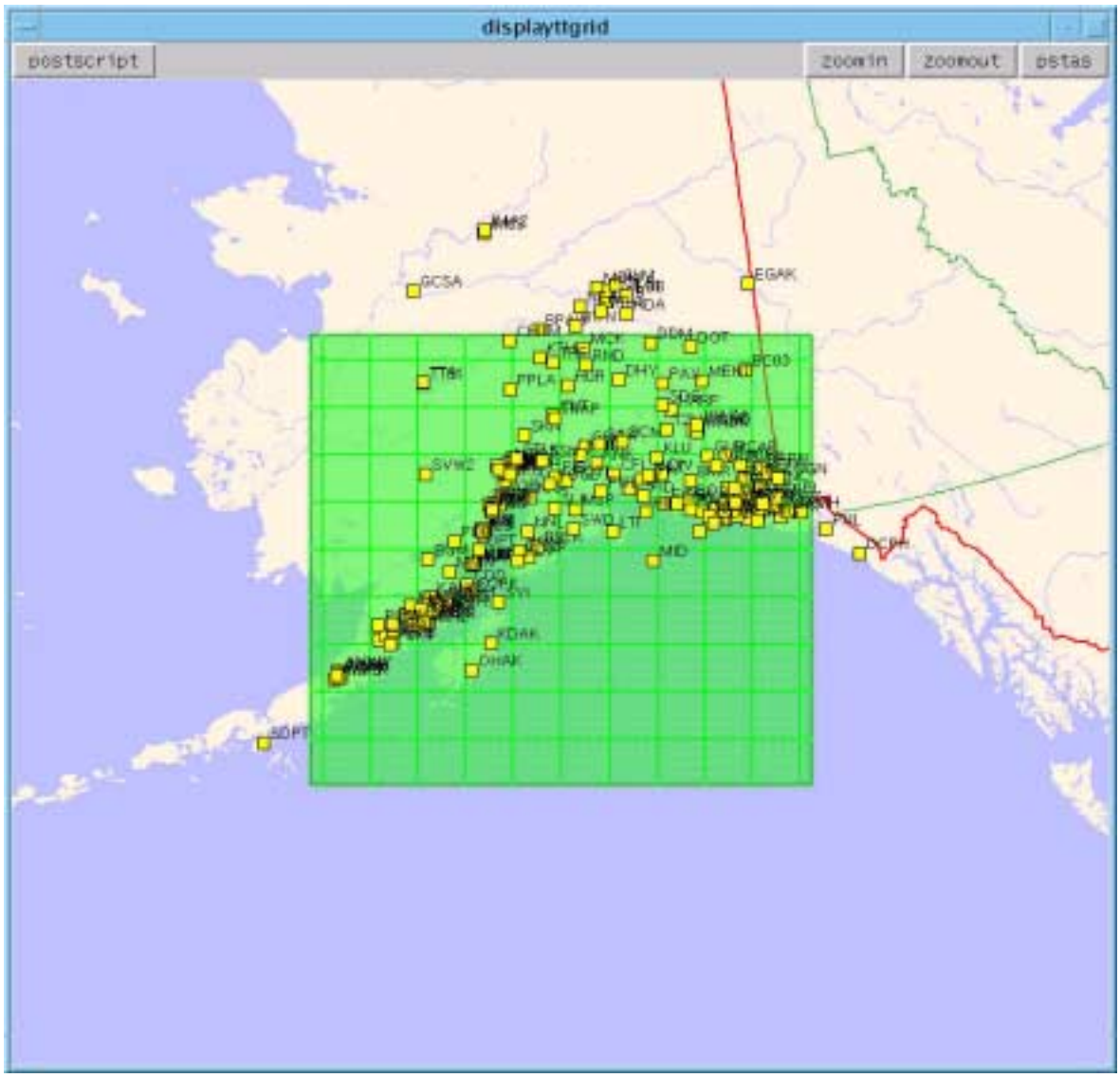

Figure A5. South-central deep grid (60-200 km).

