AGU Advances

Peer Review History of

Robust Earthquake Early Warning at a Fraction of the Cost: ASTUTI Costa Rica

Benjamin A. Brooks[1], Marino Protti[2], Todd Ericksen[1], Julian Bunn[3], Floribeth Vega[2], Elizabeth S. Cochran[4], Chris Duncan[5], Jon Avery[6], Sarah E. Minson[1], Esteban Chaves[2], Juan Carlos Baez[7], James Foster[6], Craig L. Glennie[8]

[1]Earthquake Science Center, US Geological Survey, S 245 N. Akron Rd., Moffet Field, CA 94035, USA
[2]Observatorio Vulcanologico y Sismologico de Costa Rica Campus Omar Dengo, Heredia, Apartado Postal: 2386-3000, Costa Rica
[3] Center for Data Driven Discovery, California Institute of Technology, Pasadena, California, USA
[4] Earthquake Science Center, US Geological Survey, 525 S. Wilson Ave., Pasadena, CA 91106
[5] GISmatters
[6] Hawaii Institute for Geophysics and Planetology, School of Ocean and Earth Science and Technology, University of Hawai`i, 1680 East-West Road, POST Building #602 Honolulu, HI 96822
[7]Centro Sismologico Nacional de Chile
[8]NCALM, University of Houston, Houston, TX

Files Uploaded Separately

Original Version of Manuscript (2021AV000407)
First Revision of Manuscript (2021AV000407R)
Second Revision of Manuscript [Accepted] (2021AV000407RR)
Author Response to Peer Review Comments

Peer Review Comments on 2021AV000407R

Reviewer #1

The author provide a smartphone-based EEW system and show its performance in Costa Rica. What I am most interested in is the data that generates by smartphones (or precisely the MEMS sensor in the phone). As the author didn't mention the phone model that was used in this EEW system, readers may want to know more about the data situation. So I suggest you add some more details about the phone model selection, such as what's the considerations about the phone model? Did you take any tests on different models (MEMS)? and what are the main differences? And I partly-disagree with the authors' view about "Smartphones used in a fixed network can provide earthquake early warning performance similar to scientific-grade instrumentation". The author indeed shows the reliablity of this smartphone-based EEW system, but it will be more convincing if the author could provide the lower limit magnitude of this system.

[Please see attachment that begins on the next page.]

The author provide a smartphone-based EEW system and show its performance in Costa Rica. What I am most interested in is the data that generates by smartphones (or precisely the MEMS sensor in the phone). As the author didn't mention the phone model that was used in this EEW system, readers may want to know more about the data situation. So I suggest you add some more details about the phone model selection, such as what's the considerations about the phone model? Did you take any tests on different models (MEMS)? and what are the main differences? And I partly-disagree with the authors' view about "Smartphones used in a fixed network can provide earthquake early warning performance similar to scientific-grade instrumentation". The author indeed shows the reliablity of this smartphone-based EEW system, but it will be more convincing if the author could provide the lower limit magnitude of this system.

Reviewer #2

Review for AGU Advances
Smartphone-based Earthquake Early Warning with a Fixed Network: ASTUTI Costa Rica
Brooks, B. et al.

I consider that the paper submitted by Brooks and his co-authors constitutes a novel and interesting idea that will be of interest to a broad range of readers in the Earth and social sciences communities. As the authors stress, and I fully agree with them, this technology, once perfected and more developed, may represent an interesting alternative to the deployment of Earthquake Early Warning Systems (EEWS) in countries that have neither the funding nor the communications infrastructure to install and operate an EEWS with more sophisticated instrumentation.

I believe the paper is very interesting and innovative and I would recommend its publication in AGU Advances. I would encourage the authors to consider some of the comments written below, as well as those that I added as corrections to the *Word* document attached to this review. I used this *Word* document because it was easier than to comment on the tightly spaced *pdf* pre-print available. Also, it has comments included by other reviewers (?). I hope that this marked document makes it easier for the authors to produce a revised manuscript.

Other comments are listed below:

1. The paper is extremely long. At times, it seems like the authors want to publish a review of EEWS. It is always important to provide the background, but the authors go to unnecessary lengths. I would encourage them to shorten some sections substantially, I elaborate below.

2. The paper seems to be written in haste and not thoroughly checked before submitting. Some sentences are awkward, confusing or simply unnecessary; I point out many of them in the marked document.

3. An example of this impression that it was rapidly put together is that in many cases the authors cite some references in the text with the full name and last name of the authors and not simply by last name as is the norm. This certainly gives the impression that the paper was not proofed critically by the authors before submitting.

4. I agree with the authors that for a country the size of Costa Rica, an EEWS that warns for every earthquake detected may be acceptable. However, what the authors call the ¨boy cry wolf¨ syndrome may wear down the population in time. Having worked in collaboration with the Mexican EWS, I know that the public is very demanding and unforgiving with systems like this. In addition, the authors only touch lightly on the presence of intermediate-depth and crustal earthquakes in Costa Rica. They admit that

ASTUTI may not work for them, but this is done only in passing. I would ask the authors to consider a future magnitude threshold and mention it at least in this initial paper.

