

A fast algorithm for automatic phase picker and event location: Application to the 2018 Hualien earthquake sequences

Yu-Hsuan Chang^{1,2}, Shu-Huei Hung^{2,*}, and Yen-Ling Chen¹

¹Central Weather Bureau, Taipei City, Taiwan

²Department of Geosciences, National Taiwan University, Taipei City, Taiwan

Article history:

Received 31 August 2018

Revised 28 November 2018

Accepted 23 December 2018

Keywords:

Automatic phase picker, Associator, AIC, 2018 Hualien earthquake, Mainshock, Foreshock, Aftershock

Citation:

Chang, Y.-H., S.-H. Hung, and Y.-L. Chen, 2019: A fast algorithm for automatic phase picker and event location: Application to the 2018 Hualien earthquake sequences. *Terr. Atmos. Ocean. Sci.*, 30, 435-448, doi: 10.3319/TAO.2018.12.23.01

ABSTRACT

With the vast amounts of data recorded by the CWB seismic networks, a fast and precise algorithm for near real-time earthquake locations as well as fundamentally urgent source and hazard studies of large damaging earthquakes in Taiwan becomes increasingly important. Here we present a novel method for automatic phase identification and event detection, modified from a Python-based PhasePAPy package to be more computationally efficient and suitable for the high-rate seismicity and large-scale dense network in Taiwan. The performance efficiency is enhanced by substituting the Python code with a Fortran subroutine to calculate the characteristic function for phase picker. It is further improved by amalgamating the picks close enough in time but on different components, associating pairs of the amalgamated picks with the origin times of candidate events through an empirical linear relation between the P and S-P times reported in the CWB phase catalog, and isolating clustered candidate events and corresponding phase picks to determine the locations of real earthquakes. The algorithm is applied to detecting in almost real time over two thousands of events that occurred within four days accompanying the 4 February 2018 M_L 5.89 foreshock and 6 February 2018 M_L 6.25 Hualien mainshock, which far surpasses the capability of experienced human analysts. The temporal and spatial distributions of the detected aftershocks provide the timely and first-hand information to estimate the aftershock decay rate and verify the orientation of the fault plane and rupture extent of the foreshock and disastrous mainshock as a guide for fast risk assessment and source characteristics analysis.

1. INTRODUCTION

The currently active orogeny in Taiwan as a result from the arc-continent collision is accompanied by two nearly-orthogonal subduction zones to the northeast and south of Taiwan (Suppe 1984; Teng 1990). Such a complex tectonic setting makes the island suffer from frequent earthquake activities and potential risks of great subduction-zone megathrust events (Hsu et al. 2012; Chen et al. 2016). Every year, on average three earthquakes of magnitude 6 or larger hit Taiwan and near offshore region and posed imminent threats to public safety and property (Jian et al. 2017, 2018a; Kanamori et al. 2017). Real-time monitoring of seismicity and fast determination of accurate source parameters of strong earthquakes and their aftershock distributions are of

urgent importance for both scientific and societal concerns.

Since the early 1990s, the Central Weather Bureau (CWB) of Taiwan has operated dense seismic networks including short-period instruments, and later-added strong-motion and broadband stations, to monitor restless earthquakes in and around Taiwan (Shin et al. 2013). One of the major and challenging tasks is to provide timely and accurate earthquake source information to help the community respond appropriately to possible seismic hazards. Currently an Earthquake Early Warning (EEW) system has been operated at CWB for quick detection and characterization of large earthquake within a few to tens of seconds after they occur and issuance of alerts to people and infrastructures before strong shaking arrives (Chen et al. 2019a; Wu et al. 2018). Different from the EEW system only concerning the earthquakes that meet a specific trigger criterion, the CWB

* Corresponding author
E-mail: shung@ntu.edu.tw

Seismological Center also processes continuous digital records in real time transmitted directly from every field station and manually cuts the time-series segments that contain possible event-generated signals used for the following phase picking and event location. The main aim of this task is to both monitor minor earthquakes and measure major earthquakes accurately and build up a complete event catalog that serves as a key foundation for the studies of seismotectonics, historical seismicity, earthquake statistics and source physics, and seismic risk mitigation plans in Taiwan (Chen et al. 2016, 2019b; Jian et al. 2017; Wen et al. 2019).

Since 2012, the CWB integrated the existing short-period, broadband, and strong-motion stations distributed on more than 200 sites across Taiwan and at some offshore islands which now attain over 400 instruments with 3-component seismic records used in regional earthquake monitoring and other seismic studies (Fig. 1; Shin et al. 2013). Because of the rapidly expanding datasets, in recent years the CWB has started to run a prototype automatic procedure in parallel with the manual detection and location of seismic events. However, the official report of the earthquake hypocenters and magnitudes still relies heavily on subjective manual judgement. For large events such as the 2018 Hualien earthquake sequence starting with a M_L 5.89 (M_w 5.9) foreshock two days before the M_L 6.25 (M_w 6.4) mainshock that struck just offshore eastern Taiwan on 6 February 2018 and severely damaged the nearby coastal city of Hualien (Jian et al. 2018b; Huang et al. 2019), there were estimated to be over 20000 events occurring afterwards in one month above the detection threshold of the CWB network, almost a half the average annual number of earthquakes in the Taiwan region. The current processing procedures in CWB become intolerantly inefficient and impractical to identify and locate a huge number of aftershocks clustering in time and space after such destructive earthquakes. Therefore, an advanced and efficient automatic algorithm for such type of the works is very much needed.

In general, there are two important parts that make up the algorithm: an automatic phase picker and one kind of the associators capable of effectively separating the phases arriving almost at the same time from different earthquakes. The first part of phase picking is essential to a variety of seismological studies on earthquake location, focal mechanism determination, travel-time based seismic tomography, and etc. Manual picking subject to the experience of analysts can introduce inconsistencies in the picked arrival-time data, particularly those used in the routine catalog locations (Leonard 2000). With the improved digital seismic data and computing facilities, many automatic phase pickers based on different mathematical operations have been developed. These algorithms are mainly classified into, but not limited to, energy transient methods in time or frequency domain (e.g., Withers et al. 1998; Vassallo et al. 2012), autoregressive (e.g., Leonard and Kennett 1999), neural network (e.g.,

Gentili and Michelini 2006), and wavelet transform (e.g., Zhang et al. 2003; Bogiatzis and Ishii 2015) methods, and each of them has its own advantages and limitations.

The second is to assign the phases of selection to associated earthquakes. Usually, it is very difficult to identify multiple earthquakes occurring close in time less than one minute, because the S arrivals of the preceding event are likely to be misjudged as the P of the later event (JMA 2017). It is thus necessary to have a reliable method to automatically allocate the picked signals to correct seismic phases associated with individual earthquakes.

In this study, we present an efficient and stable algorithm of automatic phase picking and event identification suitable for the current seismic data acquisition and processing systems at CWB. Our algorithm is adapted from a Pure Python code package, called PhasePapy (Chen and Holland 2016), originally developed and used in the Oklahoma Geological Survey for near real-time earthquake monitoring. The package not only takes advantage of very useful open-source Python libraries, but also adopts Obspy, a Python

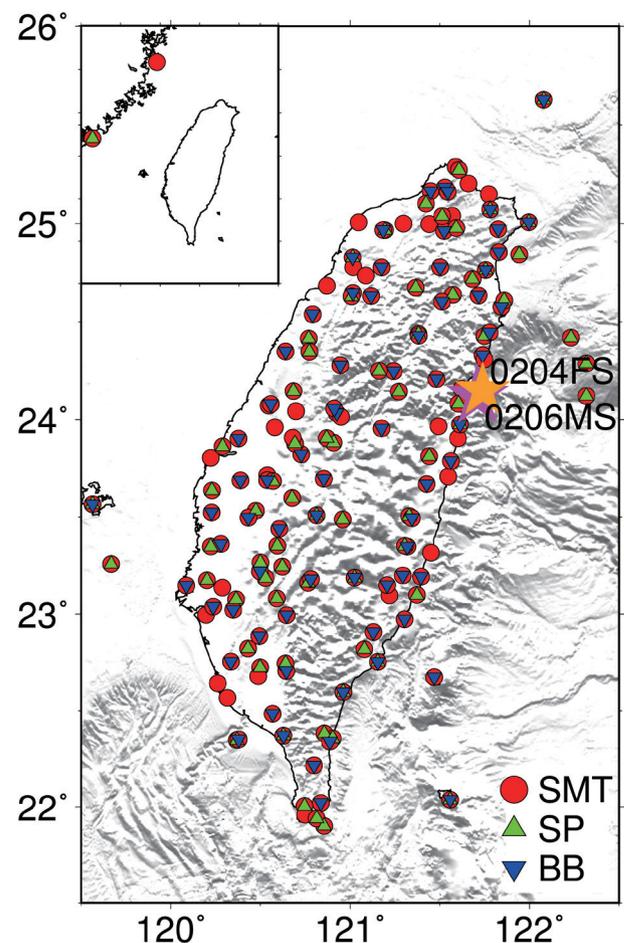


Fig. 1. Map showing the distribution of short-period (SP), broadband (BB), strong-motion (SMT) stations from the CWB seismic network. The star symbols denote the epicenters of the largest foreshock (0204FS) and mainshock (0206MS) of the Hualien earthquake sequences.

toolbox for processing seismological data (Beyreuther et al. 2010; Megies et al. 2011). Therefore, it supports all common data formats used in the seismology community and others with the time-series forms, and provide clients to access data directly from the Earthworm through waveform web service. The algorithm is first tested and applied to the 2018 Hualien foreshock-mainshock-aftershock sequence, which we use to demonstrate the potential capability and applications of the algorithm in real-time monitoring and locations of earthquake activities in Taiwan.

