Journal of Scientific Exploration, Vol. 4, No. 2, pp. 137–152, 1990
Pergamon Press plc. Printed in the USA
0892-3310/90 $3.00+.00

Some Observations of Electromagnetic Signals Prior to California Earthquakes*

MARSHA HANCOCK ADAMS

Time Research Institute, P.O. Box 620198, Woodside, CA 94062

Abstract — Electromagnetic (EM) signals in the frequency range below 1,000 Hz have been monitored since 1981 for the purpose of earthquake forecasting. Signal strength increased more than 7 standard deviations above the mean prior to 3 major California Earthquakes; Coalinga (1983), Whittier Narrows (1987), and Lorna Prieta (1989). The signal increases occurred 10 days to one month prior to the earthquakes. They were continuously elevated until after each earthquake occurred.

An effort to forecast time and location of smaller earthquakes in the magnitude 2–4 range is underway. Expert system software has been developed to interpret the EM signals in near real time. The expert system makes forecasts on a daily basis for selected areas in California. A preliminary statistical analysis of recent forecasts appears promising, yielding probabilities of p $1*10^{-4}$ or better.

On August 7, 1990 another series of strong signals began. They have continued for an unprecedented length of time and are still present at the time of submission of this paper on October 16, 1990.

This paper is a summary of research in progress, on a prototype earthquake forecasting system based on EM signals below 1 KHZ. Historically, the hypotheses that electro-magnetic EM (/n) signals may precede earthquakes came not from seismology, materials sciences or physics, but rather from the perspective of biology. While doing research at Stanford Medical School in the early 1970s, I was struck by the unexplained variability that occurs in biological experiments and wondered about its cause. I found that biological experiments sometimes showed aberrations prior to seismic activity and began searching for a variable that was capable of causing biological effects, and that might possibly precede seismic activity. After much research, I came to hypothesize that EM radiation below 1 khz might be associated with these effects.

The evidence of a link between EM, seismic and biological processes appeared so compelling that I decided to monitor EM initially without funding. In 1981, I contacted Tony Fraser-Smith and he kindly provided technical knowledge and advice, and loaned me a coil antenna for awhile. We saw the first signals on a strip chart recorder in Los Altos, California in 1981 when the first crude monitoring station was implemented.

* Paper presented at the Ninth Annual Meeting of the Society for Scientific Exploration, August 9–11, 1990, Stanford University, Stanford, California.

Over the last nine years since these observations began, we have developed and tested a series of antennae or sensors, on an ongoing basis, to best detect the EM signals. The detector system has been refined and the data acquisition process is now computerized rather than taken on a strip chart recorder.

In the early 1980s, related investigations were proposed twice to the United States Geological Survey (USGS). They declined to participate in its sponsorship, and they were not encouraging in later verbal contacts. Because funding was not forthcoming from the government, we turned to private sources. Because this project has been privately sponsored, at this time the exact methodology used and level of detail I can present at this presentation is temporarily restricted. This includes the configuration and location of the sensors, and specific frequencies monitored. My ultimate goal is full disclosure of these details to the scientific community, but until more results are known about system performance this information is temporarily proprietary. This may invite skepticism, but keep in mind that the ultimate validation of theory and methodology is the evaluation of performance of a test system in real time.

Rather than focus on methodology, I will focus on function, performance, and results of the system. I will show you functional diagrams of the system, and data prior to the 1989 Loma Prieta earthquake to compare to other observations, namely Tony Fraser-Smith's data, which he presented earlier at this meeting. In addition, this system has had stations in operation in California for a long time. During that time, some other interesting seismic activity has occurred. We will explore the data for those time periods. Finally we will examine the performance of the system by statistical assessment. Statistical methods will be presented in detail.

Instrumentation

Figure 1 shows a diagram of the system. A sensor system feeds analog data into an A/D converter then the digitized data is processed on an IBM PC and stored on floppy diskettes. The sampling is done continuously. The system runs 24 hours a day, and has since 1982 with a hiatus of a few months. It is protected by means of an uninterruptable power supply (UPS) and back-up generator. Except for the few times that there has been malfunction in the sensor itself, the system is very reliable and the data stream is continuous.

Other variables are known to introduce aberrations in EM measurements below 1 KHZ. To explain as much nonseismic variability as possible, several factors must be considered. The site location must be chosen carefully. It must be in an electromagnetically and vibrationally quiet area (e.g., Sites under power lines with heavy traffic are not good). Two other variables that must be considered are weather and the consequences of solar activity. For instance, last year, a four-fold change in the signal values could easily have been misinterpreted to signify seismic activity. This was caused by a solar flare. On August 16, 1989, two months prior to the Loma Prieta earthquake,

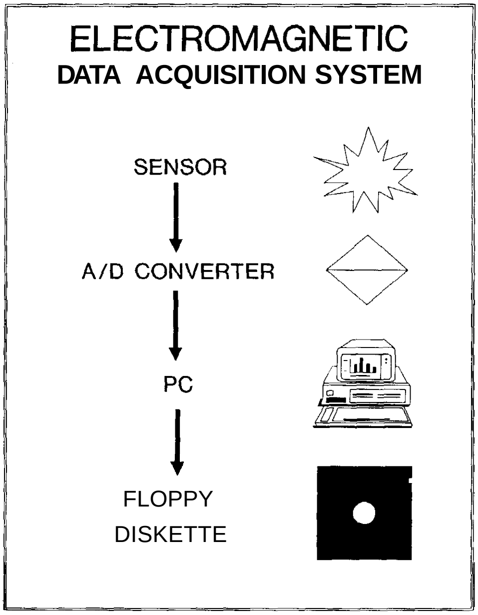

Fig. 1. Data is acquired from the sensor, digitized and processed on a PC, then stored on a floppy diskette.

an X20 solar flare occurred. It was the largest X-ray output ever recorded from a flare and certainly caused perturbations in this system. These perturbations could have been misinterpreted as indicating earthquake activity.