5. There are some lengthy discussions that could be summarized in Tables for the benefit of the authors. I found it difficult and boring to read, all the statistics regarding the delay times for the earthquakes tested. Also, the discussion towards the end of the paper of percentages of the population that felt shaking or not and its relation to the DYFI data is extremely dry. I would encourage the authors to consider putting the data on Tables and refer the readers to them.

6. Following on the last comment, the Discussion section is too long and repetitive of what is said previously in other parts of the paper. Also, it makes the Conclusions section seem unnecessary and gratuitous. The authors know well that many scientists today read carefully the abstract and conclusions of papers. The authors should give a critical review of how this section is written today and shorten it, emphasizing their findings and innovations.

7. Like all EEWS, to become a useful tool for society, the public has to be educated in its use. I know this is no easy task. However, I believe the authors´ definition of Missed Alerts covers too broad a range of earthquakes. I think the authors have to be realistic as to where earthquakes need to occur for ASTUTI to be an effective and useful tool. Whenever they roll it out, this has to be made clear. I encourage them to think about it and include it in the paper.

8. In terms of the False Alerts, the authors point out to power failures of the grid line; a common problem in many countries. Although the idea is to keep it as a low-cost system, an inexpensive power supply would solve this problem. It is not clear frome their discussion how many times this happened.; it would be useful to know. I cannot help but to suggest to the authors also reference the statistics of SASMEX in terms of FA: one false alert during its operational life.

9. Some figures are confusing and of low quality. For example, Figure 7a shows both the average ASTUTI latency for all events (or at least this is what I understand of it) together with a discussion on the performance of the system during the 27 July 2020 earthquake. Figure 7a is not referred to in the text and the reader, as I did, is forced to speculate what it means and what it represents. If Figures 7b, 7c and 7d refer only to the 27 July quake and 7a reflects the median latency for all quakes they should not be in the same Figure. This needs to be corrected.

10. In Figure 8 what authors call copper population is confusing. There are too many color bars. The ones for DYFI are clear. The others, however, are confusing. Three color bars have the title Pop. What do they mean? I would encourage the authors to improve the figure making it more didactic and clarifying the caption.

11. The main question to me is whether Amazon Web Service Notification will work when the system is implemented fully, and thousands of users need to be notified (maybe a few million). The authors touch on this in passing. I admit to being totally ignorant of this web service, however, having witnessed how pilot programs using cell phone notification worked well for a reduced number of users but failed dramatically when the number of users increased, I have a grave concern whether this solution will work. Needless to say, if you do not notify broadly and on timely, the system is futile, no matter how sophisticated it may be technically. Not understanding how this web notification works, I would encourage the authors to include tests that confirm that it is a usable alternative.

12. Finally, the question of why the authors, particularly the local ones, do not mention the option to use the excellent national strong motion network data is puzzling. Costa Rica has perhaps the best network of this nature in Latin America. I understand it is run by an independent organization but building a valuable system like this the question of whether it can be integrated and how to a future EEWS pops up naturally.

Gerardo Suárez
Instituto de Geofísica UNAM

[Please see the attachment that begins on the next page.]

Smartphone-based Earthquake Early Warning with a Fixed Network: ASTUTI Costa Rica

Benjamin A. Brooks[1], Marino Protti[2], Todd Ericksen[1], Julian Bunn[3], Floribeth Vega[2], Elizabeth S. Cochran[4], Chris Duncan[5], Jon Avery[6], Sarah E. Minson[1], Juan Carlos Baez[7], James Foster[6], Craig L. Glennie[8]

[1]Earthquake Science Center, US Geological Survey, S 245 N. Akron Rd., Moffet Field, CA 94035, USA

[2]Observatorio Vulcanologico y Sismologico de Costa Rica Campus Omar Dengo, Heredia, Apartado Postal: 2386-3000, Costa Rica

[3] Center for Data Driven Discovery, California Institute of Technology, Pasadena, California, USA

[4] Earthquake Science Center, US Geological Survey, 525 S. Wilson Ave., Pasadena, CA 91106

[5] GISmatters

[6] Hawaii Institute for Geophysics and Planetology, School of Ocean and Earth Science and Technology, University of Hawai`i, 1680 East-West Road, POST Building #602 Honolulu, HI 96822

[7]Centro Sismologico Nacional de Chile

[8]NCALM, University of Houston, Houston, TX

Corresponding author: Ben Brooks (bbrooks@usgs.gov)

Key Points:

- Smartphones used in a fixed network can provide earthquake early warning performance similar to scientific-grade instrumentation
- Operating in Costa Rica, over 6 months the ASTUTI (Alerta Sismica Temprana Utilizando Teléfonos Inteligentes; Earthquake Early Warning Utilizing Smartphones) project detected 5 events that caused wide-spread felt shaking and had 0 false alarms
- The detections and alerts would have provided time for drop-cover-hold-on (DCHO) protective actions to be taken before S-wave arrival for large percentages of the Costa Rican population
- 2 of the 5 events were triggered by P-waves on the phones, suggesting that smartphone-based EEW could be more effective than previously thought

PEER-REVIEW DISCLAIMER

This draft manuscript is distributed solely for the purpose of scientific peer review. Its content is merely being considered for publication, so it is not to be disclosed or released by reviewers. Until the manuscript has been approved for publication by the U.S. Geological Survey (USGS), it does not represent any official USGS finding or policy.