2. METHOD

2.1 Phase Picker

The PhasePapy package provides three different algorithms for automatic phase identification: the transient energy frequency-band picker (FBpicker) (Lomax et al. 2012), the AICDpicker, and the KTpicker (Chen and Holland 2016). The first method measures the relative change of transient energy between consecutive windows, while the last two respectively calculate the first derivative and kurtosis of the Akaike information criterion (AIC) function (Akaike 1974) to identify phase arrivals. We have applied each of them to waveform data from the CWB seismic network that contain phase arrivals from a collection of known events. Comparisons of the picking results demonstrate that the AICDpicker achieves better performance with consideration of accuracy and efficiency. Therefore, we implement the AICDpicker in our automatic picking algorithm.

The principle of the AIC picker is to find a global minimum of the AIC function of single or multiple component seismograms. For a discrete time-series seismogram $s(t)$ of length N , it can be subdivided into two segments by a moving division point D ($0 \leq D \leq N$). A simple way to calculate the AIC function directly from the variance of the divided time series is expressed as follows (Maeda 1985; Sleeman and Van Eck 1999)

$$\text{AIC}(D) = D \log\{\text{var}(s[1, D])\} + (N - D - 1) \log\{\text{var}(s[D + 1, N])\}, \quad (1)$$

where $\text{var}(s[1, D])$ and $\text{var}(s[D + 1, N])$ represent the variances of the two subdivided segments of a seismogram. If the entire seismic record of investigation only contains one P-wave phase, the choice of D that leads to a global minimum of the AIC function would simply give the corresponding phase pick. In reality, a seismogram is often composed of both the P- and S-waves excited from a single or multiple events occurring close in time to each other. It can become tricky and time-consuming to determine seismic phase arrivals by searching global or more than one local minima of the AIC function. Therefore, the PhasePapy alternatively adopts the absolute value of the first deriva-

tive (gradient) of AIC function as the characteristic function (CF) to identify the phase arrivals. Only those exceeding the dynamic threshold (DT) level, obtained with the short-time to long-time moving average of the CF, namely STA/LTA, are the accepted phase picks. In Fig. 2, we use 7-minute long, 3-component accelerogram data recorded by a strong motion station in the CWB network (Fig. 2a) to illustrate our automatic phase picking and earthquake location procedures. The AIC function for each component and the respective characteristic function and dynamic threshold are displayed in Fig. 2b.

2.2 Associator Algorithm

After the phases are picked, an Associator algorithm is employed to ascribe them to specific earthquake events. Likewise, the PhasePapy has two built-in Associators: the 1D Associator and the 3D Associator. Both use pre-calculated travel-time tables of P and S arrivals to determine the earthquake location associated with the picked arrivals. However, none of them meets our needs. The hypocentral depth is fixed in the 1D Associator which is obviously inappropriate for a wide range of depths of earthquakes in the Taiwan region. The 3D Associator is designed for a small seismic network with tens of stations only, which neither is suitable for the dense CWB network containing over two hundred stations and more than four hundred seismograph instruments. The computational time required to extract travel-time information from a large database would be very demanding and unacceptably long. For all of the above reasons, we have modified the original algorithm to make the ‘‘Associator’’ procedure more efficient using the following steps: pick amalgamation, candidate event creation, earthquake origin time analysis, and event-phase association/location.

The first step is to amalgamate the picks from all the components of the same station. The same station name with different location codes are considered as different stations. Though P and S arrivals are most unambiguously picked on the vertical and horizontal component, respectively, the automatic phase picker sometimes identifies the same phase on the counterpart component. To remove redundant picks of the same phase, the original picks arriving within a time interval less than one second on different components are coalesced into one single modified pick (top panel in Fig. 2c). Only the first arrival time of the amalgamated modified pick is determined for the following event association analysis. This procedure would reduce the number of the original picks and greatly improve the computational efficiency of the Associator.

The second step is to create the candidate events for each station from all possible combinations of pairs of the modified picks found in the computed CF time series, but limited to those with their arrival-time differences no greater than 80 s (bottom panel in Fig. 2c). Considering that the earlier

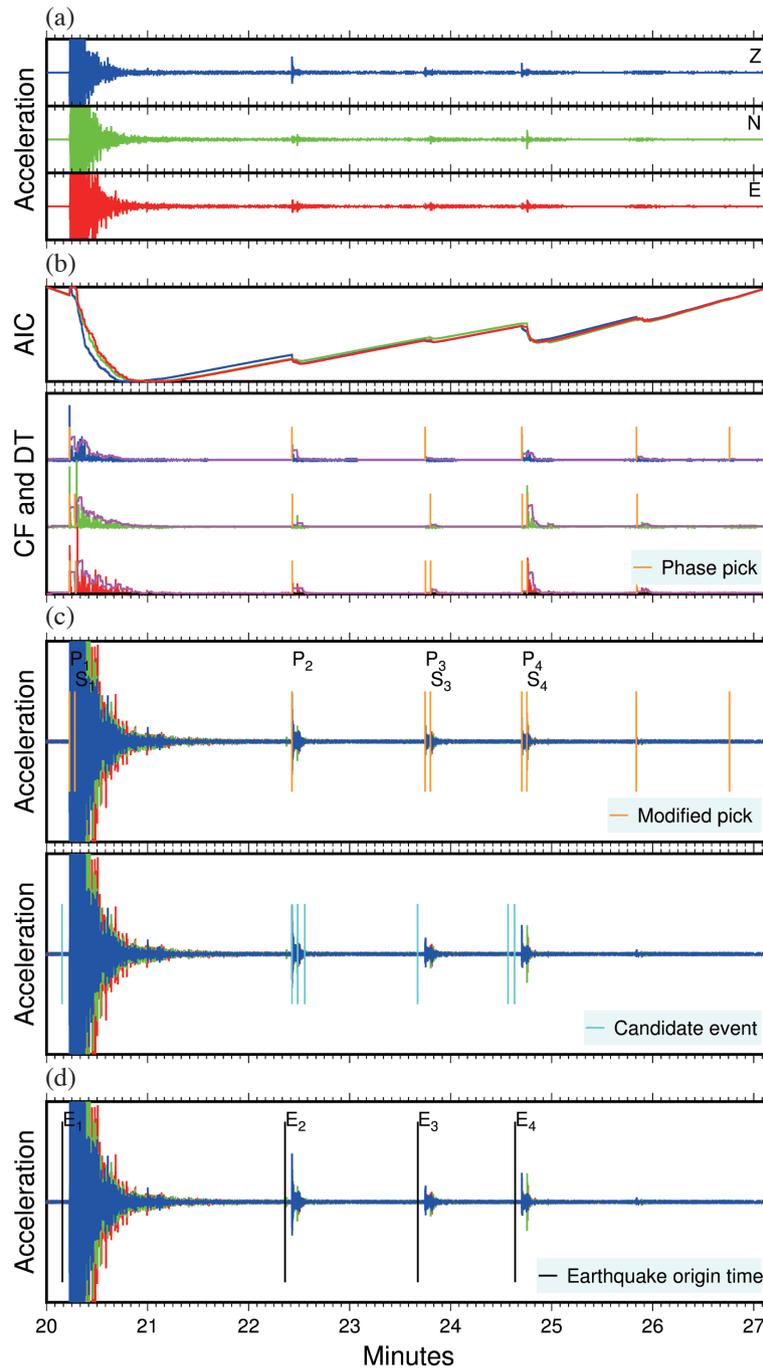


Fig. 2. (a) Examples of accelerograms recorded on the vertical (Z), north (N), and east (E) components of the strong-motion station, ALS, shown by blue, green, and red lines, respectively. (b) The AIC function of each seismic trace (top panel) and the corresponding characteristic function (CF) (bottom panel) obtained by taking the absolute value of the first derivative of AIC. The dynamic threshold (DT) calculated by STA/LTA of the CF is drawn with the magenta curve. Only those picks exceeding the DT level are the accepted phase arrivals marked with the orange vertical lines. (c) The modified picks (top panel) and the origin times of the candidate events (bottom panel) marked on the 3-component accelerograms with the orange and cyan vertical lines, respectively. The modified picks are obtained by amalgamating the phase picks arriving close in time within 1 s on all the components of the same station, while the origin times of the candidate events are estimated by the arrival-time differences between pairs of the modified picks no larger than 80 s. (d) The origin times of the four real earthquakes (E₁-E₄) marked on the 3-component accelerograms with the black lines, determined by the final source parameter inversion using the P and S phase picks shown in (c) at all the recording stations. Note that the S-arrival of the event E₂ is not picked at this station and thus not marked in (c). However, it can be observed and picked at the other stations used for the earthquake location.