To account for as much variability as possible, two other systems are run concurrently with the EM monitor; a computerized weather station is in continuous operation. It measures temperature, wind speed and direction, precipitation, and barometric pressure. To monitor solar activity, another computer is attached to the feed from the space environment laboratory satellite broadcast system. This downlink is a real-time system that provides data on solar activity and the terrestrial and space environments. These data are accumulated on floppy diskettes. Figure 2 shows a diagram of the complete EM, Weather, Solar Data acquisition system.

Data for Large California Earthquakes

There is an additional expert system component which we will explore later. But, at this point let's look at the output of the EM system prior to the Lorna Prieta earthquake. Figure **3** shows daily averaged data for the time period prior to and after the Lorna Prieta earthquake. It begins April 1, 1989 and Ends November 17th. The Y axis represents output units of the A/D converter. The horizontal lines are the mean and 2 standard deviation levels. The mean is the signal mean for the 2 years prior to the rise. The upper line

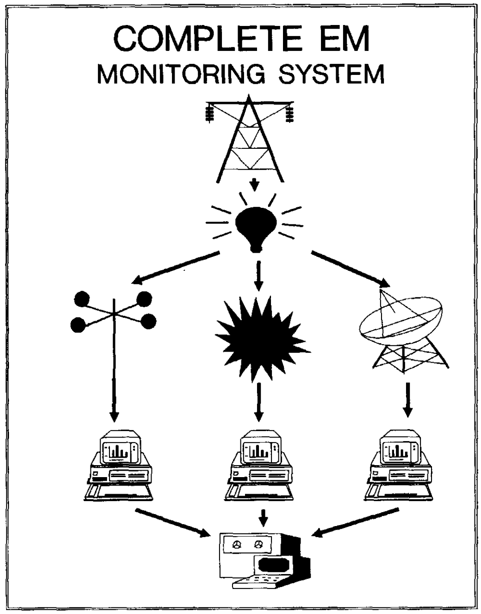

Fig. 2. The complete monitoring system is designed to remain in operation. Power goes through an uninterruptable power supply and is distributed to a weather station, the sensor station, and to a solar-terrestrialdownlink. PCs process and record all data continuously, 24 hours a day.

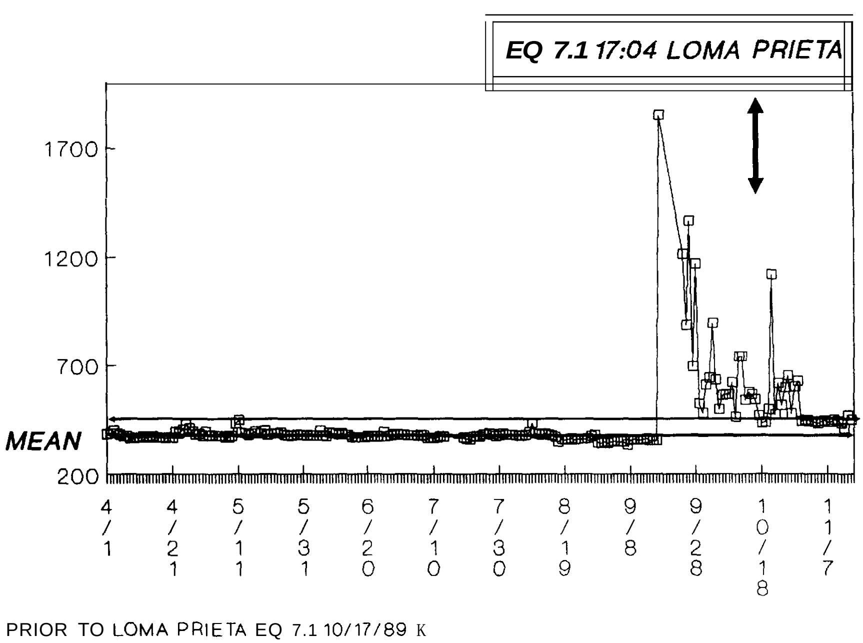

Fig. 3. Electromagneticsignals prior to the October 17, 1989 Lorna Prieta earthquake.

represents the 2 standard deviations above the mean. The amplitude of the highest peak was well over 34 sigma. Note that the exact highest part of the peak is missing from the data.

