Software Development: Digitizing Historical Analog Seismograms for Wave Climate Analyses

Peter D. Bromirski
Integrative Oceanography Division
Scripps Institution of Oceanography
8650 Discovery Way
La Jolla, CA 92093-0209
phone: (858) 822-0552; fax: (858) 534-0300 email: *pbromirski@.ucsd.edu*

Award Number: N000140110403

FINAL REPORT

LONG-TERM GOALS

The severe winter storm cycle along the California coast during the 1997-98 El Niño-Southern Oscillation (ENSO) focused attention on the variability, and recent apparent intensification, of the wave climate along the West Coast. Few instrumental wave records exist before systematic buoy measurements began in the early 1980's. The number of large storm events reported by *Seymour* [7] for 1982-83 ENSO is conspicuously greater than earlier strong ENSO's during the 1940-41 and 1957-58 winters (produced from hindcasts using meteorological data). This suggests that either the effect of strong ENSO's on the wave climate of California has significantly changed, or the wave-climate record prior to 1980 is seriously deficient.

Accurate estimates of the West Coast wave climate can be obtained by inversion of the double-frequency microseism spectrum from broadband seismometer data (*Bromirski et al.* [4]). Analyses of NOAA buoy data indicate that the wave climate is very similar along much of the California coast (*Bromirski et al.* [4] and [5]), implying that reconstructed wave measurements from the San Francisco Bay area using seismic data from Berkeley, CA can be extrapolated to the Eel River coastal region, a study area for the ONR sponsored STRATAFORM program. Analog paper seismograms archived at UC Berkeley from 1930-1980 can be used to reconstruct the historical wave record.

Quantitative wave climate reconstructions from 1930 onward are important from two perspectives:

• *Shelf sediment transport and deposition:* The resolution of wave events, their spectral characteristics, and wave climate statistics over time are necessary to understand and model the processes controlling the formation and evolution of event-scale stratigraphy on the continental shelf. Since surface waves induce bottom oscillatory flow that is wave frequency dependent, knowing the spectral character of the wave climate over time is important to understand and model the resuspension, transport, and redisposition of sediments on the sea floor.

• *Global climate change*: Changes in the character and occurrence of extreme storms in the Northeast Pacific during strong ENSO's can be important. In this regard, wave reconstruction from Berkeley data will be absolutely unique in terms of continuity and stability of the measurement system, and will thus provide an unbiased record of extreme events in the Northeast Pacific. Reconstruction of the winter wave climate prior to 1980 will establish the variability of the wave climate and the magnitude of any intensification that has occurred since the mid-1970's.

OBJECTIVES

Quantitative estimates of wave energy and climate change from archived seismograms require digital data for spectral processing. The goal of this proposal was to modify the publicly available version of digitizing software (*NXScan V. 8.1*) and substantially improve digitizing efficiency, with the aim of digitizing winter months during the 1940-41 and 1957-58 El Niño's. *NxScan* was written for Sun Microsystem workstations during the 1980's and uses the obsolete, unsupported "pixrect" primitive graphics library written for the 1980's Sun architecture and operating system. Additionally, *NxScan* was not designed for the efficient digitization of day-long seismograms for climatological studies (e.g. *Bromirski et al.* [4]). Consequently, digitizing scanned images of seismograms using *NxScan* was extremely problematic. After considerable effort was spent attempting to improve and stabilize *NxScan*, incompatibilities of pixrect with current Sun architectures continued to cause repeated crashes, making the efficient digitization of month-long time series impossible. Thus, *SeisDig* was developed using Matlab (see *Bromirski and Chuang* [1]), which makes *SeisDig* essentially computer and operating system independent. The hope is that *SeisDig* can be relatively easily maintained as Matlab and computer systems evolve.

APPROACH

Scanned seismograms for portions of the 1940-41 and 1957-58 time periods were obtained from the BSL archives. These scanned images are stored as "tif" images, which are imported directly into *SeisDig* for digitizing. *SeisDig* was developed and tested primarily on these images, but also on test seismograms made from recent broadband digital data from the Berkeley, CA (BKS) seismic station. Matlab coding was done by UCSD undergraduate computer engineering major Steve Chuang.

