

as smaller amounts of mono- and dihydroxydicarboxylic acids. The 2-methyltetrols were identified as major oxygenated organic compounds in the fine size fraction, and it is estimated that they represent a SOA source strength of about 2 Tg per year. These compounds have low vapor pressure and are hygroscopic; they can therefore contribute to particle growth (28), enhance the ability of aerosols to act as cloud condensation nuclei, and result in the formation of haze (29) above forests. The 2-methyltetrols can be regarded as specific molecular markers for the photooxidation of isoprene in the ambient atmosphere and are, as such, of potential interest for source apportionment and air quality modeling studies. Contrary to widespread assumption, we suggest that photooxidation of isoprene emitted by forest vegetation results in substantial SOA formation.

References and Notes

1. T. Novakov, J. E. Penner, *Nature* **365**, 823 (1993).
2. M. O. Andreae, P. J. Crutzen, *Science* **276**, 1052 (1997).
3. T. Hoffmann *et al.*, *J. Atmos. Chem.* **26**, 189 (1997).
4. T. Hoffmann, R. Bandur, U. Marggraf, M. Linscheid, *J. Geophys. Res.* **103**, 25569 (1998).
5. I. G. Kavouras, N. Mihalopoulos, E. G. Stephanou, *Nature* **395**, 683 (1998).
6. I. G. Kavouras, E. G. Stephanou, *Environ. Sci. Technol.* **36**, 5083 (2002).
7. A. Guenther *et al.*, *J. Geophys. Res.* **100**, 8873 (1995).
8. S. N. Pandis, S. E. Paulson, J. H. Seinfeld, R. C. Flagan, *Atmos. Environ.* **25A**, 997 (1991).
9. H. Geiger, I. Barnes, I. Bejan, T. Benter, M. Spittler, *Atmos. Environ.* **37**, 1503 (2003).
10. A. Limbeck, M. Kulmala, H. Puxbaum, *Geophys. Res. Lett.* **30**, 1996 10.1029/2003GL0177738 (2003).
11. Materials and methods are available as supporting material on Science Online.
12. J. Kesselmeier *et al.*, *Atmos. Environ.* **34**, 4063 (2000).
13. D. H. Lewis, D. C. Smith, *New Phytol.* **66**, 143 (1967).
14. A. L. Torres, H. Buchan, *J. Geophys. Res.* **93**, 1396 (1988).
15. R. Atkinson, W. P. L. Carter, *Chem. Rev.* **84**, 437 (1984).
16. L. Ruppert, K. H. Becker, *Atmos. Environ.* **34**, 1529 (2000).
17. J. H. Seinfeld, S. N. Pandis, *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change* (Wiley, New York, 1998), pp. 724–743.
18. S. Cohen, Y. Marcus, Y. Migron, S. Dikstein, A. Shafran, *J. Chem. Soc. Faraday Trans.* **89**, 3271 (1993).
19. J. Zhou, E. Swietlicki, H. C. Hansson, P. Artaxo, *J. Geophys. Res.* **107**, 8055 10.1029/2000JD00203 (2002).
20. G. C. Roberts, M. O. Andreae, J. Zhou, P. Artaxo, *Geophys. Res. Lett.* **28**, 2807 (2001).
21. B. Graham *et al.*, *J. Geophys. Res.* **107**, 8047 10.1029/2001JD000336 (2002).
22. K. Kawamura, R. Sempéré, Y. Imai, M. Hayashi, *J. Geophys. Res.* **101**, 18721 (1996).
23. B. R. T. Simoneit *et al.*, *Atmos. Environ.* **33**, 173 (1999).
24. Z. Zdráhal, J. Oliveira, R. Vermeylen, M. Claeys, W. Maenhaut, *Environ. Sci. Technol.* **36**, 747 (2002).
25. B. Graham *et al.*, *J. Geophys. Res.* **108**, 4766 10.1029/2003JD003990 (2003).
26. P. Saxena, L. M. Hildemann, *J. Atmos. Chem.* **24**, 57 (1996).
27. J. T. Houghton *et al.*, Eds. *Climate Change 2001: The Scientific Basis* (Cambridge Univ. Press, Cambridge, 2001).
28. W. R. Leitch *et al.*, *J. Geophys. Res.* **104**, 8095 (1999).
29. F. W. Went, *Nature* **187**, 641 (1960).
30. This work is a contribution to the Large Scale Bio-