Abstract

We show using the ASTUTI (Alerta Sismica Temprana Utilizando Teléfonos Inteligentes) network that smartphones deployed in a fixed network can provide Earthquake Early Warning performance comparable to scientific-grade systems. Using the phones' accelerometers, we implement a ground-motion detection and alerting strategy focusing on subduction zone earthquakes in Costa Rica. Our strategy considers that much of the country's population is concentrated near the capital San Jose and that nearly the entire population experiences shaking during earthquakes of ~ M_w 6. Rather than waiting until an earthquake reaches a high intensity, we evaluate issuing alerts when an acceleration threshold is exceeded at multiple stations. From more than six months of observations we find data latency is 0.35-0.45 secs. We simulate the 2012 M7.6 Nicoya earthquake using on-phone vibrations and find median first-alert latency of ~9-13 secs after origin time and alert-receipt latency ~~using Amazon Web Services Simple Notification Services~~ of ~4 secs. During our assessment there were 13 earthquakes that caused felt shaking. From offline reanalysis, we detected and alerted on 5 of these, all of which produced felt shaking above ? g in San Jose. The system did not produce any false alerts and undetected events (define these events) did not produce wide-spread felt shaking. If we alerted the entire population for each event, 70% and 15% of the population would receive alerts in time to undertake drop-cover-hold-on for events inside and outside of the network, respectively. ~~Complementary~~ Some population ~~ranges~~ centers would receive alerts and not feel shaking. This strategy may be effective if users are tolerant of feeling no shaking when they receive alerts from correct detections.

Plain Language Summary

We show that a network of smartphones deployed in fixed locations can provide Earthquake Early Warning performance on par with scientific grade instrumentation. This approach makes EEW ~~to~~ accessible to resource-limited countries that may ~~have~~ not been ~~to~~ able to ~~previously~~ benefit from some form of EEW. In particular, combining the fixed-network smartphone sensors with a strategy targeting low cost ~~of action~~ protective measures such as Drop Cover Hold On could be very effective.

1 Introduction

~~Earthquake early warning (EEW) attempts to rapidly detect earthquakes and to alert people and systems with enough time for protective actions to be taken before damaging shaking arrives [*Heaton*, 1985]. EEW systems are operational in a handful of regions and they are now rapidly being developed and adopted globally [*Allen and Melgar*, 2019]. As with many techno-scientific advances with the potential for large societal impacts, however, the earliest projections of EEW system performance and benefits to society have been modified and scaled back as the theoretical~~ Earthquake early warning's (EEW) fundamental promise is that earthquakes can be rapidly detected so that people and systems can be alerted to take protective action before shaking arrives at their location [*Heaton*, 1985]. In order to maximize warning time and to minimize the population not receiving sufficient warning, EEW requires a dense sensor network so that earthquakes can be detected closest to ~~wherever they may~~the point of nucleation~~e~~. ~~The locations where earthquakes nucleate and where people reside, however, might be quite removed from one another.~~ This requirement for dense sensor networks, combined with the high cost of expensive scientific-grade sensors currently limits EEW systems to wealthy countries [*Allen and Melgar*, 2019(I don't consider Mexico to be a wealthy country and it built the first public EEWS in the world. The financial problem was solved using home made, low cost instruments)] . Alternatively, a new generation of low-cost accelerometer and geodetic sensors [*E. S. Cochran*, 2018] could make EEW generally accessible. In particular, utilizing smartphones, including ~~via~~ crowd-sourcing, is a potentially transformative way to provide EEW [*Minson et al.*, 2015; *Finazzi*, 2016; *Kong et al.*, 2016]. Although crowd-sourcing removes sensor cost from EEW budgets, it is fundamentally limited, in contrast to fixed networks, by its inability to consistentely leverage the distance between earthquake source regions and population centers (This sentence is awkward. Please rewrite). Smartphones can~~, of course,~~ also be employed in fixed networks. With fixed networks, depending on the location of the event and the position of the sensors, all users could potentially get warnings (Figure 1a). ~~By definition~~In contrast, ~~however,~~ some users in a crowd-sourced EEW system ~~will~~may ~~never~~not get warnings (Figure 1b). Combining the detection advantage with the low capital and communications costs of smartphones may make fixed networks an ~~the most~~ optimal low-cost EEW configuration, perhaps even comparable to scientific grade systems. To the best of our knowledge, however, smartphone fixed network performance has never been examined. In this paper, with a fixed network in Costa Rica, we demonstrate that judiciously placed smartphones can provide operational EEW for a country's population centers at a cost drastically reduced from, and at a performance level equivalent to, scientific grade instrumentation.