Examples and illustrations in this report were made from Galitzin seismograms from the Berkeley Seismological Laboratory (BSL) archives that were electronically "scanned" at 400 dpi (dots/inch) using an IDEAL FSS8000 full length scanner and stored as "tif" files. Each scanned seismogram image consists of a series of quasi-parallel sequential 30-minute or 60-minute traces that generally cover about a day, with two partial traces at the beginning and end of each image. One-minute timing marks usually cause small gaps at the end of each minute. In addition to high amplitude earthquake signals, the arrival of high amplitude storm waves at nearby shorelines (*Bromirski and Duennebier* [3]) causes elevated noise levels called "microseisms", both types of signals resulting in the occurrence of adjacent trace intersections, overlaps, and reduced line thickness that complicate digitization.

Storage problems over time have damaged some seismograms (particularly at edges), further complicating digitization of full seismograms. Degradation of the seismogram paper quality (especially for many pre-1950 seismograms), and the difficulty in perfectly aligning the 40 inch long seismograms in the scanner feed groove, commonly result in either slightly skewed or slightly curved trace images, typically exaggerated at the trace ends. These scanning artifacts become apparent if the horizontal:vertical image ratio is augmented, but are difficult to identify in the raw tif images. The effects of these scanning artifacts were minimized by fitting a third order polynomial to the image traces to establish a reference line for each trace to determine the digitized trace amplitudes.

Because of presumably varying voltage, Galitzin seismograph drum speed variability resulted in the Time Scale Factor (TSF) varying from about 450 to 480 pixels/min. Impact of the variable TSF is minimized by using an average pixels/min determined for the entire seismogram. When digitizing relatively short duration earthquake signals, estimates of the TSF should be made either just prior to the event or from an average of the TSF for traces before and after the event.

WORK COMPLETED

SeisDig is a Graphical User Interface (GUI) based interactive digitizing tool. Digitizing operations are mouse selected. Because digitizing numerous entire seismogram images efficiently is critical for studying the characteristics of extreme storms and their occurrence for climate change analyses, automated trace image tracking is necessary. Efficient digitizing necessitates a minimum of user interaction.