and later arrivals from a given pair of the modified picks are associated with the P and S phase arrivals of the same event, their time difference equal to the S-P time can be directly used to determine the P arrival time and thus the origin time of the event through the Wadati diagram (Wadati and Oki 1933), which is constructed a priori by finding the best-fit linear relationship between the observed S-P times and P arrival times of local earthquakes occurring before. For each station, we choose the historical events which have both the P and S arrival-time picks reported in the CWB phase database to obtain the linear regression relation. Figure 3 shows a typical Wadati diagram plotting the P against the S-P arrival times from the events between July and December of 2016 recorded at station ALS and the best least-squares fit line through all the data points. It is worth mentioning that the P arrival times predicted from the linear regression equation and used to determine the origin times of the candidate events would deviate from the actually observed arrival times because there exists inconsistency between the observed and predicted S-P times with the root-mean-square (RMS) error of ~ 1.47 s due to data noise, phase picking errors, and uncertainty of the velocity structure used in earthquake location. However, such deviation won't cause misidentification of the candidate events unless the origin times of the successive candidate events are too close less than two seconds, which would be difficult to distinguish them separately even by an experienced analyst. The actual origin time of each potential real earthquake together with its hypocenter and local magnitude will be determined simultaneously in the final step of earthquake location using the arrival times of the modified picks and the same 3-D velocity model (see details in the following introduction).

On the other hand, because of the inevitable uncertainties in the picked P and S arrival times, the modified picks recorded at different stations and associated with the actually same earthquake are normally assigned to a cluster of candidate events with the origin times all very close to that of the actual earthquake. In particular, the earthquakes with larger magnitude and/or well recorded by the dense stations nearby would yield a substantial number of the candidate events. Therefore, the following step is to merge those clustering around the peak number of the candidate events into a single trial event and collect all the associated modified picks to locate the respective hypocenter.

After assigning the origin times of the candidate events detected by all the stations and putting one marker for each of them at the specified origin time along the time axis, the algorithm would employ a moving window of 3-s in length shifted by 0.01 s, the sampling time interval of the seismograms, and count the total markers, i.e., the number of the candidate events in every 3-s window, which are shown as a function of time in Fig. 4. If the total number exceeds a given threshold, it will consider a trial earthquake occurring at this time instant. Ideally, the minimum threshold is three

since at least three stations are required to determine the earthquake location. However, the number of the clustered candidate events largely depend on the earthquake magnitude and station distribution around the event epicenter. Thus, rather than using a global fixed threshold, we choose local dynamic thresholds to make the event location procedure more computationally efficient. The threshold is set to increase with the total numbers of the candidate events and stations recording the paired modified picks and tested with some earthquake data by trial and error. We then choose to pre-specify the least number of stations recording the paired modified picks for a given number range of candidate events as a threshold to declare a trial earthquake for the following location procedure. In general, the larger the total number of candidate events, the higher the least number of the stations with the paired modified picks or the acceptance threshold number is.

The final step is to locate the hypocenter of each trial earthquake at the time instant when the total number of the counted candidate events exceeds the local threshold number. It starts with the one which has the maximum number of the candidate events by collecting all the pairs of the modified picks associated with those events occurring within this particular 3-s time interval. The respective auto-picked P and S arrival times are then given as input to an external location program, HYPO71 (Lee and Lahr 1972) to invert for the source parameters, the origin time and hypocenter of the trial earthquake using the 3-D P and S velocity models obtained from local earthquake tomography (Wu et al. 2007). If the resulting solution passes all the quality control checks, that is, the RMS errors in the origin time, epicenter, and depth of the located event have to be less than 9 s, 10 km, and 10 km, respectively, the Associator assigns it as a real earthquake and calculates the predicted P and S arrival times at all the stations.

In some cases as exemplified in Fig. 2c, the automatic phase picker may solely detect either the P or S picks at a few stations which are unable to form pairs of the modified picks used in the previously-described location procedure. To include the arrival-time data from these unpaired modified picks which may help improve the azimuthal coverage of stations and the estimated hypocenter, the algorithm further finds the unpaired modified picks that arrive closest to and within a time interval centered around the predicted P or S arrival times based on the source parameters inverted with the paired modified picks only. The width of the time interval is set as a fraction (10%) of the predicted arrival time. By adding more data from the unpaired modified picks, the HYPO71 program is rerun to obtain the final origin time and hypocenter of the earthquake. The Associator and earthquake location are then repeated for the remaining modified picks and candidate events. However, before taking the second largest number of the candidate events as the next trial earthquake, in order to avoid producing the duplicate events, the algorithm

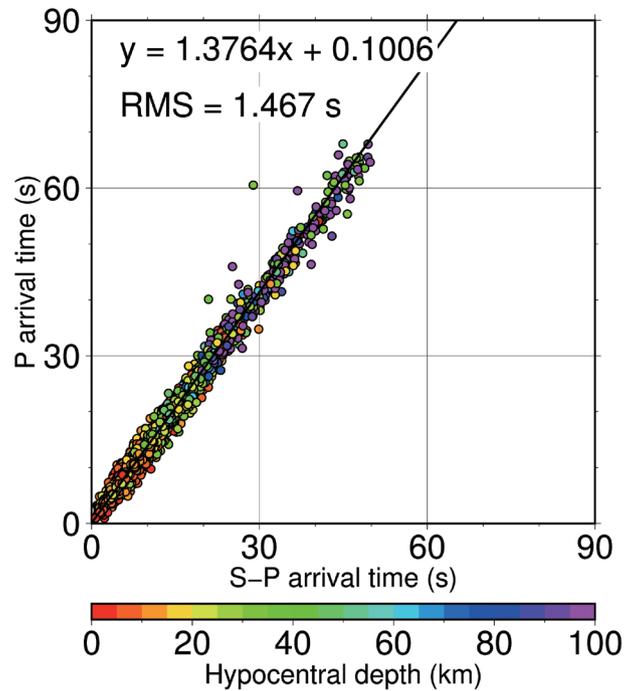


Fig. 3. A “reverse” Wadati diagram plotting the P arrival time against the S-P time of the events between January and June of 2016 recorded at ALS station from the CWB event phase catalog. The equation of the best-fitting straight line through these data points is used to quickly determine the origin time and P phase arrival time of each candidate event from a given S-P time obtained with the separation time between a pair of the modified picks.

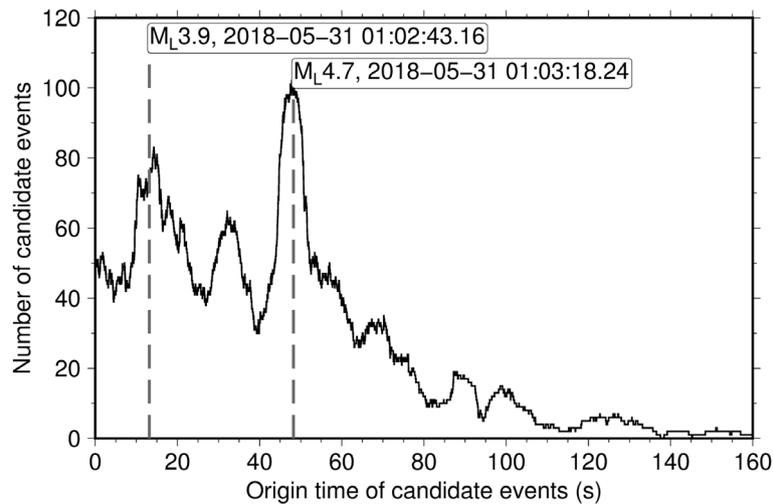


Fig. 4. The total number of the candidate events detected by all the stations counted as a function of event origin time estimated from the pairs of the modified picks and the Wadati diagram. The two largest peaks correspond to the two biggest earthquakes that occurred in Taitung ($M_L 4.7$) and the Nanao Basin ($M_L 3.9$) during the processing time period spanning from GMT 01:02:30 to 01:05:10 on 31 May 2018.

would first exclude those with the origin times that fall within a pre-specified time window centered at the origin time of the earthquake being located in the previous step. Considering the width of the peaks of the event counts shown in Fig. 4 and detection limitations for the phase arrivals of the later event mixed or hidden within stronger long coda waves of the preceding large earthquake which are seen in the waveform data, we choose 10 s for the half-width of the time window. The whole procedure is conducted iteratively until either all the modified picks have been assigned to the located earthquakes or the number of stations detecting the paired modified picks for the given number of the candidate events remained are all lower than the specified dynamic threshold.