sphere-Atmosphere Experiment in Amazonia (LBA). We thank the Belgian Federal Science Policy Office, the University of Antwerp, the "Fonds voor Wetenschappelijk Onderzoek-Vlaanderen," the European Commission, the Max Planck Gesellschaft, Germany, and the Fundação de Apoio a Pesquisa do Estado de São Paulo and Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil, for research support; the National Botanic Garden of Belgium for tree leaf specimens; and J. Kesselmeier and anonymous reviewers for constructive comments.

Supporting Online Material

www.sciencemag.org/cgi/content/full/303/5661/1173/DC1
Materials and Methods
SOM Text
Figs. S1 to S4
Schemes S1 and S2
Table S1
References

21 October 2003; accepted 21 January 2004

Repeating Seismic Events in China

David P. Schaff^{1*} and Paul G. Richards^{1,2}

About 10% of seismic events in and near China from 1985 to 2000 were repeating events not more than about 1 kilometer from each other. We cross-correlated seismograms from ~14,000 earthquakes and explosions and measured relative arrival times to ~0.01 second, enabling lateral location precision of about 100 to 300 meters. Such precision is important for seismic hazard studies, earthquake physics, and nuclear test ban verification. Recognition and measurement of repeating signals in archived data and the resulting improvement in location specificity quantifies the inaccuracy of current procedures for picking onset times and locating events.

Each day, hundreds of seismic events are located by the traditional method of picking onset times of seismic waves recorded by seismometer networks and then interpreting these data with a travel time model to infer the epicenter (latitude/longitude), depth, and origin time of each event. Pick errors (i.e., errors in picking onset times) and errors in the models lead to uncertainties in event location amounting to a few percent of the station spacing. These errors frustrate efforts to understand the interactions between neighboring events and to obtain more precise models of Earth structure. Model error may be reduced by locating many earthquakes simultaneously, and pick error may be reduced by measuring relative arrival times by cross-correlation when possible (instead of onset times). Reduction in both types of error can lead to improvement in location precision of up to three orders of magnitude in specialized studies of small regions, typically using short windows (lengths of a few seconds) based on the *P*- and *S*-wave onsets (*I*–8). Here, we report improvements in epicentral locations over a broad area of diffuse seismicity, namely China and surrounding regions, using waveform cross-correlation on a large scale.

Repeating earthquakes or doublets arise when two events display nearly identical seismograms at a common station, implying that the events have similar focal mechanisms and similar locations. Doublets have primarily been observed in the creeping zones of major faults (*9*–*12*) and more recently in subduction zones (*13*–*15*). Clusters of repeating events are often called multiplets.

The cross-correlation coefficient *CC* quantifies waveform similarity, with values from 0 (no similarity) to 1 (a perfect match). We processed 12 GB of waveform data for about 14,000 events in and surrounding China as listed in the Annual Bulletin of Chinese Earthquakes (ABCE) from January 1985 to April 2000 (*16*, *17*). The waveforms were acquired from several networks of stations that are archived by the IRIS Consortium (*18*).

We define a doublet to be an event pair having $CC \geq 0.8$ for time windows from 5 s before the *P*-wave to 40 s after the *Lg*-wave recorded at least at one station. An example of a pair of events in China (magnitude ~4.5) with similar waveforms, recorded at a station ~1100 km away, yields $CC = 0.8$ for a 220-s window bandpass-filtered from 0.5 to 5 Hz (Fig. 1). *Lg* is the largest amplitude phase and shows a high degree of correlation, even though it is a complex, scattered wave. The waveforms are similar even for unknown phases and coda as well as for the *P*- and *S*-waves.