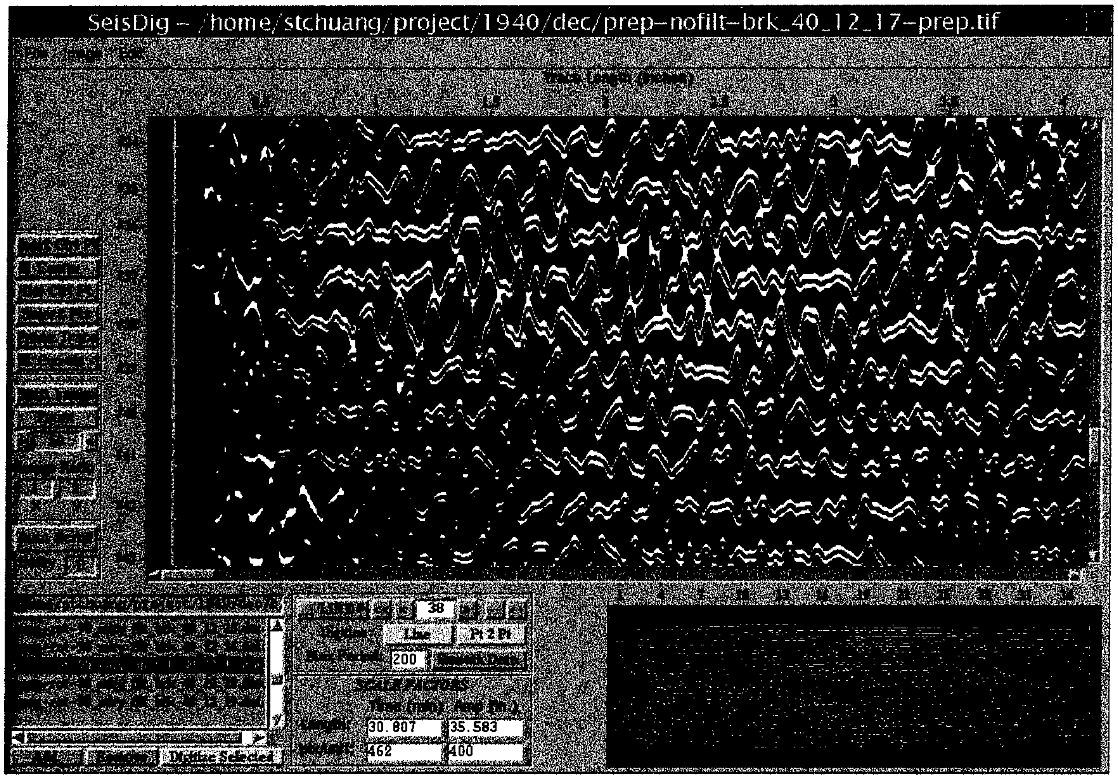

Figure 1: A portion of the seismogram for Dec. 17, 1940 viewed with *SeisDig* shows the trace variation observed in pre-1962 Galitzin seismograms, with an example of faded trace lines at the seismogram edge that sometimes occurs in older seismograms. The scanned traces (thick white lines) show examples of adjacent trace intersections during relatively high amplitude noise levels. The projected digitized traces (blue lines) successfully track past these irregularities. The magenta trace is the current (active) trace being modified. The *Edge Line* (vertical green line) is shown on the left, its position adjusted to conserve the total elapsed time of the seismogram. The full seismogram image is shown in the locator panel at the lower right, with the portion of the image shown in the main image panel outlined by the red box and the active trace identified by the dashed magenta line. The right Edge Line is set at the well defined trace ends. The paper seismogram physical trace length in inches (determined using the pix/min value displayed) is shown at the top of both the main *SeisDig* and the locator panels.

Timing marks, occuring at 1-minute intervals in most Galitzin seismograms, cause small gaps in the seismogram image, and high amplitude signals result in trace intersections and overlaps (see Fig. 1). The trace width can vary within a trace as well as from trace to trace. Each of these, as well as other gaps and thin or intermittent lines, can cause incorrect trace-tracking. The necessity of setting control points at gaps and intersections was substantially eliminated with a trace-tracking algorithm that follows the center of the trace image (in the vertical cross-section) and uses a combination of slope estimates from both the preceding digitized points and the trace amplitude range to constrain the forward digitizing search. Trace tracking is further constrained to recross the baseline within a user selected time period. Incorrect handling of trace intersections and other image irregularities (commonly occurring when high amplitude signals cause repeated trace intersections or overlaps) are easily identified by visual inspection of the projected digitized trace overlaying the raw image trace. The GUI allows interactive modification of the digitized trace(s) (as well as setting control points at trace ends for pre-processing) prior to saving to an output file for cases when image quality and/or complexity warrant user interaction.