As with many techno-scientific advances with the potential for large societal impacts, the earliest projections of EEW system performance and its benefits to society have been modified and scaled back as EEW's theoretical, empirical, practical, and social limitations are better understood [*M A Meier*, 2017; ~~*Sarah E.*~~ *Minson et al.*, 2018; *S*~~*.E.*~~ *K McBride et al.*, 2019; *Minson et al.*, 2019; *Nakayachi et al.*, 2019; *Trugman et al.*, ~~2019~~2019; *Becker et al.*, 2020; *S McBride et al.*, 2020; *David J Wald*, 2020b]~~, empirical [*M A Meier*, 2017; *Trugman et al.*, 2019], practical [*David J. Wald*, 2020a], and social [*Nakayachi et al.*, 2019; *Becker et al.*, 2020; *S McBride et al.*, 2020] limitations of EEW are elucidated. In parallel with this evolution in understanding, EEW instrumentation and network design has also rapidly evolved. Although most systems, such as ShakeAlert in the U.S. [*Given et al.*, 2018; *Kohler et al.*, 2020] use traditional seismic instrumentation, it has recently become clear that low-cost, consumer-grade~~

Comment [GS1]: I

~~sensors in smartphones including MEMs accelerometers [*E Cochran et al.*, 2009; *J Evans et al.*, 2014; *Kong et al.*, 2016] and GNSS chips [*S.E. Minson et al.*, 2015] are sensitive enough to be effectively used in EEW systems, even in crowd-sourced modes [*Lawrence et al.*, 2014; *S.E. Minson et al.*, 2015; *Finazzi*, 2016; *Kong et al.*, 2016; *Finazzi and Fassò*, 2017; *Kong et al.*, 2018; *Kong et al.*, 2019b; *Finazzi*, 2020; *Kong et al.*, 2020].~~

~~Given this rapidly evolving context, it is important to consider that a one-size-fits-all approach to EEW may not be appropriate, in particular for resource-limited populations. Varied EEW objectives and performance requirements could likely be acheived with different combinations of instrumentation, algorithms, alerting strategy, and messaging strategy [*Y. Fujinawa et al.*, 2011]. For instance, EEW objectives for automated industrial systems with large action-costs (such as nuclear power plants) may require highly sensitive instrumentation, high ground-motion prediction accuracy levels, low false alarm tolerance, and large annual budgetary allocations [*Strauss and Allen*, 2016]. Alternatively, if the EEW objective is to promote population participation in an immediate low cost-of-action response such as Drop-Cover-Hold-On (DCHO) [*Porter and Jones*, 2018], then sensor sensitivity may be lower, ground-motion prediction accuracy may not be as important, false alarm tolerance may be higher, and budgetary allocation may be much smaller. Our work here examines this low-budget end of the spectrum.~~

~~Specifically, we ask: can off-the-shelf smartphones deployed in a fixed network provide EEW in a reliable and effective enough manner so that significant populations receive warnings in time for DCHO protective actions to be undertaken? Instead of investing in the institutional overhead of designing, building, and maintaining our own sensing hardware and communications platform [*Anthony et al.*, 2019], the smartphone network approach leverages the massive research and development programs of smartphone providers such that sensors, field computing, and communications platforms are provided in a single compact unit. Moreover, the constant competition in this consumer space assures frequent technological updates as instrumentation becomes obsolete. Although there are now multiple examples of smartphone-based EEW methodologies and capabilities [*S.E. Minson et al.*, 2015; *Finazzi*, 2016; *Kong et al.*, 2016; *Finazzi and Fassò*, 2017; *Kong et al.*, 2018; *Kong et al.*, 2019b; *Kong et al.*, 2020], to the best of our knowledge, there have been no evaluations of smartphones used, over extended periods of time, to provide EEW for a specific region.~~
. To motivate and place in context our work, we review these recent developments, in particular as they pertain to smartphone-based EEW.

Ideally, an EEW system would issue actionable alerts prior to P~~-~~ (compressional)-wave arrival, affording the most time possible for protective actions. Even though a felt P-wave would hopefully serve as a natural EEW system alerting people to imminent shaking [*David J. Wald*, 2020a], site and path effects can preclude P-waves being felt ubiquitously, so a minimum criterion for an EEW system to be effective is for it to deliver alerts prior to the subsequent arrival of strongest shaking associated with S-waves. Currently, many operational EEW systems detect earthquakes using some permutation of point-source ~~parameter estimation from the early P-wave. For instance, ShakeAlert uses a small amount of the P-wave (>0.2 secs) from 4 stations and empirical scaling relations to estimate point-source event location and magnitude~~magnitude estimation from P-wave information [*Chung et al.*, 2019]~~. In Japan, P-wave parameter estimation is augmented with a particle filter technique to combine ground motion observations and~~