More than a month before the Loma Prieta main shock, a few days prior to September 16, 1989, the signal strength began to rise. Then, on the 16th it suddenly shot up to the highest levels recorded since the system was in put operation. Before the day was out, the sensor was damaged in a storm. It was replaced with a similar sensor within hours. This sensor recorded even higher levels. In about four days the damaged sensor was repaired and placed back into service where it continued to record exceedingly high signal levels. For the sake of continuity, data from the other sensor was not included in this figure. The elevated signals continued to the time of the earthquake and remained afterwards for a few weeks. At the time I was perplexed. My assumption was that the recently repaired sensor had not been repaired correctly and was giving me erroneous data. The engineer who works on the system had gone to the east coast on a business trip and would not be back for three weeks. I thought the sensor was still malfunctioning and that I had to wait until the return of the engineer for further repairs.

Shortly before October 17th, I was surprised to see the signals drop to near normal levels. As I had observed this phenomenon prior to many other earthquakes, only then did I realize that the sensor had not been malfunctioning, and that an earthquake was impending but I was still skeptical. Although the earth shook violently, we were lucky that there was only a brief power outage, and the UPS had protected the system from going down even through the earthquake. Immediately after the earthquake, I rushed to look at the sensor data screens. They were fluctuating wildly. It was interesting to see that one of the sensors, which may be motion sensitive, showed bursts of increased activity 5 to 15 seconds before the aftershocks were felt.

I synchronized my watch with WWV and noted the times of dozens of these bursts. Sometimes I could feel the aftershock, sometimes not. At a later date I will compare these notes to the actual times of the aftershocks to try to determine whether they were due to actual EM signals occurring at the time of the aftershock or due to the mechanical vibration of the sensor at the time of the P-wave arrival. The P-wave of the earthquake may be too small to be felt, it precedes the S-wave which is usually the perceptible part of the earthquake motion.

During the aftershock period, the feed from the sensors were split off the A/D converters to various monitoring equipment. Lacking more expensive recording devices, about four hours of videotape were recorded of the signal activity from the sensors on dual channel displays during the few days after the main shock.

EM observations prior to earthquakes can raise as many questions as they answer. The question has been posed to Tony Fraser-Smith that the signal he recorded prior to the earthquake may have been a coincidence; that it needs corroboration. The signal Tony recorded was concurrent with the signal recorded by this system. Although we use different sensor technology, the sensors were at different locations, and incoming data is processed differently, nonetheless, two recordings of exceedingly high EM signals just prior to a large earthquake appears not to be coincidental.

In further support of the hypothesis that strong EM signals precede large earthquakes, lets look at the signals prior to the two other large California earthquakes that have taken place since the system has been in operation; the Whittier (1987) and Coalinga (1983) earthquakes. Figure 4 shows that about 10 days prior to the Whittier earthquake a very high spike occurred in the data, then it resumed normal levels. Several days prior to the earthquake it increased suddenly again to what was then record levels for the system. This data was so aberrant that I also believed at that time that the system had malfunctioned.

Note the similarity of the signals before Loma Prieta and Whittier Narrows earthquakes. One interesting thing to note here is that the signal strength remained very high for a few weeks after the Whittier Narrows quake and then returned to normal. Although the signal level has decreased markedly now, after the Loma Prieta quake, it has still not quite returned to pre-earthquake levels.

Figure 5 shows the first large California Earthquake before which elevated EM signals were observed by this system. It was the 1983 Coalinga earthquake. This earthquake occurred very early in the EM observation effort using the first prototype sensor. Data then was recorded on a half-daily basis.

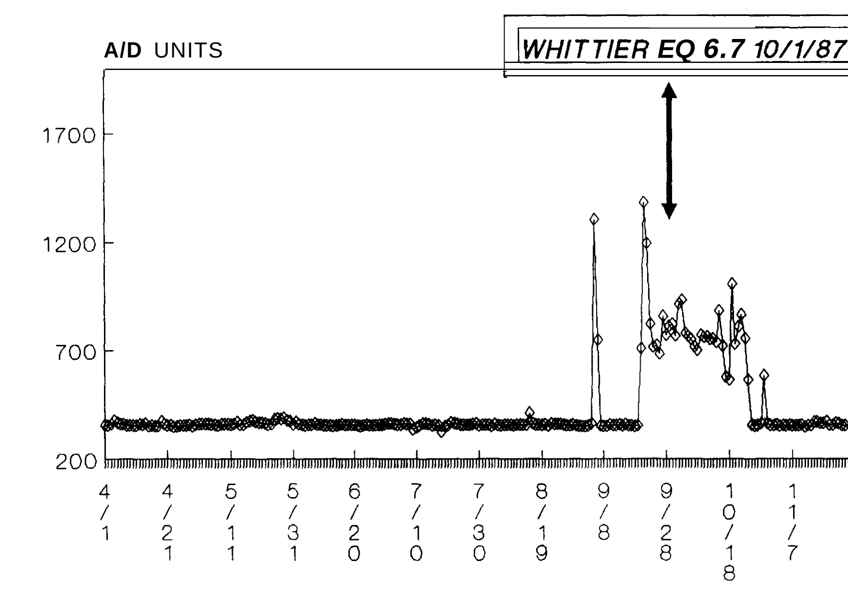

Fig. 4. Electromagnetic signals prior to the October 1, 1987 Whittier earthquake.