3. ALGORITHM TESTING

To demonstrate the performance of our automatic phase picker and event association/location algorithm, we first validate the accuracy and improved efficiency of our modified algorithm by running a comparison test with the 3D PhasePapy package. Both the algorithms are employed to automatically pick phase arrivals and associate them with located earthquakes using the same waveform data, which comprises seven hundreds of ~7-minute long, 3-component seismic traces recorded by the CWB stations (Fig. 2a). The CPU times spent on the eight processors of Intel Xeon E5-2690 v4 with 8G RAM running the AICDpicker and Associator are summarized in Table 1a. It is obvious that the efficiency of the automatic picker is substantially improved by replacing the original Python module with a Fortran code. The Associator that adopts the linear regression relation between S-P versus P arrival times of historical events also speeds up by almost a factor of 3. Both the algorithms detect 4 out of 5 manually located earthquakes during the 7-minute seismic records. The origin times and hypocenters of four detected earthquakes determined by the two automatic algorithms and human analysts are compared in Table 1b.

The second test is that we apply our algorithm to systematically detect the earthquakes during the first 6 months of 2017 and compare the results with those from routine event locations by the CWB analysts. Among 16675 earthquakes reported in the CWB catalog, our algorithm detects 10169 events. Figure 5 shows the cumulative percentage and number of the automatically detected earthquakes of a given magnitude or greater. The percentage rate of detection for the events within each 0.1 magnitude interval is also displayed for comparison. For those larger events with $M_L \geq 4.0$, 100% of them can be located by our algorithm. For the events with $M_L \geq 2.0$, the detection rate can still reach about 75%.

Comparing the epicentral distributions between the detected and undetected earthquakes shown in Fig. 6, we find that most of manually located events missed by our automatic algorithm occurred in the offshore areas of Taiwan,

primarily because the azimuthal gap of the recorded stations exceeding 180° often fails to obtain the qualified hypocentral solutions. Some of the undetected events occurred inland Taiwan, mostly associated with aftershock clusters or earthquake swarms which have relatively small magnitudes ($M_L < 3$). Their corresponding phase arrivals are often too close in time or too weak to be isolated or identified by the automatic picker. We notice that a few larger events with $M_L > 3$ in the south-central Taiwan are undetected inside the network. By checking the waveforms and manual phase picks used in locating these earthquakes, we find their P and S phase arrivals are usually hidden or mixed within the stronger long coda waves of the preceding large earthquake. For such situations, the analysts can judge to apply a suitable filter to waveform data to enhance the phase signals of the later event for manual picking, but there would be a large uncertainty in the estimated magnitude.

In Fig. 7, we compare P and S phase arrival-time picks, local magnitudes, and hypocentral locations of the detected earthquakes determined by our automatic algorithm and human analysts. The average arrival-time difference between the automatic and manual picks and the respective standard deviation are 0.1622 ± 0.346 s for 148568 P and 0.100 ± 0.465 for 120007 S phase picks (Figs. 7a - b). The local magnitude estimated by the maximum amplitude on the horizontal components between the automatic and manual picks generally agrees with each other except for the events with magnitude smaller than 1.5 where the automatic algorithm tends to overestimate the magnitude (Fig. 7c). The differences in longitude, latitude, and depth of automatically and manually determined hypocenters are on average 3 - 4 km in the lateral directions ($0.003 \pm 0.040^\circ$ in longitude and $0.0006 \pm 0.032^\circ$ in latitude), and about 7 km in depth (-0.366 ± 6.98 km).

4. APPLICATION TO THE 2018 HUALIEN EARTHQUAKE SEQUENCE

Since November of 2017, the novel algorithm has been formally implemented in the CWB automatic event detection system, providing in near real time the origin time, location, and local magnitude of earthquakes in the Taiwan region. On 6 February 2018, a powerful earthquake with $M_L = 6.25$ struck off the northeast coast of Taiwan, causing severe injuries, fatalities, property damages, and financial losses along the shoreline of the Hualien region. It is the largest event ever recorded since the new system started to operate. Within a few hours after this disastrous mainshock hit the Hualien city heavily, hundreds of felt aftershocks were triggered in the initial rupture and surrounding area. As they may further devastate the damaged houses, buildings, and public facilities, not only the academics but the government, media, and public are all eager to know the spatiotemporal distributions of the aftershocks in order to

Table 1. (a) Comparison of the performance efficiency between the 3D PhasePapy and our modified algorithm based on the CPU times spent in running the AICDpicker for automatic phase picking and the Associator for associating the modified picks with the detected earthquakes. Both the algorithms detect 4 out of 5 manually located earthquakes during ~7-minute long seismic records shown in Fig. 1. (b) Comparison of the origin times, epicenters, and depths of the four detected earthquakes determined manually and by the two automatic algorithms.

(a)

	3D PhasePapy	Modified algorithm
AICDpicker	4394 s	72 s
Associator	43 s	16 s
# of events: Auto (Manual)	4 (5)	4 (5)

(b)

	Origin Time (hr:mn:sec)	Epicenter (lat, lon)	Depth (km)
Manual	14:20:09.40	(23.58°N, 120.72°E)	15.24
	14:22:21.39	(23.59°N, 120.72°E)	14.32
	14:23:40.39	(23.58°N, 120.71°E)	15.58
	14:24:38.03	(23.60°N, 120.74°E)	13.72
Modified algorithm	14:20:09.45	(23.58°N, 120.72°E)	14.81
	14:22:21.64	(23.60°N, 120.72°E)	11.77
	14:23:40.49	(23.58°N, 120.72°E)	12.52
	14:24:38.34	(23.60°N, 120.74°E)	11.41
3D PhasePapy	14:20:08.76	(23.55°N, 120.75°E)	5.0
	14:22:21.65	(23.55°N, 120.75°E)	15.0
	14:23:40.52	(23.55°N, 120.75°E)	15.0
	14:24:38.14	(23.65°N, 120.75°E)	15.0

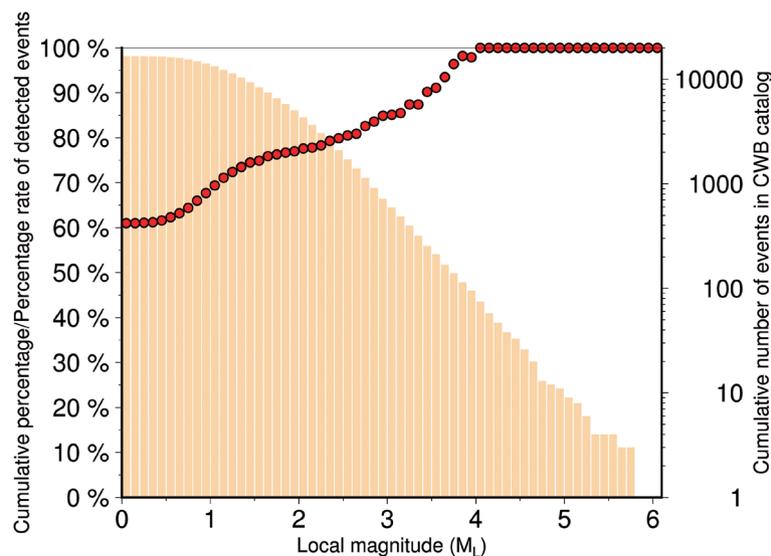


Fig. 5. The cumulative percentage (left y-axis) and number (right y-axis) of earthquakes of a given magnitude or greater detected by our automatic algorithm, which are counted in every 0.1 magnitude bin and recorded between January and June of 2017. The red dots indicate the percentage rate of detection with respect to the number of the manually located events.