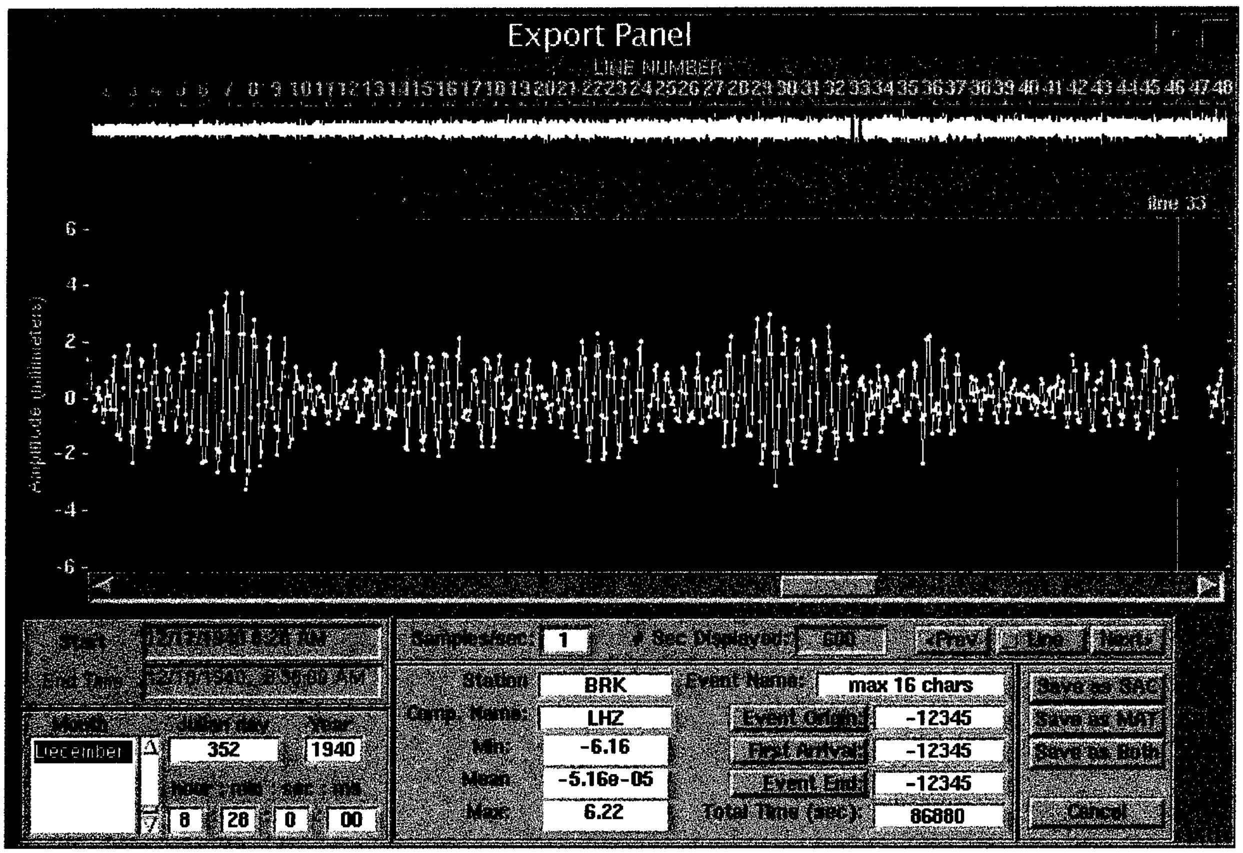

Figure 2: Layout of the "*Export*" pop-up window. The top panel shows the concatenated output of the digitized data for the full seismogram shown in Fig. 1, containing 48 traces of approximately 30 min duration each. The portion of the digitized trace outlined by the red locator box (600 s) in the upper panel is expanded in the main panel for detailed inspection. The yellow dots indicate the sampled amplitudes to be exported. Note the semi-regular coda-like arrivals, indicative of microseisms generated by the arrival of ocean swell at nearby shorelines.

Digitized data can be exported in either the common Seismic Analysis Code (SAC) format or as a Matlab *.mat file that mimics the SAC2000 Matlab output. SAC2000 can be downloaded from *http://www.llnl.gov/sac/*. Figure 2 shows the *SeisDig* "Export" window for previewing the digitized output and setting header and digitizing parameters. The output digitized trace amplitudes are in millimeters and correspond to trace amplitudes of the paper seismogram.

RESULTS

The temporal variation of raw spectral levels for digitized data for Jan. 1941 (Fig. 3) shows energy concentrations on Jan. 8-9, 19-20, 22-23, and 24-25 that are characteristic of double-frequency microseisms (high amplitude energy in the [0.09,0.25] Hz band at twice the wave frequency). These signals result from wave-wave interactions at nearby coastal locations at the arrival of dispersed gravity waves (swell) from large storms, with the spectral variation in Fig. 3 similar to that observed in modern digitally-recorded seismic data [see e.g. *Bromirski and Duennebier* [3]; *Bromirski et al.* [4]].

Associated primary microseisms at the frequency of the waves are not clearly identifiable in the [0.05,0.09] Hz band, partially because these spectra do not include an instrument response correction (not yet available). Relatively short-duration high-amplitude signals at frequencies less than 0.09 Hz are earthquake surface wave arrivals. The Galitzin seismometer that recorded these signals has a natural period of 12 sec, causing the peak in the raw spectra near 0.1 Hz to be elevated relative to 0.06 Hz and 0.25 Hz levels. However, the character of the spectral variation observed and the percentage of the data digitized indicate that these data can be inverted to obtain gravity-wave parameter estimates at a more than adequate sampling density to enable quantitative characterization of the wave climate prior to 1980 when instrumental buoy measurements are unavailable.