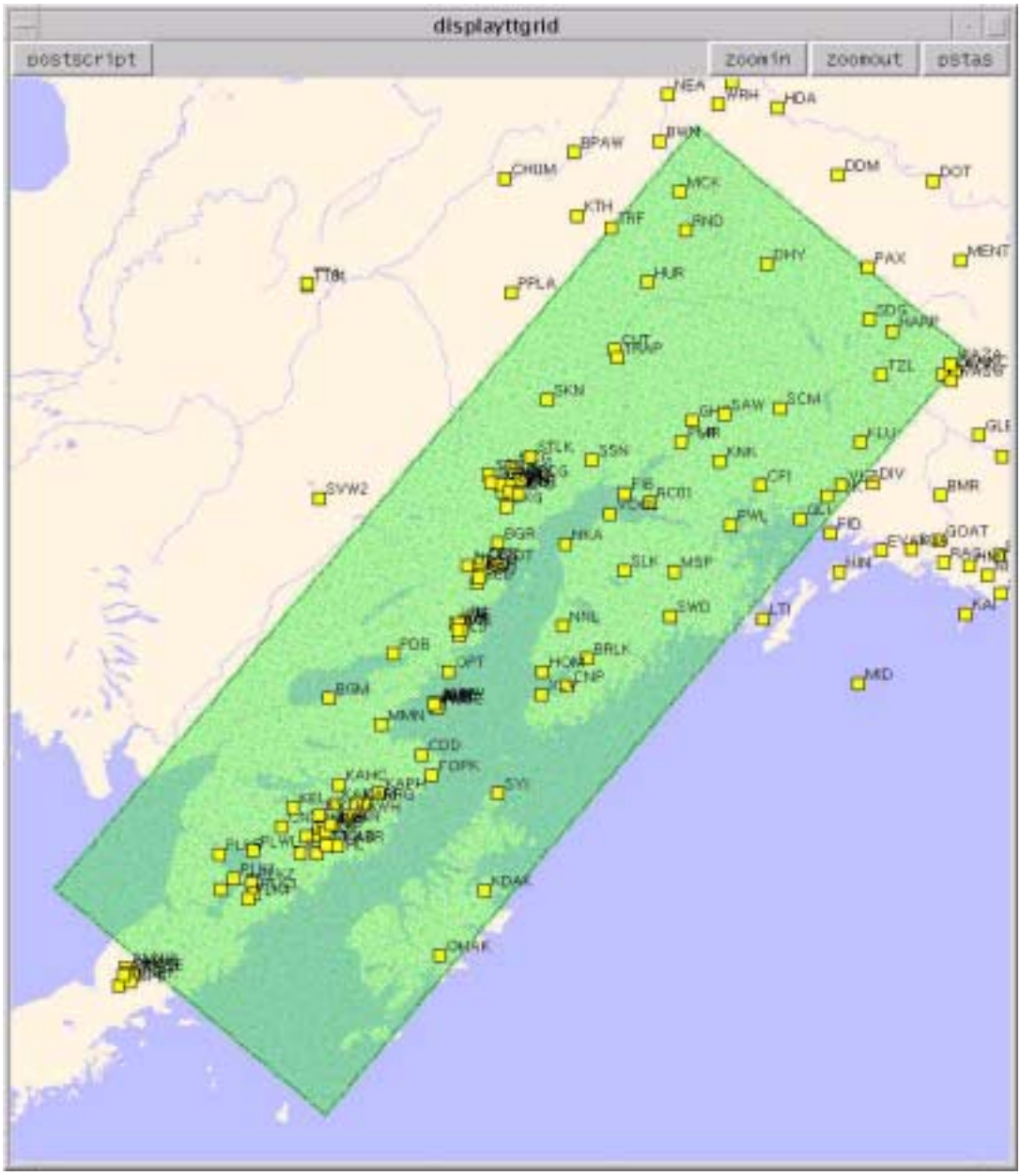

Figure A5. South-central deep grid (60-200 km).Continued.