An automated search isolated 1301 events that have $CC \geq 0.8$ with at least one other event (Fig. 2). These comprise 950 doublets, which can be grouped as 494 multiplets of various sizes (table S1). They are well distributed throughout seismic zones in and near China and represent more than 9% of the ABCE events from 1985 to 2000. Most events satisfy $CC \geq 0.8$ at only one station because of the sparse station coverage. As a general rule, high similarity of complex multiply scattered waves for an event pair requires the events to be no farther apart than about one-quarter of the dominant wavelength, which in our case is ~0.8 km (*19*, *20*). If the signal-to-noise ratio is good enough to see different phases overlaid, it further supports the hypothesis that the

¹Lamont-Doherty Earth Observatory, ²Department of Earth and Environmental Sciences, Columbia University, Palisades, NY 10964, USA.

*To whom correspondence should be addressed. E-mail: dschaff@ldeo.columbia.edu

events are closely located because the *S-P* and *Lg-P* times are nearly identical (Fig. 1).

For subsets of our 950 doublets we can give more quantitative analyses of the spatial separation. Seventy-five events (39 pairs) have at least three stations meeting the $CC \geq 0.8$ criterion, which is sufficient to solve for the relative epicenters (21). Epicentral locations where the azimuthal gap is sufficiently well constrained to provide robust locations (38 events, 20 pairs) confirm that these event pairs are separated by no more than 1 km, with location precision on the order of 100 m.

Evidence that the events (Fig. 2) are collocated at the 1-km level also comes from a detailed study of 28 events in the 1999 Xiuyan sequence (22) in which all events had subkilometer separations with at least one other event, constrained at the 100-m precision level by *Lg*-wave cross-correlation and double-difference location. Fifteen of those 28 events were found by the automated search reported here, in which the criteria are more selective. Therefore, not all of the events separated by less than 1 km can be retrieved by our current automated search procedure. Other factors that may cause an underestimate of the true number of repeating events separated by less than 1 km are *CC* detection thresholds and network variability over time (fig. S4) (17).

An independent study of underground nuclear explosions at the Lop Nor Test Site used waveform cross-correlation to relocate two explosions, one in 1992 and one in 1996, and found them to be separated by about 100 m, implying that they may be collocated and may have occurred within the same tunnel (23). These two events were also found in the 1301 events derived from our automated processing and are the only known explosions in this selected subset of events. None of the shaft explosions for Lop Nor were $CC \geq 0.8$ doublets, and all of those events were independently found to be separated by more than 1 km (23).

For the 293 events (178 doublets) recorded by only two stations, we are not able to perform a relative location, but we can put some bounds on separation distances given certain assumptions (figs. S1 and S2) (17). The data do not require the majority of these 293 events to be located more than 1 km from each other. Extrapolating to the 1301 events meeting $CC \geq 0.8$ at one or more stations and also applying the quarter-wavelength rule to single recording stations, we infer that almost all these events are separated by less than 1 km.

Events tend to cluster in space and in time. The 1301 events include 950 doublets having a wide range of recurrence intervals (Table 1). A large percentage, however, are causally related in that about 25% of the events occurred within a day of their corresponding doublet pair, and nearly 50% occurred within a week. For comparison, the partially creeping Calaveras Fault, California, has about 10% of the events occurring within 100 m and 1 day after their doublet pair (7). This does not necessarily indicate that events have repeated-

ly ruptured the same fault patch, because these fault scales are beyond our location precision. Examination of doublet pairs with recurrence intervals less than a week shows a slight tendency for the later event to have a lower magnitude (aftershock of the first), but the proportion is not significant given the uncertainties in the magnitude estimates. Most pairs have magnitudes that typically differ by not more than 0.5 units.

Doublets can be used to empirically assess location errors of published bulletins (Table 2). The 1301 events were matched to possible events in four existing bulletins. For an event to be counted as a match, the listed hypocenter must occur within 500 km of that listed in the ABCE and the origin time must not differ by more than 30 s. The number of matching doublet pairs (maximum 950) is listed in the fourth