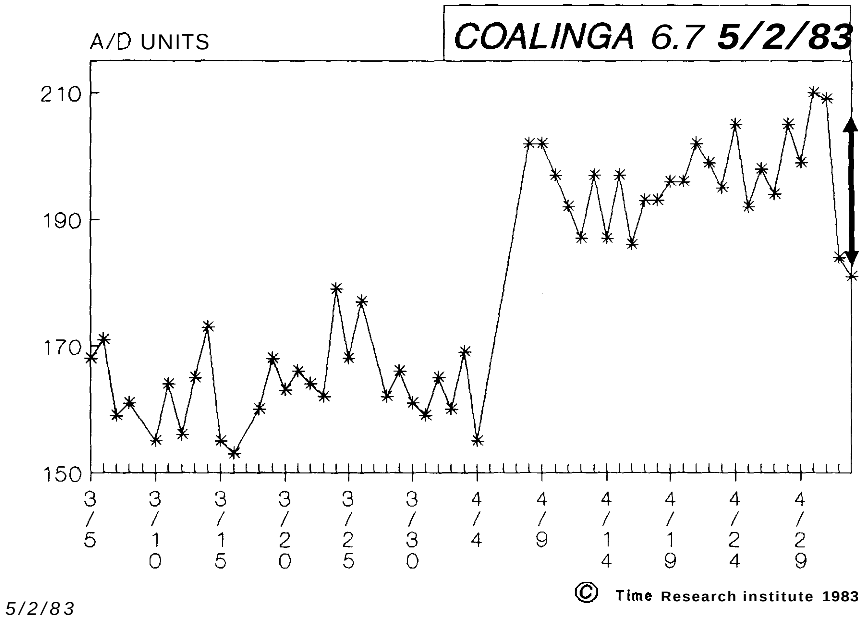

Fig. 5. Electromagnetic signals from the prototype station before the May 5 1983 Coalinga earthquake.

Since that time, sensors and data acquisition procedures have been markedly refined. Though an increase in amplitude is clearly visible, the quality of the data contrasts with data acquired more recently. Nonetheless, the largest increase seen to that date occurred within a two-week period prior to the Coalinga earthquake.

It does not appear that observation of exceedingly large EM signals followed by large earthquakes is mere coincidence, especially now that these observations have been corroborated by independent measurements. The question that remains is: How useful is observation of EM signals for the purposes of practical earthquake forecasting? Do they only precede large earthquakes? Can EM signals be detected before smaller, but potentially damaging earthquakes, if so, what are the magnitude limitations?

Forecasting Smaller California Earthquakes

These questions introduce us to the rest of the system I promised to describe earlier. Its purpose is to process the data through an expert system that interprets various inputs, resulting in experimental forecasts for earthquakes in the magnitude 2–4 range. This magnitude range was selected because there is more data to work with due to the frequency of occurrence of earthquakes in this range.

The incoming data are input to an IBM PS/2 model 80 or a Wyse 25 Mhz 386 computer where it is processed by expert system software that we have developed. The output is an experimental forecast that identifies a start date, and location for potential earthquakes. All forecasts are for a standard magnitude range of 2.0–4.0. To assess the feasibility and performance of the expert system, the forecasts of are output to both a paper document and a computer data file.

Many people mistakenly ask what the percent success rate of the forecasts is. The most important question to ask is not the percent success rate of the forecasts, but are the forecasts better than chance, and if so what percent better than chance are they? As you all know earthquakes are extremely common in California. Approximately three earthquake's of magnitude 1 or less, and 1.5 earthquake's in the magnitude 2, range occur in Northern and Central California each day. About five magnitude 3's occur within a 1,000-mile radius of San Francisco each week. Because of the high probability that earthquakes will occur in a given area, it is necessary to calculate the success rate of actual forecasts then compare the success rate to the chance success rate for that area. Depending on the location and time frame allowed, chance can range from a few percent to 100%, so it is necessary to restrict the time frame and radius of the forecasts to insure that they do not have a high probability of being successful by chance alone.

To determine the difference between actual forecast success rates and chance success rates and the corresponding statistical probability, a Monte Carlo procedure was used to analyze the test forecast output. The Monte Carlo analysis software was developed at Time Research Institute (TRI). The assessment procedure is as follows:

First, a time interval, for example June first to July first, is selected. Every forecast made for a given location generated by the expert system is included in the test data set. A number of parameters are defined that qualify earthquakes for a hit for a forecast. The parameters include the time interval allowed for the forecast to come to fruition, the maximum radius from the center of the test location, and the minimum acceptable magnitude for an earthquake.

These criteria are applied to each forecast of the set of forecasts. The test data set is then matched with earthquakes that meet these criteria. A success rate, expressed as a per cent is calculated. A success is defined as at least 1 earthquake occurring during a forecast period that meets the selected criteria. No added consideration is given for multiple hits.

Next, the same number of random dates are generated as there were real forecasts in the data set. They are generated to fall within the exact time interval as the actual forecast data, for example, five random dates falling between June 1 and July 1. Duplicate dates are omitted. Earthquake occurrence is compared to the random dates using exactly the same selection criteria as for the actual forecasts. Usually about 100 sets of random dates are generated and compared to earthquake incidence to determine the "chance success rate." This is also expressed as a percentage.

Other relevant considerations are: The distribution of the pseudorandom date generator appears flat, and the distribution of the percent success values appears normal. A z-statistic is used to determine the probability for the difference between actual forecasts and chance.