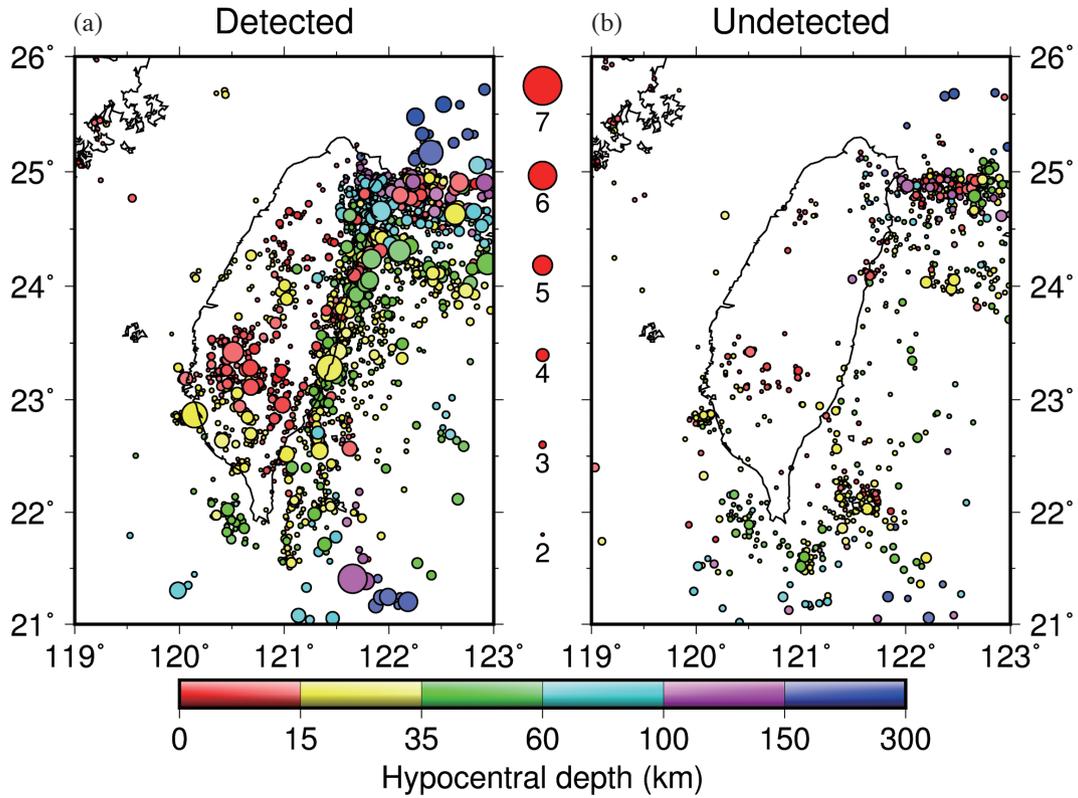


Fig. 6. Epicenters of earthquakes with local magnitude equal to or greater than 2.0 that occurred between January and June of 2017 and were (a) detected and (b) missed by our automatic location algorithm.

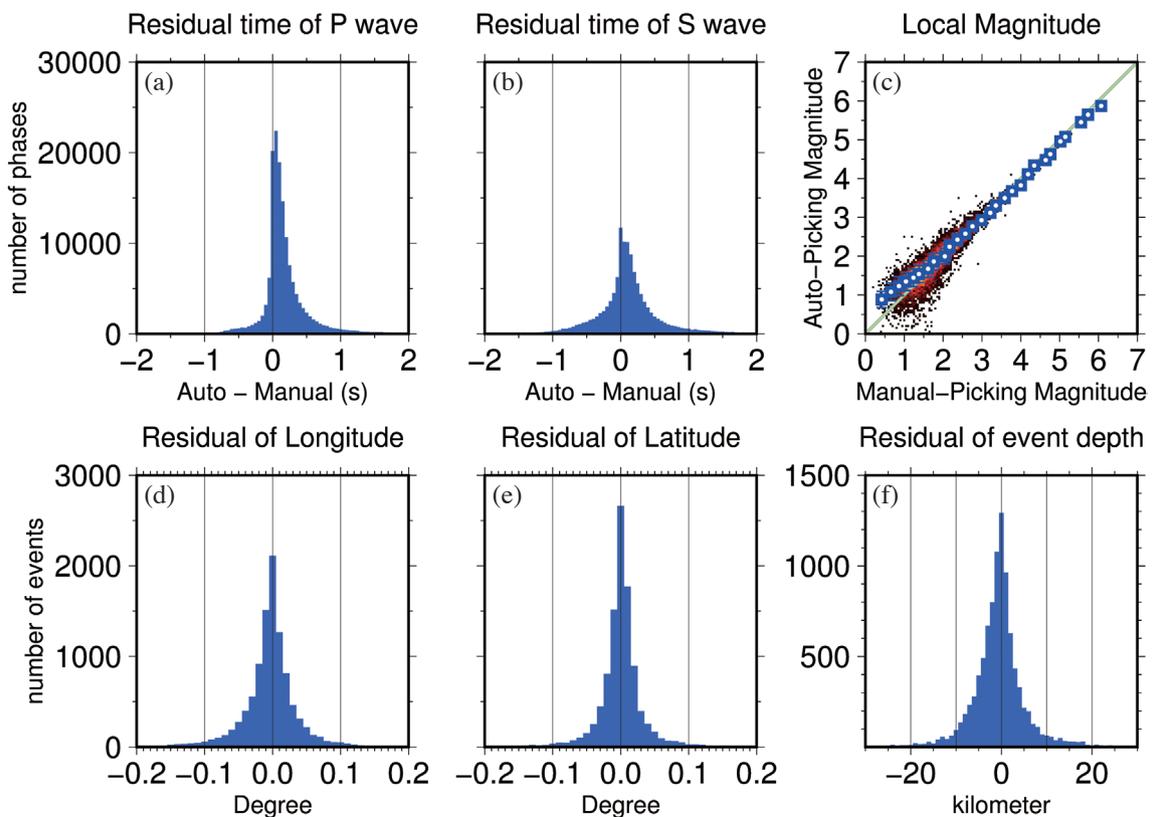


Fig. 7. Comparison of (a) - (b) P and S phase arrival-time picks, and (c) determined local magnitudes, (d) - (e) epicenters, and (f) depths of our automatically detected earthquakes between January and June of 2017 with those from manual picking and location.

quickly make post-earthquake risk assessment and mitigation plans. On the other hand, the Hualien mainshock was preceded by foreshock activities starting about three days earlier with the largest foreshock ($M_L = 5.89$) located about 6 km NNE of the mainshock epicenter on February 4. As the foreshocks and aftershocks are key to unravel the pre-nucleation stage, the geometry of the subsurface causative fault, and the rupture extent of the mainshock, it is of essential importance to provide the community with the rapid and precise information of their hypocentral and magnitude distributions. The novel algorithm offers a timely and efficient solution to respond these urgent needs, reporting near

real-time locations and magnitudes of the 2018 Hualien foreshock-aftershock sequences.

Figure 8 shows the local magnitudes, numbers, detection rates, and spatial distributions of earthquakes located in the Hualien mainshock area during February 2018 by the new and old automatic algorithms, compared with those determined manually. The results demonstrate that the new algorithm surpasses the old one in terms of the minimum magnitude and number of events detected. Compared to the old one, the new algorithm is capable of detecting more aftershocks with the magnitude lower than 1.5 (Fig. 8a). The detection rate calculated by the percentage ratio of the number