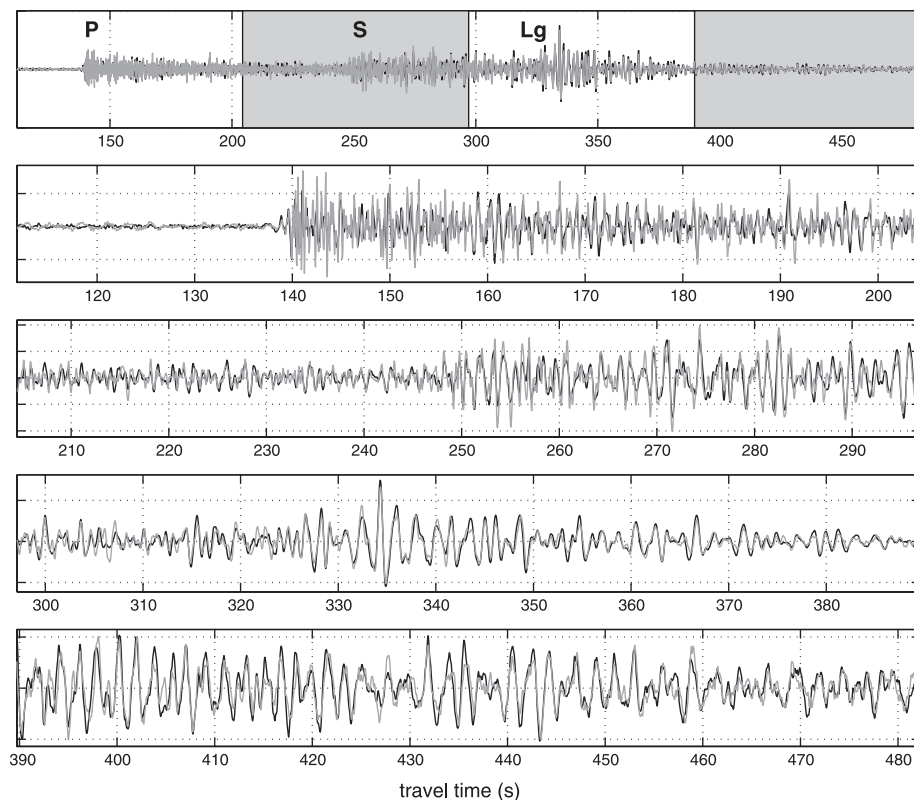


Fig. 1. A pair of similar events in China (black and gray waveforms) filtered from 0.5 to 5 Hz, with y axes normalized to unit amplitude. The topmost panel is a condensation of the data in the lower panels. The predicted *P*-wave arrives at 143 s, the *S*-wave arrives at 256 s, and the *Lg*-wave arrives at 315 s.

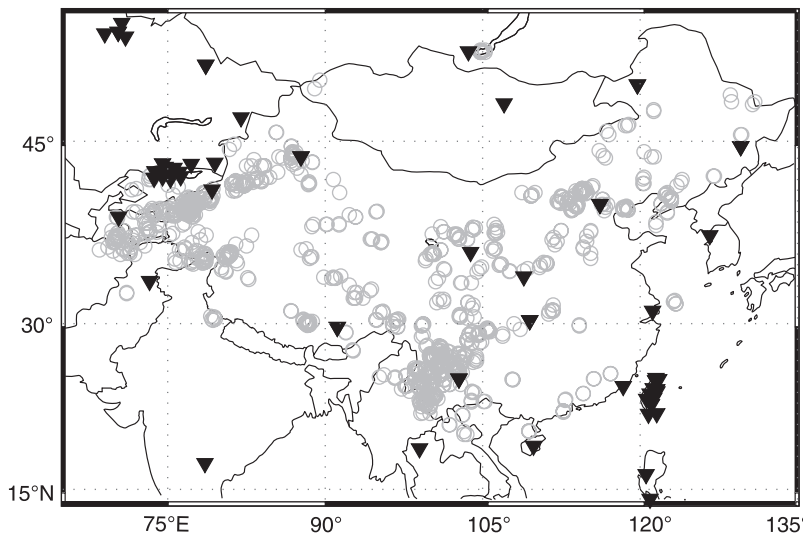


Fig. 2. 1301 events (9% of the ABCE), 950 doublets satisfying the criteria of $CC \geq 0.8$ for long windows from 5 s before the *P*-wave to 40 s after the *Lg*-wave on waveforms that are filtered from 0.5 to 5 Hz. Recording stations are denoted by solid triangles. Events (circles) are plotted at the ABCE absolute locations.

row for each bulletin. Our doublets are almost all separated by less than 1 km. The location difference computed from a published bulletin (between event pairs in a doublet) therefore indicates relative location error (24).