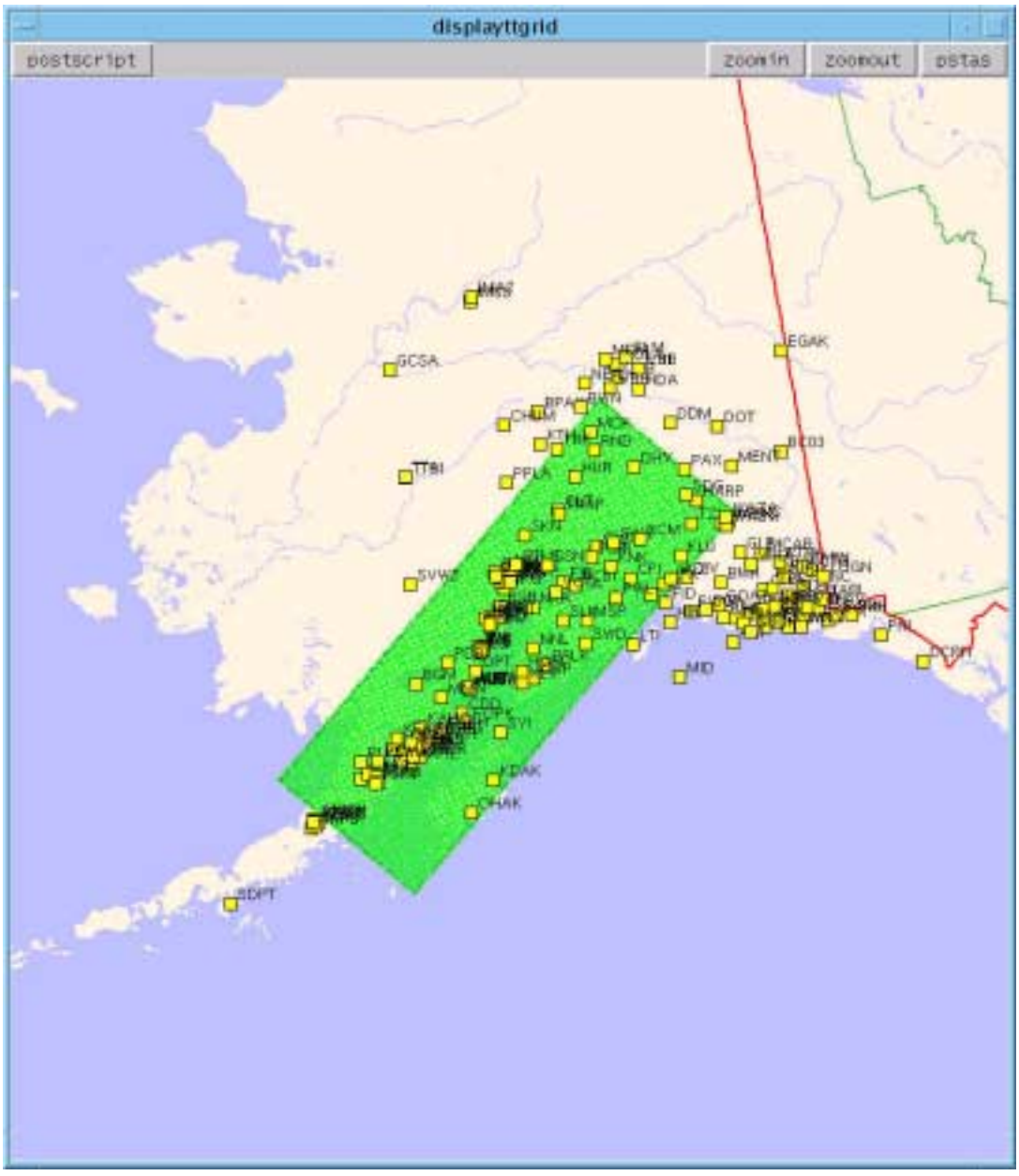

Figure A5. Southeast Alaska grid (1-40 km).

Figure A6. Northern alaska (40-200 km).