~~information about where shaking has not been observed [*Tamaribuchi et al.*, 2014]. In Mexico, an event's body wave magnitude, m_b, is estimated empirically from the rate of seismic energy released between P- and S-wave arrivals at a series of free-field acceleration stations along the country's subduction zone coast [*Cuéllar et al.*, 2017].~~. Point-source-based parameter estimation, albeit fast at initial detections, exhibits degraded performance when events become large and magnitude estimates saturate [*Hoshiba et al.*, 2010]. Moreover, theoretical studies have recently found that the conventional EEW approach of using source parameters to forecast shaking causes alerts to be too slow for higher levels of shaking [*M A Meier*, 2017; ~~*Sarah E.*~~ *Minson et al.*, 2018; *Trugman et al.*, 2019]~~,~~ and to mostly produce missed and false alerts due to the intrinsic ~~factor-of-two~~ variability of ground motion ~~[*Gregor et al.*, 2014; *S.E. Minson et al.*, 2019]. Accordingly, EEW practioners are exploring and/or incorporating alternate methodologies including line-source estimates from strong-motion data [*Böse et al.*, 2012], surface displacements from GNSS data [*S. E. Minson et al.*, 2014], and~~[*Gregor et al.*, 2014; *Minson et al.*, 2019]. Accordingly, EEW practioners are developing ground-motion based approaches that issue alerts when one or more stations observe shaking above a threshold [*Kodera et al.*, 2018; ~~*E*~~*Elizabeth S Cochran et al.*, 2019].

~~In particular, the Propagation of Local Undamped Motion (PLUM) algorithm is a ground-motion-based approach that was developed as an alternative to source-parameter EEW methodologies [*Kodera et al.*, 2018; *E S Cochran et al.*, 2019; *M A Meier et al.*, 2020]. PLUM arose in response to the 2011 M9.0 Tohoku-oki earthquake, whose productive aftershock sequence led to many missed events and false alarms from the point-source algorithm when simultaneous small earthquakes were combined into a single larger event (Hoshiba et al., 2011; Kodera et al., 2018). Unlike most EEW methods that attempt to characterize source parameters and forecast expected shaking, PLUM uses shaking to directly forecast shaking. By~~For instance, by using strong ground motion that accrues sometime between P- and S-wave arrivals to directly forecast ground motion, alerts can be issued for strong shaking as soon as it is observed without having to wait for the rupture (and earthquake magnitude) to grow in size~~. PLUM handles inter-event variability well, reducing false and missed alerts due to inaccuracies in the ground motion forecast~~ [*Kodera et al.*, 2018; ~~*E*~~*Elizabeth S Cochran et al.*, 2019].

Formatted: Font: Not Italic

~~In addition to detection methodology, alerting criteria are also quite variable. ShakeAlert issues alerts for ground motions predicted to be greater than Modified Mercalli Intensity (MMI) III [*Thakoor et al.*, 2019]. In Japan, warnings are issued by the JMA (Japan Meteorological Agency) to seismic hazard blocks [*Yukio Fujinawa and Noda*, 2013], roughly four per prefecture, when any evaluation point inside a block is predicted to experience shaking greater than JMA IR (intensity reading) 4.0 (MMI VI), but only after an IR of 5L (MMI 7) is forecast in at least one area [*Kamigaichi et al.*, 2009; *Kodera et al.*, 2018]. In Mexico, if rapidly estimated magnitude is greater than m_b 6, public warnings are issued in Mexico City and other inland locales. Alerting criteria are further~~In addition to detection methodology, EEW alerting criteria are also quite variable. Alerting criteria are complicated by the simple physical fact that earthquake magnitude is not static; rather, it increases as an earthquake rupture evolves. The promise of EEW systems has been amplified by the idea that differences between the ways large and small earthquakes start~~-~~ could potentially be used to predict the final size of a growing earthquake, rather than relying on rapidly updating contemporaneous estimates of source parameters and ground motion [*Olson and Allen*, 2005; *Melgar and Hayes*, 2019]. However, a growing number of studies have shown that the early P-wave holds little to no predictive power [*M-A Meier et al.*, 2017; *Goldberg et al.*, 2019; *Trugman et al.*, 2019]. Furthermore, *Minson et al* (2018) recently showed that given the velocities at which earthquake ruptures grow and seismic waves propagate, together with the exponential decay of ground-motion with distance from a propagating rupture, users' warning time for the strongest shaking will often be brief if EEW systems wait to issue alerts until earthquakes grow large enough that their forecast shaking is strong. ~~Accordingly, it is becoming clear that EEW offers the most potential for successful mitigating actions~~Accordingly, it is becoming clear that EEW offers the most potential for successful mitigating actions, such as Drop-Cover-Hold-On (DCHO) [*Porter and Jones*, 2018] to be taken if alert thresholds are set at lower ground-motion levels than those expected to cause damage [~~*Sarah E.*~~ *Minson et al.*, 2018; *Saunders et al.*, 2020].

~~It is important to consider that while~~Although an EEW system with low alert thresholds would likely produce alerts without ~~subsequent~~ strong shaking in some locations, the number of missed alerts would be minimized and the system would have the best chance of being effective ~~when it mattered most~~ during medium to larger events. For such a strategy to be successful, however, users must be more tolerant of receiving unnecessary alerts (the 'Boy Who Cried Wolf' phenomena) than is often assumed. Although there are essentially no studies explicitly addressing the 'Boy Who Cried Wolf' phenomenon for EEW, other scientific communities with more established alerting relationships with users caution against the implicit assumption that user false alarm tolerance is low [*Roulston and Smith*, 2004]. In fact, the small amount of evidence that exists suggests that EEW users may be ~~surprisingly~~ false-alarm tolerant. *Nakayachi et al* (2019), found from surveys associated with alerts from two Japanese events that, indeed, the user population was surprisingly tolerant of false alerts. Anecdotal surveys in Mexico found similar attitudes [*Allen et al.*, 2018].