The greatest challenge to accurate assessment is to locate a database that is truly descriptive of the seismic activity of an area. Historical data sets are accurate, but do not cover the time periods of EM monitoring. More recent databases are preliminary and contain known errors in date, magnitude, and location of the earthquakes. Other recent databases are selective; including only seismic events that are large, or that are felt in populated areas.

Two databases are in-house for the Monte Carlo analysis. The analysis process is in its preliminary stages. The bulk of the analysis has been done on one database, the other, has only been available for about two weeks. The first data set is a compilation of data from all available near real-time sources. It includes all Northern and Central California earthquakes from the following sources; earthquakes downloaded from the USGS computer in Golden Colorado, earthquakes from the University of California Berkeley information line, earthquakes from the USGS Menlo Park information line, and earthquakes extracted from the weekly Preliminary Determination of Epicenters. Duplicate earthquakes from two or more sources are omitted. No other selection criteria or screening of the data has been applied to this database.

The shortcoming of this database is that there is no absolute lower magnitude cutoffs for the earthquakes, and the many lower magnitude earthquakes in the magnitude 2 range and below are missing or not represented accurately. Nonetheless, because no selection criteria were imposed on these data, one would expect this database to be equally biased towards good and bad results for both actual and random forecasts in the Monte Carlo analysis.

The second database, which has only been available a short while, was obtained from the USGS Menlo Park with the caveat that it is preliminary and is known to contain errors. This database, however, has better coverage of lower magnitude activity. It is not available in real time so it cannot be used to analyze current forecasts. The most recent recordings are December of 1989.

Results from the Monte Carlo analysis for the real-time database were encouraging. Figure 6 shows data for a set of 22 forecasts made for a 25 km (or about 15 mile) radius from a point in the Northern part of the Loma Prieta zone, from May 28, 1990 through July 11, 1990. The X axis is the time interval in days after the forecast was made and the Y axis is the cumulative percent success rate for each time interval. The stippled area represents the success rate of the actual forecasts while the striped area below represents the chance success rate as determined by 100 Monte Carlo runs for each data point. The line shows the probability levels determined by a z-statistic. The best success difference over chance for the Loma Prieta area was 152%, for an eight-day interval after the forecasts were made. The success rate was 64% for the actual forecasts and random success was 42% ($p < .004$).

Figure 7 shows the same data presented a little differently. The stippled area shows the percent the actual forecasts were better than random forecasts, for each cumulative time period, while the line again shows the respec-

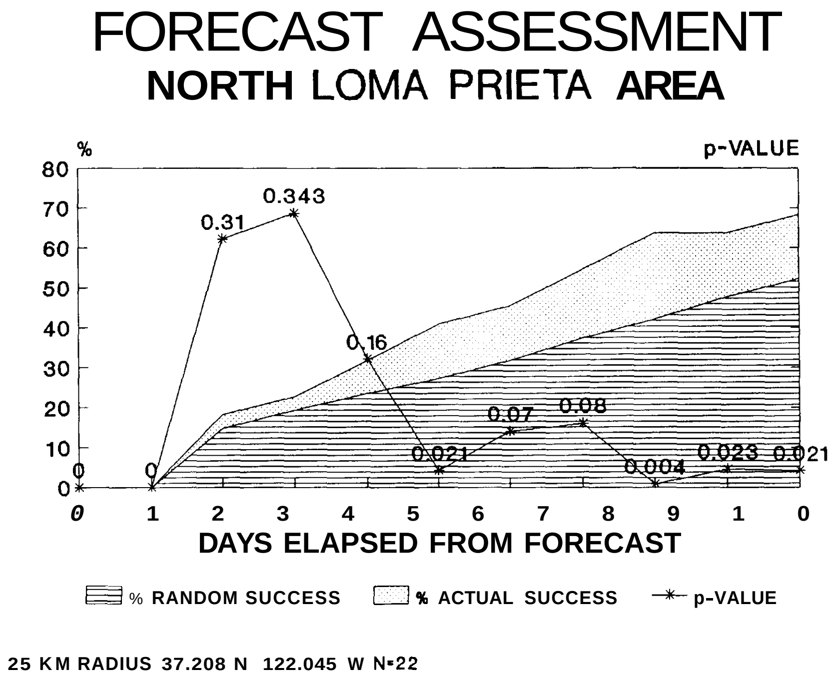

Fig. 6. Loma Prieta area: Stippled area shows the success rate better than chance cumulatively over a ten day period. Striped area shows chance success rate.

tive *p* values. You will note that the actual success rate increases above chance at two days and remains better than chance by about 150% for at least a 10-day period.

Figure 8 shows results for the East Bay area, the central point is north of Livermore, east of Danville. There were nine forecasts made during the same time interval. The best percent difference over chance occurred after one day when actual forecasts were almost 600% better than chance. However, a maximum of 44.4% total success was achieved by the fourth day of the forecast period compared to random success of 11.89%. Results were highly significant from day 1 through day 10. Note the better probabilities early in the forecast period compared to the Loma Prieta area.

Figure 9 shows the East Bay data in terms of percent better than chance. The figure shows a jump to about 600% above random at the first day. It is interesting to note that these relationships were very consistent. When the data set was split, similar results were obtained for earlier data and later data.