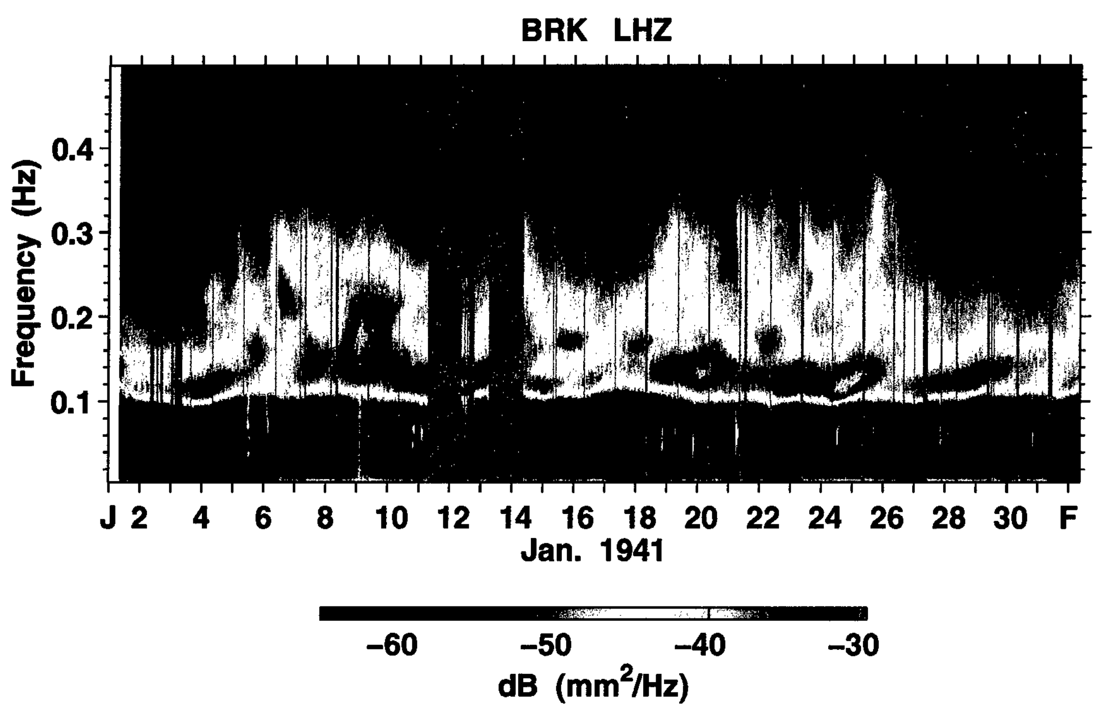

Figure 3: Power spectral variation of digitized seismograms during Jan. 1941, not corrected for either instrument response or calibration constants. Vertical stripes indicate either missing data or time periods where the scanned image quality was not adequate for digitizing.

After determining the transfer function to convert the digitized data from millimeters to microns, inversion of the microseism signals in these data (following the methodology of *Bromirski et al.* [4]) will give estimates of wave climate parameters during these earlier winters to be compared with the buoy-measured wave climate during the 1982-83 and 1997-98 El Niño's. Results from this demonstration study will show the feasibility of the seismic analog-to-digital-to-wave climate methodology. The next step will be to scan and digitize the remaining analog seismograms from 1930 onward so that a complete wave record can be constructed.

The long-period Wilip-Galitzin seismometer at Berkeley (that produced the seismogram in Fig. 3) was replaced with a long-period Sprengnether seismometer in 1962, with subsequent seismograms having a thinner pen width that makes trace-image tracking more difficult. Further software development is required to efficiently digitize post-1962 seismograms and to automate digitizing across trace overlaps. Digitizing scanned pre-1962 seismogram images with *NXScan V. 8.1* required about 15 min (or more, depending on image quality) per 30 min trace. *SeisDig* requires 30-45 min per day-long seismogram (48 traces in pre-1962 seismograms), a very significant improvement.

IMPACT / APPLICATIONS

Studies modeling nearshore sediment transport over time require good estimates of wave climate statistics. Extending the wave record backwards to 1930 along the California coast will provide a substantially longer time series for sediment transport modeling efforts than currently exists.

Reconstruction of the coastal wave record is important to the State of California. Analyses of the wave climate over the past decades (using meteorologically-based reconstructions and buoy data) show major upward trends in large, long-period wave episodes, especially since the mid-1970's (*Seymour* [7]). Because these trends obviously have important impact on design criteria for coastal construction, beach nourishment and evaluations of coastal erosion, wave climate data from earlier decades will be of significant economic value from both "design-wave" and insurance liability perspectives to establish whether the observed variation in recent decades is in fact a long-term trend or cyclicity at time scales longer than the currently available wave record.