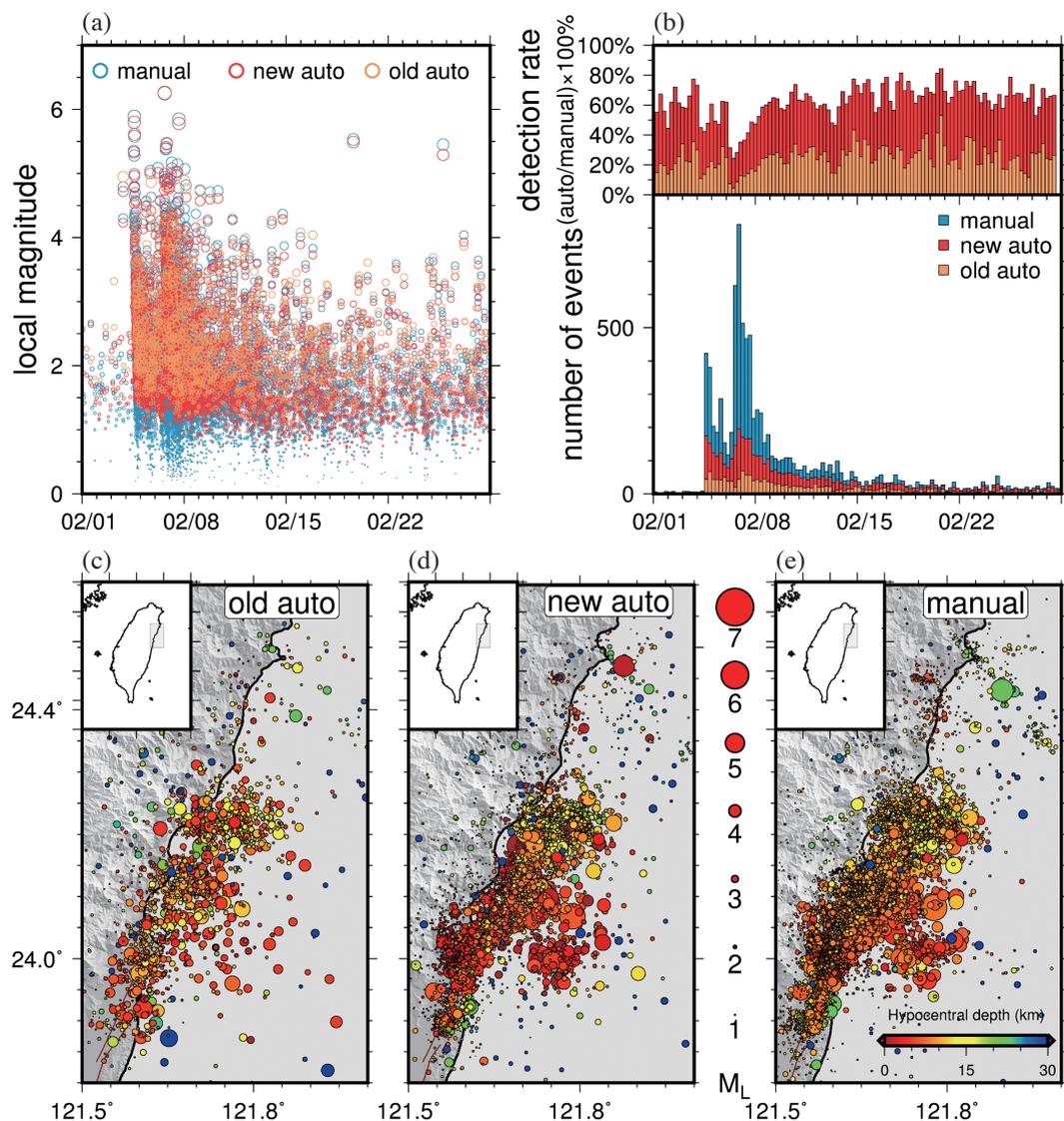


Fig. 8. Local magnitudes, numbers, detection rates, and spatial distributions of the 2018 Hualien foreshock-mainshock-aftershock sequences occurring in February 2018 and determined automatically by the new and old algorithms, compared with those from manual location. (a) Local magnitude versus origin time with the blue, red, and orange open circles corresponding to the earthquakes located manually and by the new and old automatic algorithms, respectively. (b) The number of earthquakes counted in every 6-hr bin (bottom panel), which are detected manually and by the new and old automatic algorithms shown with the blue, red, and orange bars, respectively. The corresponding detection rates calculated by the percentage ratio of the number of events from the automatic to the manual location are displayed in the top panel. (c) - (e) Spatial distributions of the earthquakes with their hypocenters and local magnitudes determined by the old and new automatic algorithms and manual location, respectively. The color and size of circles are respectively scaled with the hypocentral depths and magnitudes estimated by each location method.

of events from the automatic relative to the manual location is also effectively improved and increased almost by a factor 3 (Fig. 8b). In particular, the old algorithm seems to miss a large number of aftershocks clustering in time accompanying the occurrence of the largest foreshock and mainshock with a sudden drop of the detection rate below 10% (the top panel of Fig. 8b). Though the detection rate of the new algorithm is also degraded below 50% during the bursts of the aftershocks in the first few hours to one day after the mainshock occurrence which leaves room for improvement, the aftershock activity generally following a power-law decay with time (Omori 1894) is better characterized by the time sequence of the aftershocks detected by our new algorithm (the bottom panel of Fig. 8b).

In terms of fast hypocenter determinations, the epicenters of the aftershocks obtained with the new algorithm (Fig. 8d) appear to be less scattered and distributed along some linear trends similar to those seen in the manually located aftershocks (Fig. 8e). In Fig. 9, we further make quantitative comparisons of the P and S arrival-time picks, magnitudes, and hypocentral locations of the Hualien earthquake sequences detected in February by our automatic algorithm with those by the human analysts. The differences between the picked phase arrival times, the determined source parameters and magnitudes all exhibit similar distributions as those seen in the previous 6-month test period (Fig. 7). However, we notice the residuals between the automatic and manual P and S picks shows much larger peaks at zero times (Figs. 9a - b), suggesting that many automatic phase picks checked by the manual analysts are considered to be accurate enough and directly adopted in the later manual location. If the automatic algorithm is further improved, we expect that it can greatly speed up the manual phase picking and location procedures particularly during the seismically very active period following the occurrence of a large mainshock.

To better illuminate the aftershock distributions associated with the largest foreshock and mainshock, we plot the 2-day aftershocks following the largest foreshock but before the mainshock occurrence and the aftershocks after the mainshock until the end of February 2018 on the cross sections cut along the strikes of two nodal planes of the foreshock in Fig. 10b and the strike of the fault plane of the mainshock in Fig. 10c. The nodal plane orientations are determined from the point-source moment tensor inversion using regional broadband waveform data (Jian et al. 2018a). The depths of the two-day aftershocks following the largest foreshock show a gradual decrease from northeast off the east coast toward the southwest closer to the shoreline (Fig. 10b). We thus inferred that the foreshock probably nucleated and ruptured on an NE-dipping low-angle fault plane, in general agreement with that determined by a more sophisticated back-projection (BP) rupture imaging method (Jian et al. 2018b). On the other hand, the distribution of

the aftershocks occurring within the first few hours forms a subhorizontal streak extending southwestward, implying that the mainshock rupture primarily propagated horizontally toward the SW direction, which is also consistent with the focal mechanism of the mainshock with the predominant NE-SW striking, left-lateral strike-slip motion and the rupture propagation history revealed by the BP analysis (Jian et al. 2018b).

5. DISCUSSION AND CONCLUSION

Toward real-time monitoring of very frequent earthquake activities in Taiwan using the dense, island-wide seismic networks operated by the CWB, we have developed a rapid and precise algorithm based on PhasePapy developed in Oklahoma Geological Survey for automatic phase identification and event location, which improves the performance efficiency and limitations of the network size and number of stations of the original Python-interface package. The algorithm mainly consists of two parts: the AIC-based phase picker which identifies and picks phase arrival times and the Associator which assigns the pairs of picks to candidate events and then isolates the local peaks of the numbers of the clustered events for further location of real earthquake events.

We have tested our automatic algorithm with the database of the CWB earthquake catalog between January and June of 2017. The results demonstrate that our algorithm is capable of detecting 100% of the reported earthquakes with $M_L \geq 4$ and 75% of the earthquakes with $M_L \geq 2$ in the selected time period. The difference in the determined local magnitude is quite small except for the small events with $M_L < 1.5$ that reach the detection threshold of the Taiwan network (Chen et al. 2016). The statistics show that the differences between automatically and manually picked P and S phase arrival times and located hypocenters are a few tenths of one second and a few kilometer, respectively, on the same order of the manually-picking errors and uncertainties of event locations.

Our algorithm is also capable of detecting in near real time the bursts of aftershock activities within the first few hours to one day following the largest foreshock and mainshock of the 2018 Hualien earthquake sequence. The temporal distributions of the detected events clearly exhibits the frequency of the aftershocks decays with the reciprocal of time elapsed since the mainshock. The locations and magnitudes of the aftershocks provide the almost synchronized guides to quickly distinguish the fault-plane ambiguity, diagnose the source characteristics and rupture extent of the foreshock and devastating mainshock as well as assess seismic hazards and make risk mitigation decisions.