Figure 10 shows the analysis for the Vallejo area. There were two forecasts made for this area in the same period as the other forecasts. After five days a 100% success rate was achieved compared to random at 18%. This figure shows probabilities for days 5 through 10 only due to limitations of graphical scaling. The p values for days 1 through 4 ranged from $< .04$ to $< .097$.

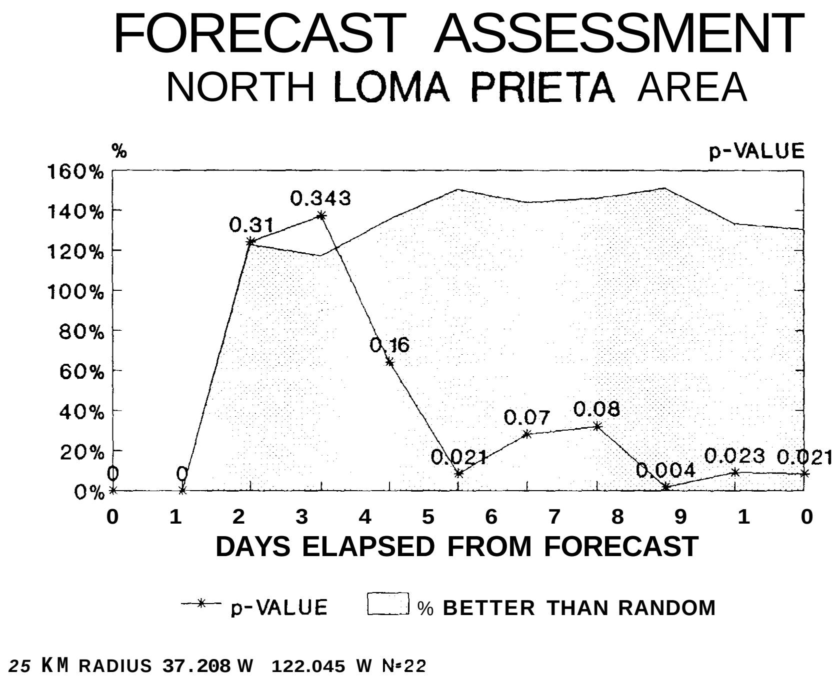

Fig. 7. Lorna Prieta area: Stippled area shows percent better than chance over cumulative time.

Figure 11 shows the actual success rate was 800% better than random on the 5th and 10th days after the forecast was made. Overall the forecasts did about 600% better than chance for the entire period.

The Vallejo area overlaps slightly with the East Bay area previously discussed. The point of showing the Vallejo area is to demonstrate that redundancy occurs within the system. When activity occurs between or near two areas, it is usually reflected by individually determined forecasts for both areas, yet each area analyzed by itself, can produce a significant result.

During the time interval analyzed, in the above areas anywhere from 2 to 6 qualifying earthquakes occurred for each forecast set.

Discussion

As stated earlier, these are the first results of analysis of this system. This research in process, and I have no final answers. Many haunting questions remain to be explored such as:

Duplication of these results with other time periods and data bases.
- The success rate for larger earthquakes in the 4 to 6 range.
- How many of the earthquakes that occurred were actually forecast?

These and other questions have high priority in the research plan. The question of duplication looks encouraging though only a few runs using the

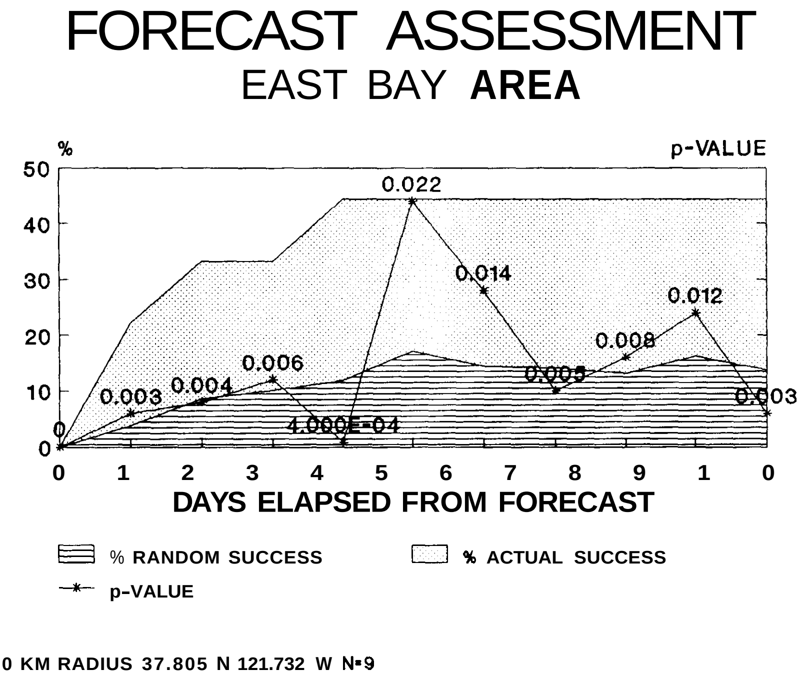

Fig. 8. East Bay Area: Stippled area shows the success rate better than chance cumulatively over a ten day period. Striped area is chance success rate.

USGS data base have been done so far. (a set of 100 runs takes about 1–7 hours of computing time depending on the length of the time interval, number of forecasts made, and number of earthquakes in the data base). Some of the analyses showed positive results ranging from 120–250% above random and significance levels of $p < .01$.