In this contribution, we implement the ASTUTI (Alerta Sismica Temprana Utilizando Teléfonos Inteligentes; Earthquake Early Warning Utilizing Smartphones) network ~~and we evaluate the coupling of a fixed smartphone network and ground-motion based detection methodology with a low-threshold alert strategy. Smartphone MEMs accelerometers are well-suited for ground-motion based EEW methodologies, having been shown to have noise levels sufficiently low to permit detection of small (M4-5) earthquakes [*Kong et al.*, 2016; *Finazzi and Fassò*, 2017; *Kong et al.*, 2018; *Kong et al.*, 2019a; *Kong et al.*, 2019b; *Finazzi*, 2020; *Kong et al.*, 2020]. Although we introduced the concept of smartphone crowd-sourced EEW [*S.E. Minson et al.*, 2015], we maintain that it is prudent to question crowd-sourcing's operational reliability and overall practicality. Even though efforts such as The Earthquake Network Project [*Finazzi*, 2016] and MyShake [*Kong et al.*, 2016] report high initial enrollment rates, and even though Google has recently announced that the Android operating system will have native ability to use accelerometer data for EEW detection purposes [*Press*, 2020], there is no guarantee that users will continue to permit their phones to be utilized for extra-personal purposes, especially as internet privacy advocacy becomes a more penetrative societal issue [*Ketelaar and Van Balen*, 2018]. Additionally, depending on citizen subscribers, with their changing preferences and resources, likely poses a substantial risk in operating a warning system that requires continuous data from the near-source regions of the most damaging expected earthquakes. Finally, as~~in Costa Rica. ~~For~~During six ~~6~~ months of continuous operation of the 82 station network, we assess the coupling of the fixed smartphone network with a ground-motion based detection methodology~~.~~(Please clarify this sentence. In what context do you use the word coupling?) We evaluate a low-threshold alert strategy that would alert the entire country upon the unambiguous detection of an event (Any magnitude?). The spatial distribution of events during the ~~time~~ period appears to be representative of the general spatial distribution of Costa Rican seismicity and allows us to examine system performance for events that occur within and outside of the network. Specifically, for each event, we quantify the percentage of Costa Rican population that would have sufficient warning time to undertake DCHO mitigating actions. Finally, all ~~of the~~ raw data collected for this project are freely available. As commercial entities with proprietary algorithm and data policies are poised to increase by orders of magnitude the number of smartphone EEW~~-capable~~ sensors and to issue public-safety alerts themselves [*Stogaitis et al.*, 2020], we assert that it becomes ever more important that the scientific community have access to transparently collected and freely available smartphone data.

Comment [GS2]: (Please cl

Comment [GS3]:

Formatted: Font: Not Italic

2 Materials and Methods

2.1 System Design

Costa Rica's seismic hazard is due primarily, though not exclusively, to earthquakes generated by oblique subduction of the Cocos plate below the Caribbean plate at rates up to ~ 8.5 cm/yr along the Middle America Trench (MAT) [*DeMets et al.*, 2010; *Protti et al.*, 2014](Figure ~~1~~2). In contrast to other subduction zones, ~~the portion of the plate interface where~~ MAT earthquakes in Costa Rica ~~tend to~~ nucleate ~~is overlain by land~~ under the continent and not ~~rather than the~~ offshore ~~seafloor~~ [*Protti et al.*, 2014]. Since 1853, 8 earthquakes greater than *Mw* 7.0 have occurred either on the northern Nicoya (1853, 1900, 1950, 2012) or southern Osa (1856, 1904, 1941, 1983) peninsulas [*Kobayashi et al.*, 2014; *Protti et al.*, 2014]   The majority of Costa Rica's ~5M population resides in the greater San Jose region, ~60-200 km from the Pacific coast and principal seismogenic zone (Figure ~~1~~2). For greater San Jose as well as Costa Rica's coastal regions, 10% expected peak ground acceleration (PGA) exceedance over the next 50 years is 0.55-0.90g, in the upper ranges of global seismic hazard [*Pagani et al.*, 2020]. Recently, the *Mw* 7.6 2012 Nicoya peninsula earthquake ~~caused shaking throughout~~was felt in the entire country, with PGA values as high as 0.5-1.4g and MMI V-VII reported in San Jose (Figure ~~1b~~2b)~~, and it is~~. ~~generally~~ Generally, ~~accepted tha~~earthquakest $M > 6$ events are felt country-wide.