Currently our algorithm has not yet been designed to run in complete real time. This is partly because too many stations and channels in the CWB network which comprises

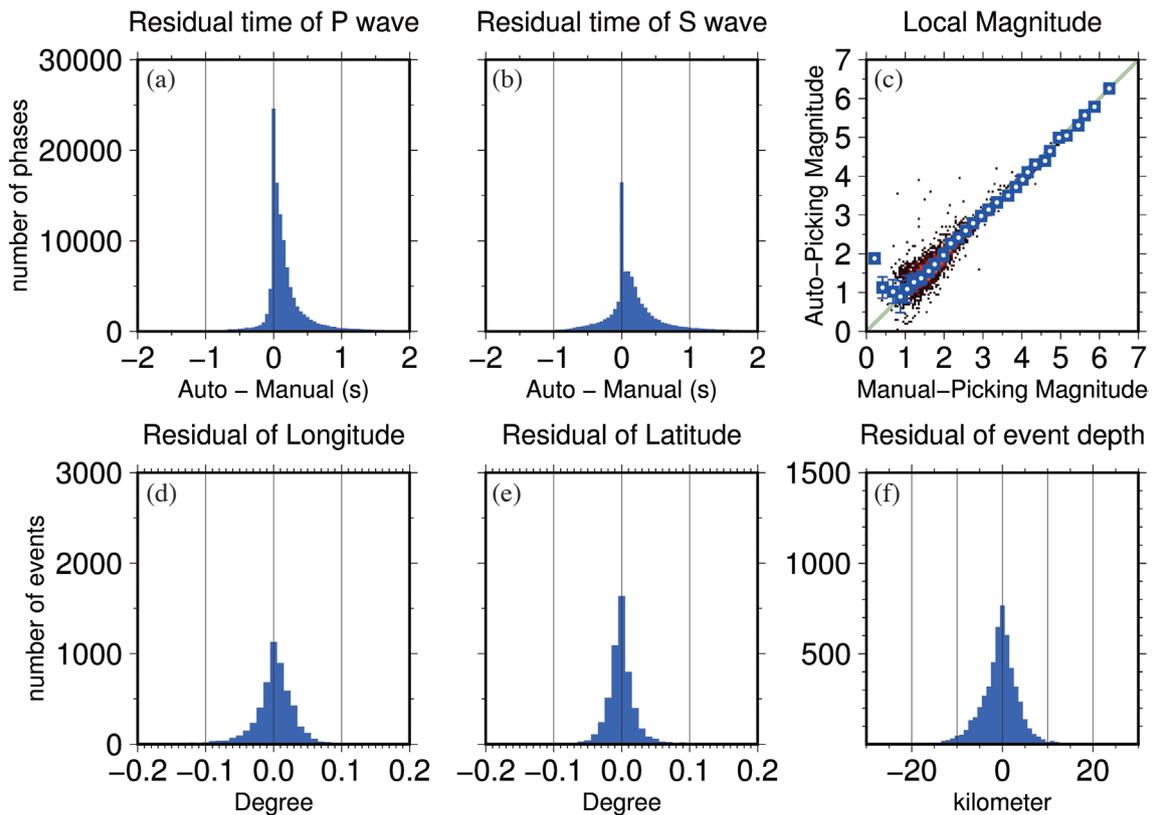


Fig. 9. Comparison of P and S phase arrival-time picks (a) - (b), and determined local magnitudes (c), epicenters (d) - (e), and depths (f) of the Hualien earthquake sequences during February 2018 from our automatic algorithm with those determined manually.

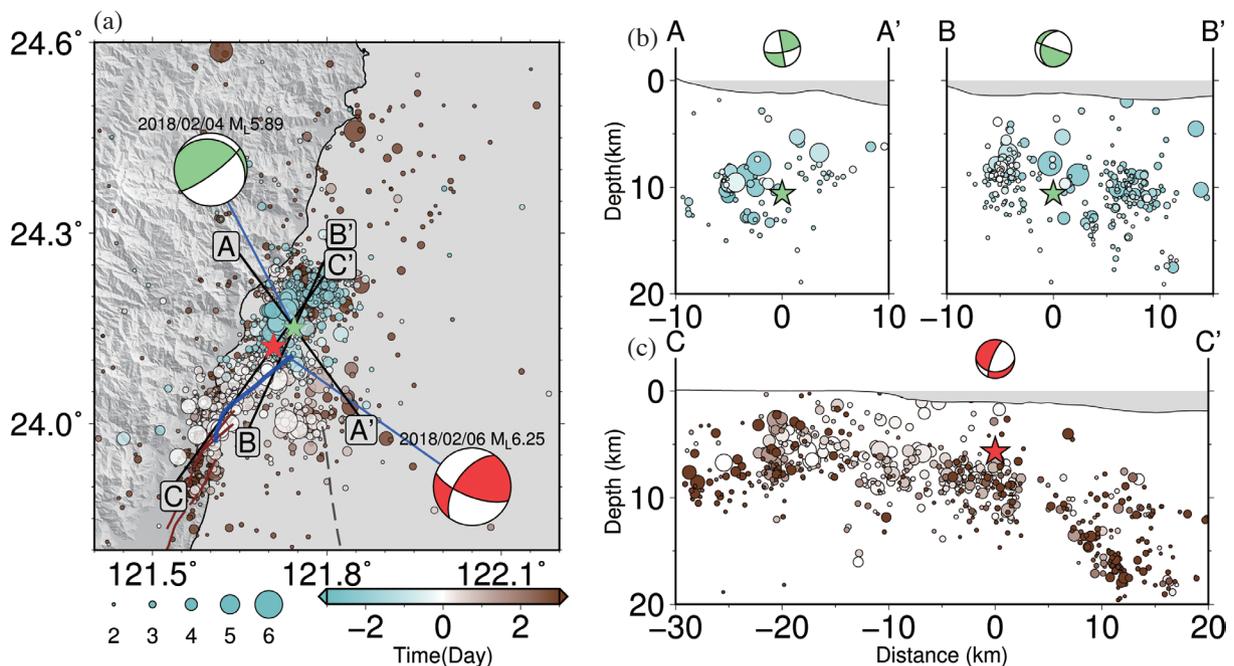


Fig. 10. (a) Map and (b) - (c) cross section views of the spatial distributions of the 2018 Hualien foreshock-mainshock-aftershock sequences with $M_L \geq 2$ automatically located by the new algorithm. The dashed and blue curves denote the Ryukyu trench and Milun fault considered as the causative fault of the mainshock (Central Geological Survey 2018), respectively. The AA' and BB' cross sections cut parallel to the strike of the fault plane and conjugate nodal plane of the 0204 largest foreshock, while the CC' is parallel to the strike of the fault plane of the 0206 main shock verified by the BP rupture images (Jian et al. 2018b). The focal mechanism diagrams of the foreshock and mainshock shown on the map and cross sections are viewed from the lower-hemisphere and side projection, respectively.

of short-period, broadband, and strong-motion instruments such that the AICDPicker and the Associator cannot run concurrently in a synchronized fashion. Instead that continuous seismic data are fed into the algorithm directly, seismic waveforms containing pulse-like oscillations considered to be triggered by possible seismic events are cut and stored into short time-series files for the automatic phase picker and event detection. Besides, the false triggered or real phases from multiple events arriving at the same time may confuse the Associator to create redundant candidate events for real earthquake locations. The current cluster analysis of candidate events provides a simple way to separate and identify the multiple events occurring close in time in a local region such as the peak aftershock activities following the 2018 Hualien mainshock. Once we modify the phase picker to individually identify the P and S phase arrival times, we can avoid the identification of the false candidate events in the cluster analysis and largely improve the processing speed of the Associator and the capability of smaller earthquake detection in the future.

Acknowledgements We thank the Central Weather Bureau (CWB) of Taiwan for providing seismogram data and event catalog used to test our algorithm. Comments from two anonymous reviewers and the Guest Editor, Ruey-Juin Rau, greatly help improve the manuscript. This work was partially supported by the Ministry of Science and Technology of Taiwan under grants 106-2116-M-002-009 (S.-H. Hung) and 106-2116-M-052-001 (Y.-L. Chen).