Does the system have a **useful application** now or in the near future? Analysis to date indicates that the system may be capable of issuing valid routine forecasts on a daily basis for earthquakes in the magnitude 2–4 range. We now have three good data points for large California earthquakes. Looking through the "retrospectoscope," these earthquakes could have been predicted based on the very strong signals observed beforehand. There has been a greater than magnitude 6 earthquake in California within a month or less, each time the signal strength has remained above about the 7 sigma level.

Based on the results of the forecasts in the magnitude 2–4 range, if a large signal were seen, it would inspire confidence to alert State disaster agencies of the high probability of an earthquake of magnitude 6 or more within a 30 day interval. There could be errors; it might not happen. But the cost of making such a forecast that did not come true would be minimal compared to the consequences of not making such a forecast. The costs of not making a

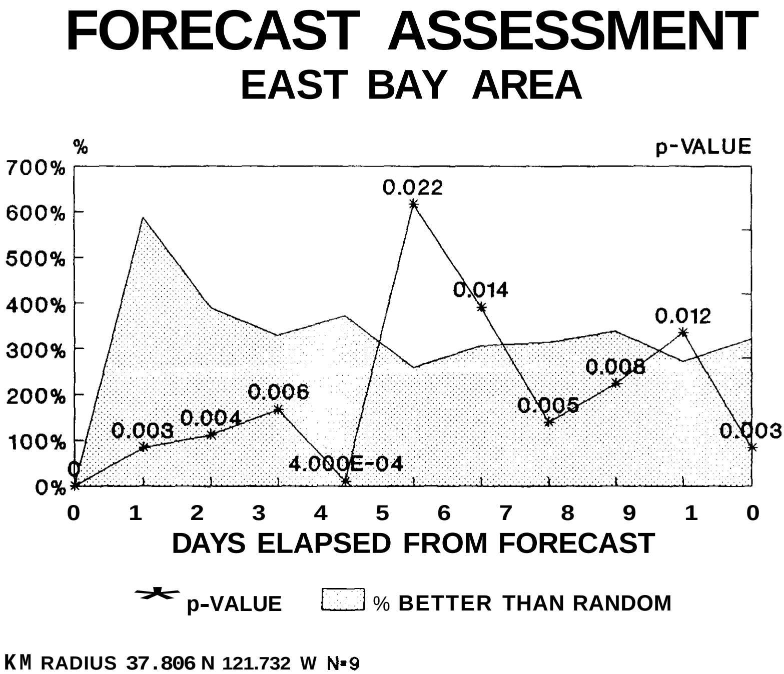

Fig. 9. East Bay Area: Stippled area shows percent better than chance cumulatively over 10 days.

forecast would be measured in lives, casualties, and property damage while the costs of an unsuccessful forecast would be to the reputation of the system.

In September of 89, a month before the Loma Prieta earthquake, at the very beginning of the signal increase, before the storm damage to the sensor, such a forecast was made. This is a research project, not production project, so I do not make public, proactive forecasts. However, I was concerned enough about this signal to alert about 30 of my neighbors. The only proactive forecast I had ever made to them to that date was: there would be an earthquake greater than magnitude 5.5 in the Loma Prieta area within the next week. I was about three weeks early on the timing but we all know what happened after that.

At that time I also realized the futility of contacting government agencies without having established a history, knowledge base, or track record of the system and its potential capabilities. Today, this paper will start that process. It may be none too soon. On Tuesday August 7th another signal increase began. It is smaller than the signals before the Lorna Prieta and Whittier earthquakes, but still remarkable; about a 4–5 sigma change. Artifacts, and malfunction of the equipment must be ruled out, and I am somewhat skeptical of the increase because it began at high noon, a time that could imply a man-made artifact. Solar activity does not appear to be a factor, however, we have had very active weather systems nearby. Nonetheless this signal in-

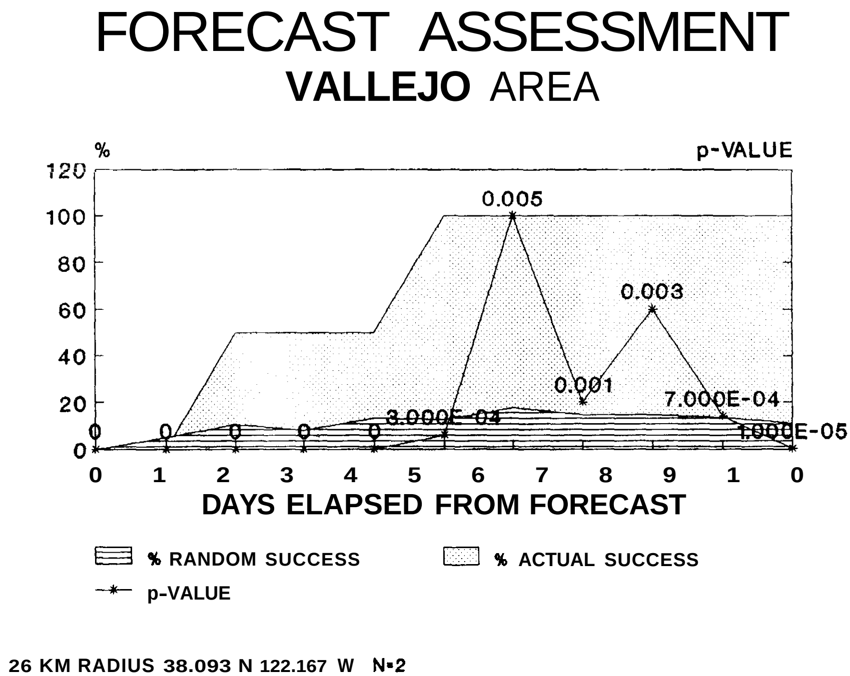

Fig. 10. Vallejo area: Stippled area shows the success rate better than chance cumulatively over 10 days. Striped area is chance success rate.