~~Given~~ Based on Costa Rica's seismotectonic framework and population distribution we designed the ASTUTI network with these principles: (1) Our EEW efforts are focused on warning people, not automated systems, and our targeted user response is DCHO. Because the cost of taking action for DCHO is so low, this implies that detection and alerting thresholds can be low. (2) We prioritize detecting and alerting for MAT earthquakes, the events that have the highest probability of affecting the largest percentage of the population, especially in San Jose. Although San Jose's central location exposes it to earthquake sources from the entire country, most non-MAT sources would be too close to San Jose to permit warnings. (3) If we wait to issue an alert until an event has grown large enough to cause shaking damage, then it will most likely be too late to issue actionable warnings ~~[*Sarah E. Minson et al.*, 2018; *Trugman et al.*, 2019]~~[*Minson et al.*, 2018; *Trugman et al.*, 2019]. Accordingly we attempt to issue warnings at the earliest detection of events of potential concern. (4) Because of local ground-motion variability ~~[*S.E.*~~

~~*Minson et al.*, 2019]~~[*Minson et al.*, 2019] and because it is unlikely that earthquake ruptures are deterministic [*M-A Meier et al.*, 2017; *Goldberg et al.*, 2019; *Trugman et al.*, 2019] our detection and alerting is entirely non-parametric. In addition to making no attempt to estimate source information (such as location and magnitude), the alert does not include information about predicted ground-motion levels. (5) Based on the previous design principles, Costa Rica's generally small areal extent, and the country-wide shaking from MAT events such as the 2012 Nicoya event, we do not attempt to designate intra-country warning precints; rather we evaluate the scenario where every alert will be issued for the entire country. (6) Because of the previous design priorities, there will likely be a number of alerts issued for smaller events when users will feel no shaking although an event was correctly detected. This will require rapid post-event messaging and constant user interaction and education to remind users that the system performed correctly even if they did not feel shaking [*S McBride et al.*, 2020]. We stress, however, that, as of January 2021, we are not issuing public alerts, aside from those sent to our small group of beta testers. We leave thorough investigation of this topic to a future paper.

2.2 Hardware, Network, and Data Architecture

From September to December 2019 we constructed the ASTUTI network (Figure ~~1a~~2a). For the duration of the testing period, the network comprised 82 stations. Closest to the MAT and most of the strongest expected sources, station spacing is ~30 km and it increases to 30-50 km away from the MAT. Higher station density closer to the subduction zone is similar to the ~~SASMEX~~ configuration in Mexico [*Espinosa-Aranda et al.*, 2009; *Cuéllar et al.*, 2017; 2018; *Suárez et al.*, 2018]. Phones are installed on the ground floor~~s~~ of buildings in protective boxes and affixed to floors or walls (Figure ~~2~~3). Previously, in addition to accelerometer data from smartphones, we have discussed and used GNSS, which typically requires outdoor installation ~~[*S.E. Minson et al.*, 2015]~~[*Minson et al.*, 2015]. We found that many devices overheated, however, so for the initial phase of ASTUTI we made the operational decision to only employ phones installed on interior walls. Accordingly, the only data we use for ASTUTI are from ~~on-board~~ the built-in accelerometers. To date, our approach utilizes smartphones with the Android Operating System ~~(Android OS).~~. Criteria for phone choice included in-country availability and cost (Supp. Mat).

~~Onboard~~ Inside the phones, our control and sensing software ~~is called QED (for 'Quick Event Detection'; Figure 3; Supp. Mat). QED~~ controls sensor sampling and logging, detects events, ~~and~~ prepares data ~~(via formatting, filtering, or on-board data reduction)~~ for use in various downstream processing algorithms, and sends data in short, labeled messages~~. QED harnesses as much on-board computation from the smartphone's CPU as possible, thereby eliminating some computational load (and potential latency increases) when large numbers of stations are active in the network. QED supports either continuous or triggered streaming at sampling rates up to 100 Hz for accelerometer data.~~ (Figure 4; Supp. Mat).From December 2019 to late August 2020, phones streamed at 10 Hz; subsequently we increased streaming rate (Streaming in this context is equivalent to the sampling rate? Not clear in the discussion) for the entire network to its current rate of 100 Hz. At 100 Hz, data rate for each site is ~70-100 Mb/day.

Comment [GS4]:

~~Data from the phones are streamed using the UDP (User Datagram Protocol) protocol to a cloud-based data receiver (Supp. Mat). We prefer UDP to TCP/IP (Transmission Control Protocol/Internet Protocol) because it eliminates latency caused by two-way communications~~

~~regarding packet completeness and, similar to others [*J R Evans et al.*, 2005], we have found UDP packet loss to be negligible. The UDP receiver passes data to a MQTT (Message Queuing Telemetry Transport) broker that distributes it to processing and archiving subscribers (Figure 3,~~As of January 2021 we have collected more than 0.5 Tb of data (please see Acknowledgments, Samples, and Data section below for information on accessing the archived data).~~ As of January 2021 we have collected more than 0.5 Tb of data. The phones actively maintain NTP (Network Time Protocol) clock corrections to ensure timing accuracy (Supp. Mat) and they periodically check in to a settings server comprising a MySQL (My Structured Query Language) database that provides configurable control parameters. Additionally, a management console provides a map- and table-based general monitoring and control interface for the network.~~