The median mislocation for the Reviewed Event Bulletin (REB) (25) is the largest at 31 km, as expected, because it emphasizes prompt reporting rather than the best locations. The International Seismological Centre (ISC) bulletin (26)—published 2 years in arrears, with more data collected and used for the locations—has correspondingly smaller median mislocations of 23 km. The culled ISC bulletin and associated locations by Engdahl, van der Hilst, and Buland (EHB) (27) produce the best locations on average.

Doublets can also be used to quantify pick measurement error in published bulletins. Because the events are collocated, they experience common but unknown perturbations to the travel times due to velocity heterogeneities. These variations can be canceled out by forming travel time differences. The resulting scatter reflects errors due only to uncertainty in the pick measurements. The relative pick error in the ABCE for *P*- and *S*-phases averages 2 and 4 s, respectively (fig. S3 and table S2). For *Lg*-phases pick error amounts to tens of seconds, whereas cross-correlation measurement error for *Lg*-waves ranges from 0.005 to 0.01 s—an improvement by better than three orders of magnitude over phase data (fig. S3 and table S2). Such excellent measurement precision results in relative location errors on the order of a few hundred meters (22).

The 1301 events were previously recorded and associated in the ABCE. An automated cross-

correlation detector could find all of the 1301 events and perhaps many more of some smaller magnitude that are not listed in the bulletin. The REB, which is designed for nuclear test ban treaty monitoring, reports even low-quality detections for all seismic events to avoid missing anything that could be an underground nuclear explosion. Of a possible 908 events, 423 were reported in the REB. The primary reason for the missed events is that the International Monitoring System was far from complete for this region during the time period of our study. The detection threshold for China and nearby regions is expected to be around magnitude 3.25 to 3.5 when all stations are operational (28). The ISC, with more reporting stations than the REB, misses 557 of the 1301 events.

By applying waveform cross-correlation to entire regional seismograms, 1301 repeating events with *M* ≥ 3 have been discovered in and near China from 1985 to 2000. These events are inferred to be collocated within 1 km. Although *Lg*-wave epicentral locations cannot be used to solve for depth separations, they are inferred to be small because of the waveform similarity (19, 20, 22). Because most of the recurrence intervals are a month or less, we infer that most of the repeating events in China are causally related, implying triggering by some mechanism. Theories to explain this type of earthquake interaction include static and/or dynamic stress transfer or pore fluid effects. Repeating events with longer recurrence intervals are usually assumed to occur as a result of creep loading of stuck patches or asperities. Location and pick errors can be independently and empirically assessed for traditional bulletins from repeat-

ing events. Cross-correlation can reduce pick error substantially (by factors of 10 to 10³). The entire regional seismogram including the complex *Lg*-wave correlates very well (*CC* ≥ 0.8) for these 1301 events across diverse tectonic settings in China for about 10% of the ABCE, providing compelling evidence for the existence of repeating earthquakes with similar waveforms on a broad scale. Such data may lead to improvements in event location and analysis of the physical properties of the source.