REFERENCES

- Akaike, H., 1974: A new look at the statistical model identification. *IEEE Trans. Automat. Contr.*, **19**, 716-723, doi: 10.1109/TAC.1974.1100705. [[Link](#)]
- Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann, 2010: ObsPy: A Python toolbox for seismology. *Seismol. Res. Lett.*, **81**, 530-533, doi: 10.1785/gssrl.81.3.530. [[Link](#)]
- Bogiatzis, P. and M. Ishii, 2015: Continuous wavelet decomposition algorithms for automatic detection of compressional- and shear-wave arrival times. *Bull. Seismol. Soc. Am.*, **105**, 1628-1641, doi: 10.1785/0120140267. [[Link](#)]
- Central Geological Survey, 2018: The 20180206 Hualien Earthquake Geological Investigation Report, MOEA. (in Chinese)
- Chen, C. and A. A. Holland, 2016: PhasePapy: A robust pure Python package for automatic identification of seismic phases. *Seismol. Res. Lett.*, **87**, 1384-1396, doi: 10.1785/0220160019. [[Link](#)]
- Chen, D.-Y., T.-L. Lin, H.-C. Hsu, Y.-C. Hsu, and N.-C. Hsiao, 2019a: An approach to improve the performance of the earthquake early warning system for the 2018 Hualien earthquake in Taiwan. *Terr. Atmos. Ocean. Sci.*, **30**, 423-433, doi: 10.3319/TAO.2018.12.23.02. [[Link](#)]
- Chen, P.-F., Y.-L. Chen, P.-L. Su, Y.-D. Peng, and L.-F. Chen, 2019b: Understanding the 6 February 2018, Hualien earthquake sequence through catalog compilation. *Terr. Atmos. Ocean. Sci.*, **30**, 399-409, doi: 10.3319/TAO.2018.11.15.02. [[Link](#)]
- Chen, Y.-L., S.-H. Hung, J.-S. Jiang, and L.-Y. Chiao, 2016: Systematic correlations of the earthquake frequency-magnitude distribution with the deformation and mechanical regimes in the Taiwan orogen. *Geophys. Res. Lett.*, **43**, 5017-5025, doi: 10.1002/2016GL069020. [[Link](#)]
- Gentili, S. and A. Michelini, 2006: Automatic picking of P and S phases using a neural tree. *J. Seismol.*, **10**, 39-63, doi: 10.1007/s10950-006-2296-6. [[Link](#)]
- Hsu, Y. J., M. Ando, S. B. Yu, and M. Simons, 2012: The potential for a great earthquake along the southernmost Ryukyu subduction zone. *Geophys. Res. Lett.*, **39**, doi: 10.1029/2012GL052764. [[Link](#)]
- Huang, S.-Y., J.-Y. Yen, B.-L. Wu, I.-C. Yen, and R. Y. Chuang, 2019: Investigating the Milun Fault: The coseismic surface rupture zone of the 2018/02/06 M_L 6.2 Hualien earthquake, Taiwan. *Terr. Atmos. Ocean. Sci.*, **30**, 311-335, doi: 10.3319/TAO.2018.12.09.03. [[Link](#)]
- Jian, P.-R., S.-H. Hung, L. Meng, and D. Sun, 2017: Rupture characteristics of the 2016 Meinong earthquake revealed by the back projection and directivity analysis of teleseismic broadband waveforms. *Geophys. Res. Lett.*, **44**, 3545-3553, doi: 10.1002/2017GL072552. [[Link](#)]
- Jian, P.-R., T.-L. Tseng, W.-T. Liang, and P.-H. Huang, 2018a: A new automatic full-waveform regional moment tensor inversion algorithm and its applications in the Taiwan area. *Bull. Seismol. Soc. Am.*, **108**, 573-587, doi: 10.1785/0120170231. [[Link](#)]
- Jian, P.-R., S.-H. Hung, and L. Meng, 2018b: Rupture behavior and interaction of the 2018 Hualien earthquake sequence and its tectonic implication. *Seismol. Res. Lett.*, **90**, 68-77, doi: 10.1785/0220180241. [[Link](#)]
- JMA, 2017: The 7th Technical Committee Meeting of Earthquake Early Warning Evaluation and Improvement Study. Available at <https://www.data.jma.go.jp/svd/eqev/data/study-panel/eev-hyoka/t07/index.html>. (in Japanese)
- Kanamori, H., L. Ye, B.-S. Huang, H.-H. Huang, S.-J. Lee, W.-T. Liang, Y.-Y. Lin, K.-F. Ma, Y.-M. Wu, and T.-Y. Yeh, 2017: A strong-motion hot spot of the 2016 Meinong, Taiwan, earthquake ($M_w = 6.4$). *Terr. Atmos. Ocean. Sci.*, **28**, 637-650, doi: 10.3319/TAO.2016.10.07.01. [[Link](#)]
- Lee, W. H. K. and J. C. Lahr, 1972: HYPO71: A computer program for determining hypocenter, magnitude, and

- first motion pattern of local earthquakes. Open-File Report 72-224, U.S. Geological Survey, doi: 10.3133/ofr72224. [[Link](#)]
- Leonard, M., 2000: Comparison of manual and automatic onset time picking. *Bull. Seismol. Soc. Am.*, **90**, 1384-1390, doi: 10.1785/0120000026. [[Link](#)]
- Leonard, M. and B. L. N. Kennett, 1999: Multi-component autoregressive techniques for the analysis of seismograms. *Phys. Earth Planet. Inter.*, **113**, 247-263, doi: 10.1016/S0031-9201(99)00054-0. [[Link](#)]
- 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. *Seismol. Res. Lett.*, **83**, 531-540, doi: 10.1785/gssrl.83.3.531. [[Link](#)]
- Maeda, N., 1985: A method for reading and checking phase time in auto-processing system of seismic wave data. *Journal of the Seismological Society of Japan*, **38**, 365-379, doi: 10.4294/zisin1948.38.3_365. [[Link](#)]
- Megies, T., M. Beyreuther, R. Barsch, L. Krischer, and J. Wassermann, 2011: ObsPy – What can it do for data centers and observatories? *Ann. Geophys.*, **54**, 47-58, doi: 10.4401/ag-4838. [[Link](#)]
- Omori, F., 1894: On the aftershocks of earthquakes. *Journal of the College of Science, Imperial University of Tokyo*, **7**, 111-200.
- Shin, T.-C., C.-H. Chang, H.-C. Pu, H.-W. Lin, and P.-L. Leu, 2013: The Geophysical Database Management System in Taiwan. *Terr. Atmos. Ocean. Sci.*, **24**, 11-18, doi: 10.3319/TAO.2012.09.20.01(T). [[Link](#)]
- Sleeman, R. and T. Van Eck, 1999: Robust automatic P-phase picking: An on-line implementation in the analysis of broadband seismogram recordings. *Phys. Earth Planet. Inter.*, **113**, 265-275, doi: 10.1016/S0031-9201(99)00007-2. [[Link](#)]
- Suppe, J., 1984: Kinematics of arc-continent collision, flipping of subduction, and back-arc spreading near Taiwan. *Mem. Geol. Soc. China*, **6**, 21-33.
- Teng, L. S., 1990: Geotectonic evolution of late Cenozoic arc-continent collision in Taiwan. *Tectonophysics*, **183**, 57-76, doi: 10.1016/0040-1951(90)90188-e. [[Link](#)]
- Vassallo, M., C. Satriano, and A. Lomax, 2012: Automatic picker developments and optimization: A strategy for improving the performances of automatic phase pickers. *Seismol. Res. Lett.*, **83**, 541-554, doi: 10.1785/gssrl.83.3.541. [[Link](#)]
- Wadati, K. and S. Oki, 1933: On the travel time of earthquake waves. (Part II). *J. Meteorol. Soc. Jpn.*, **11**, 14-28, doi: 10.2151/jmsj1923.11.1_14. [[Link](#)]
- Wen, S., Y.-Y. Wen, K.-E. Ching, Y.-L. Yeh, and Y.-H. Lee, 2019: Tectonic implications on the 2018 Hualien Earthquake. *Terr. Atmos. Ocean. Sci.*, **30**, 389-398, doi: 10.3319/TAO.2019.01.28.01. [[Link](#)]
- Withers, M., R. Aster, C. Young, J. Beiriger, M. Harris, S. Moore, and J. Trujillo, 1998: A comparison of select trigger algorithms for automated global seismic phase and event detection. *Bull. Seismol. Soc. Am.*, **88**, 95-106.
- Wu, Y.-M., C.-H. Chang, L. Zhao, J. B. H. Shyu, Y.-G. Chen, K. Sieh, and J.-P. Avouac, 2007: Seismic tomography of Taiwan: Improved constraints from a dense network of strong motion stations. *J. Geophys. Res.*, **112**, B08312, doi: 10.1029/2007JB004983. [[Link](#)]
- Wu, Y.-M., H. Mittal, T.-C. Huang, B. M. Yang, J.-C. Jan, and S. K. Chen, 2018: Performance of a low-cost Earthquake Early Warning System (*P*-Alert) and shake map production during the 2018 M_w 6.4 Hualien, Taiwan, Earthquake. *Seismol. Res. Lett.*, **90**, 19-29, doi: 10.1785/0220180170. [[Link](#)]
- Zhang, H., C. Thurber, and C. Rowe, 2003: Automatic *P*-wave arrival detection and picking with multi-scale wavelet analysis for single-component recordings. *Bull. Seismol. Soc. Am.*, **93**, 1904-1912, doi: 10.1785/0120020241. [[Link](#)]