crease is unusual, and worth noting. The expert system has made a forecast for the East bay, but this could well be for coincidental activity in the 2–4 range. It only forecasts for certain areas and does not locate impending activity in general. It is also not known how it responds to a larger magnitude earthquake.

Conclusion

In conclusion, the question has been raised as to the coincidental or chance occurrence of EM signals observed prior to large seismic events. In this system we have seen large increases in EM signals prior to three large California earthquakes. There has been a one-to-one correspondence of these signals and subsequent large earthquakes over more than a nine year period. These large increases have been the only such increases in the history of the system. There have been no false alarms. The occurrence of such signals prior to large events is corroborated by Tony Fraser-Smith's observations. Therefore the conclusion that EM signals in the frequency range below 1 KHZ precede large seismic events is probably a valid one and certainly warrants intensive study.

The USGS has increased the probability for its forecast for a large earthquake in the Bay Area in the next 20 years. Given the past performance of

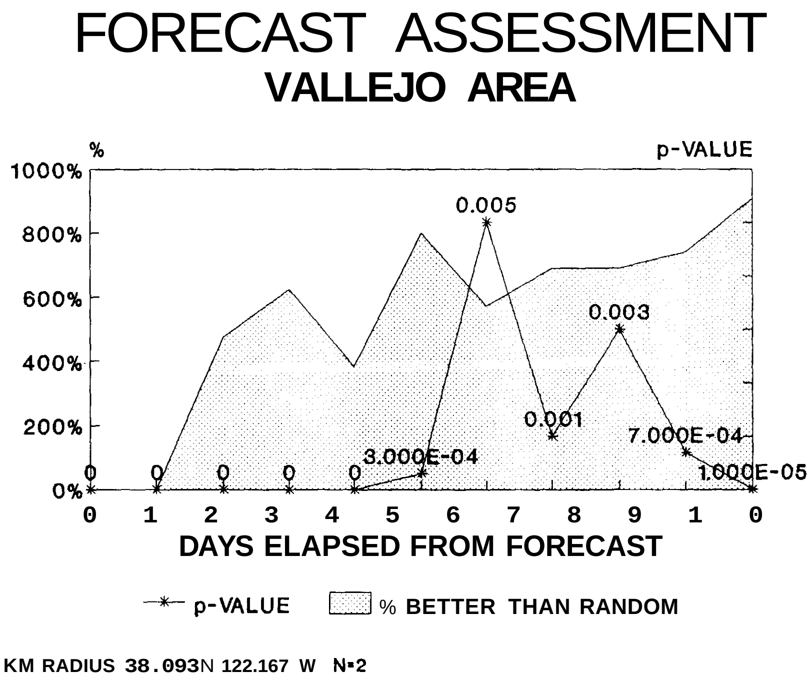

Fig. 11. Vallejo **area: Stippled area shows percent better than chance cumulatively over 10 days.**

this system one would expect to see signals similar to those observed prior to the other large California earthquakes at least a week to a month in advance of the event.

Preliminary evaluations of the feasibility of utilizing these signals in an expert system to forecast smaller earthquakes appears promising. This report on preliminary results is just the beginning of the exploration of this EM database. These are the results of the first trial of using the expert system to interpret the EM data. After almost 10 years of hands-on experience, I suspect that there is tremendous untapped potential here.

To realize the full potential of the system, increased manpower, and resources are necessary. Increasing the accuracy and shedding more light on the mechanisms by which these empirical observations function could yield enormous progress towards practical, cost effective earthquake forecasting the near future.

Post Scripts

1. The East Bay earthquake forecast in the body of this presentation occurred on August 13, 1990 four days after this paper was given. A magnitude 2.8 earthquake occurred south of Hayward.
2. The increased signal activity has continued to be observed through the

time of submitting this document on October 15, 1990. Strong signals have now been observed for more than two months. This is unprecedented for this system. The signals have increased in amplitude to levels as high as those that occurred before the Lorna Prieta earthquake. Numerous aberrations and anomalies have occurred in the signals that have not been observed before. 90% of equipment malfunction possibilities have now been ruled out.

3. Since the strong signals began on August 7, 1990 the rate of earthquakes in the Bay Area, and as far South as Hollister appears to have increased.
4. The signal source is difficult to locate as the expert system that identifies location was designed to operate on lower signal levels and is functioning suboptimally due to the extreme level of the signals.
5. The largest earthquake to occur in Northern or Central California since the Loma Prieta earthquake (ML = 5.7), happened on October 23rd near Mono Lake. Landslides closed roads to nearby Yosemite Valley.