References and Notes

1. G. Poupinet, W. L. Ellsworth, J. Fréchet, *J. Geophys. Res.* **89**, 5719 (1984).
2. N. Deichmann, M. Garcia-Fernandez, *Geophys. J. Int.* **110**, 501 (1992).
3. P. M. Shearer, *J. Geophys. Res.* **102**, 8269 (1997).
4. A. M. Rubin, D. Gillard, J.-L. Got, *Nature* **400**, 635 (1999).
5. F. Waldhauser, W. L. Ellsworth, A. Cole, *Geophys. Res. Lett.* **26**, 3525 (1999).
6. C. Thurber, C. Trabant, F. Haslinger, R. Hartog, *Phys. Earth Planet. Inter.* **123**, 283 (2001).
7. D. P. Schaff, G. H. R. Bokelmann, G. C. Beroza, F. Waldhauser, W. L. Ellsworth, *J. Geophys. Res.* **107**, 2186 (2002).
8. C. A. Rowe et al., *Pure Appl. Geophys.* **159**, 563 (2002).
9. R. M. Nadeau, W. Foxall, T. V. McEvilly, *Science* **267**, 503 (1995).
10. J. E. Vidale, W. L. Ellsworth, A. Cole, C. Marone, *Nature* **368**, 624 (1994).
11. D. P. Schaff, G. C. Beroza, B. E. Shaw, *Geophys. Res. Lett.* **25**, 4549 (1998).
12. A. M. Rubin, *Bull. Seismol. Soc. Am.* **92**, 1647 (2002).
13. B. L. Isacks, L. R. Sykes, J. Oliver, *Bull. Seismol. Soc. Am.* **57**, 935 (1967).
14. D. A. Wiens, N. O. Snider, *Science* **293**, 1463 (2001).
15. T. Igarashi, T. Matsuzawa, T. Hasegawa, *J. Geophys. Res.* **108**, 2249 (2003).
16. W. H. K. Lee, H. Kanamori, P. C. Jennings, C. Kisslinger, *International Handbook of Earthquake and Engineering Seismology, Part B* (Academic Press, London, 2003).
17. See supporting data on Science Online.
18. Incorporated Research Institutions for Seismology (www.iris.edu/data/data.htm).
19. R. J. Geller, C. S. Mueller, *Geophys. Res. Lett.* **7**, 821 (1980).
20. The dominant wavelength (for *Lg* with period ~1 s) is ~3.3 km.
21. Because the *Lg*-wave is typically the largest amplitude arrival on the regional seismogram, it is assumed to dominate the cross-correlation differential arrival time measurements, and therefore the *Lg*-wave group velocity is chosen to estimate event separation distance.
22. D. P. Schaff, P. G. Richards, *Bull. Seismol. Soc. Am.*, in press.
23. M. D. Fisk, *Bull. Seismol. Soc. Am.* **92**, 2911 (2002).
24. If the variances of the two events are equal, absolute location error can be estimated by dividing the relative location error by √2 and adding a contribution due to systematic bias from effects such as model error.
25. The REB is now included in and can be obtained from the ISC bulletin.
26. The ISC bulletin is available at www.isc.ac.uk/Products/bulletin.html.
27. E. R. Engdahl, R. van der Hilst, R. Buland, *Bull. Seismol. Soc. Am.* **22**, 2317 (1998).
28. National Academy of Sciences, *Technical Issues Related to the Comprehensive Nuclear Test Ban Treaty* (National Academy Press, Washington, DC, 2002).
29. Supported by the Defense Threat Reduction Agency. This is Lamont-Doherty Earth Observatory Contribution number 6573.

Supporting Online Material

www.sciencemag.org/cgi/content/full/303/5661/1176/DC1
 Materials and Methods
 Figs. S1 to S4
 Tables S1 and S2
 Data files (catalog of 1301 events, differential times)

7 November 2003; accepted 23 January 2004

Table 1. Recurrence interval statistics for 950 doublets (1301 events).

	Interval						
	1 day or less	1 day to 1 week	1 week to 1 month	1 month to 1 year	1 to 2 years	2 to 5 years	5 to 10 years
Number	218	197	155	271	59	44	6
Percentage of doublets	23	21	16	29	6	5	1
Cumulative percentage	23	44	60	89	95	100	100

Table 2. Relative location errors for 950 doublets (1301 events) for several seismic bulletins (16, 25–27) given as median, mean, and maximum values. (The REB was published for a period during which only 908 of the 1301 events occurred.) The last three rows indicate the number of missing events for each magnitude if available, except for ABCE, which indicates the number of events for each (Richter) magnitude.

	REB	ISC	EHB	ABCE
Median (km)	31	23	12	15
Mean (km)	59	43	26	25
Maximum (km)	549	494	159	154
Number of doublets	226	458	79	950
Number of events	423	744	216	1301
Number of unreported events	485	557	1085	0
Percentage of unreported events	53	43	83	0
Number of missing Mag3 events	2	22	59	75
Number of missing Mag4 events	10	70	144	309
Number of missing Mag5 events	1	4	7	26