RELATED PROJECTS

Tide gauge data from 1858 onward from San Francisco Bay is currently being studied with funding from the California Department of Boating and Waterways. The variance of non-tide residuals shows "storminess" cycles on time scales of 10-20 yr (*Bromirski et al.* [2]), consistent with other climatological studies (e.g. *Trenberth and Hurrell* [8]; *Graham* [6]). Extending the wave record backward to 1930 from inversion of microseism data would help characterize the wave climate variation observed during the last 25 years, i.e. have the storm and wave intensity increased in the most recent decades? Wave climate data showing decadal-scale cycles in long-period wave energy would validate results from tide gauge data and other climatological studies, and provide valuable information for coastal planning.

References

[1] Bromirski, P.D. and S. Chuang (2003). SeisDig: Software to digitize scanned analog seismogram images, User's Manual, *Scripps Institution of Oceanography Technical Report*, http://repositories.cdlib.org/sio/techreport/ .

[2] Bromirski, P.D., R.E. Flick, and D.R. Cayan (2003). Storminess variability along the California coast: 1858-2000, *J. Clim.*, **16**, 982-993.

[3] Bromirski, P.D. and F.K. Duennebier (2002). The near-coastal microseism spectrum: Spatial and temporal wave climate relationships, *J. Geophys. Res.*, **107**(B8), 2166, doi:10.1029/2001JB000265.

[4] Bromirski, P.D., R.E. Flick, and N. Graham (1999a). Ocean wave height from inland seismometer data: Implications for investigating wave climate changes in the NE Pacific, *J. Geophys. Res. C*, **104**, 20,753-20,766.

[5] Bromirski, P.D., R.E. Flick, and D.R. Cayan (1999b). The northeast Pacific wave climate: Spatial and temporal variation, EOS: Trans. Am. Geophys. Union, 2001 Fall Meeting, abs.

[6] Graham, N.E. (1994). Decadal-scale climate variability in the tropical and North Pacific during the 1970's and 1980's: observations and model results, Clim. Dyn., **10** 135-162.

[7] Seymour, R.J. (1996). Wave climate variability in southern California, J. Waterw. Port Coast. Ocean Eng., 182-186.

[8] Trenberth, K.E. and J.W. Hurrell (1994). Decadal atmosphere-ocean variations in the Pacific, Clim. Dyn., **9** 303-319.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (FROM - TO)
20-06-2003	Final Technical Report	01-02-2002 - 31-03-2003

4. TITLE AND SUBTITLE
Software Development: Digitizing Historical Analog Seismograms for Wave Climate Analyses

5a. CONTRACT NUMBER

5b. GRANT NUMBER
N00014-01-1-0403

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5F. WORK UNIT NUMBER

6. AUTHOR(S)
Peter D. Bromirski

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)
Scripps Institution of Oceanography
Integrative Oceanography Division
9500 Gilman Drive
La Jolla, CA 92093-0218

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
Attn: Dr. Jill L. Karsten, Code 322GG
800 North Quincy Street
Arlington, VA 22217

10. SPONSOR/MONITOR'S ACRONYM(S)
ONR

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

12. DISTRIBUTION AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE

13. SUPPLEMENTARY NOTES

14. ABSTRACT
North Pacific storm systems strongly affect the California coast, with long period waves representing the single largest source of energy driving coastal processes. The understanding and modeling of nearshore sediment transport over time requires a long record of wave statistics. The instrumental buoy wave record along the West Coast begins in the mid-1970's. The wave record along the California coast from 1930 to the beginning of the instumental record can be obtained from analog long-period seismograms archived at UC Berkeley. Quantitative estimates of wave energy and climate change from archived seismograms require digital data for spectral processing. New software (SeisDig) has been developed for the efficient digitization of day-long seismograms for climatological studies. SeisDig is a Matlab based graphical user interface interactive digitizing package that produces digitized data from electronically scanned "tif" images for wave climate reconstructions in Seismic Analysis Code (SAC2000) format.

15. SUBJECT TERMS
digitizing analog seismograms, wave climate reconstruction

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON / 19b. TELEPHONE NUMBER (Include area code)
Unrestricted	Unrestricted	Unrestricted	None	8	Peter D. Bromirski / 858-822-0552