2.3 ~~PGAN~~ Detection Algorithm and Alerting Strategy

Any EEW processing algorithm that operates on acceleration data can subscribe to the MQTT broker. Here, we report on a new algorithm, PGAN (for 'Peak Ground Acceleration with N vertices') that is an adaption of the PLUM method for a network of smartphones. PGAN is similar to PLUM in that it is a non-parametric ground-motion based algorithm which does not attempt to estimate anything about a detected earthquake's source, such as magnitude or epicentral location.(I believe this discussion is important to include)~~ When PLUM is run as a component of JMA's operational EEW system, however, it not only utilizes high quality seismometers, but data feeds are also monitored by JMA personnel to ensure quality control and flag noisy data. To compensate for the noisier nature of smartphone MEMS accelerometers installed in buildings susceptible to anthropogenic noise, as well as for the lack of human monitoring, PGAN requires multiple neighboring stations to experience anomalous accelerations in order to trigger an alert. The network is divided into a polygonal mesh with a configurable number of vertices greater than or equal to three. Here, we report on a quadritlateral (PGAN-4)~~Our detection algorithm is a modification of the ground-motion based Propagation of Local Undamped Motion algorithm [*Kodera et al.*, 2018]. In a polygonal mesh of station locations, we compensate for the noisier nature of smartphone accelerometers by requiring multiple neighboring stations to experience anomalous accelerations in order to trigger an alert. Here, we report on a quadritlateral mesh configuration of adjoining stations because we found an unacceptably large number of false alerts for triangular configurations.

~~In detail, PGAN-4~~ The station locations are organized into a set of unique, overlapping polygons whose sides are all less than a configurable length (currently 40 km).~~ We permit polygons to overlap.~~ Once per second PGA values are measured at each station and the values are compared with a primary threshold (currently 0.6% g). If at any station~~'s~~ PGA is above the threshold, the polygon it belongs to is marked as potentially triggered. The PGA values at each of the remaining stations in that polygon are compared with a secondary threshold (currently 0.55%g). If PGA values at all four stations in a potentially triggered polygon are above the thresholds, an alert is issued. If not, the incoming PGA values continue to be monitored for up to 15 seconds (configurable), and the polygon will trigger if the remaining station PGA values rise above the secondary PGA threshold. Otherwise, the polygon times out, and no alert is triggered. Once a polygon is triggered, an alert is sent out using ~~the~~ a cloud-based notification system (Amazon Web Services Simple Notification System~~ (AWSSNS~~). As of January 2021, alerts are only being sent to a small number of beta-subscribers in our research team.

Formatted: Font: Not Bold

Comment [GS5]:

3 Data Quality and System Latency

~~Below, we characterize the system's d~~Data quality and latency was evaluated during a six month ~~evaluation~~ period, from December 2019 to June 2020. ~~Generally, d~~During this time, daily system up-time (defined as the percentage of each day the entire network was transimitting data) was ~~reliably~~ consistengly greater than 95%. After June 2020, as Covid-19 pandemic restrictions significantly limited travel in Costa Rica, we were not able to visit stations for routine or time-dependent site visits. Accordingly, station up-time and network coverage decreased during this period to a ?% level. In the Supporting Information section we provide more details about system operation.

3.1 Data Quality

As with any seismic sensor, coupling to the ground and site-specific noise conditions control an individual phone's sensitivity. MEMS accelerometers in smart-phones have been shown to have the sensitivity to discriminate ground motions caused by earthquakes from background noise ~~[J *Evans et al.*, 2014]~~[*Evans et al.*, 2014], even in crowd-sourcing scenarios where they rest on tables [*Finazzi*, 2016; *Kong et al.*, 2016; *Finazzi and Fassò*, 2017; *Kong et al.*, 2019b]. In our fixed-network approach, phones ~~do not rest on tables nor are they placed in seismically quiet underground vaults, rather they~~ are affixed to floors, baseboards, and walls in built structures, typically homes, schools, fire departments, hospitals, or municipal buildings (Figure ~~2~~3). As such, they record accelerations through the filter of a ~~built~~ structure emplaced in locally varying soil and/or rock substrate. This type of installation means that individual stations will record earthquakes (Figure ~~4a~~5a) as well as ambient non-seismic accelerations ~~from myriad sources such~~produced by ~~as~~ road traffic, domestic motion, or thunder [*Finazzi*, 2020]. In addition to site-specific noise, we have found that some phones experience sporadic, unexplained, transient noise spikes, either on individual or multiple accelerometer components simultaneously (Figure ~~4b~~5b).

~~To construct the most appropriate data for the PGAN algorithm we~~We combine the raw acceleration data from the three orthogonal ~~MEMS~~ accelerometers into a PGA data type (the 'P' message) sampled and sent at 1 Hz ~~(the 'P', for 'PGA', message).~~. Th~~e~~ise ~~P message~~data value is the vector norm of the individually de-meaned 100 Hz acceleration values and has units of m/s^2, also expressed as a percent of gravitational acceleration, %g. To mitigate ~~the~~ transient noise spikes while permitting real seismic accelerations to pass, ~~final formation of the P message implements~~we implement a simple filter that takes the 30th percent highest value in a 1 second window (3rd highest if accelerometer sampling is 10 Hz, 30th highest if accelerometer sampling is 100 Hz; Figure ~~4b~~5b) [*Kamigaichi et al.*, 2009].

Because the ~~PGAN~~detection algorithm only requires exceedance of ground acceleration above a threshold, we characterize each site and ~~network~~ polygon (??) data quality by examining histograms of PGA data. Generally, the entire network exhibits mean ~~P-~~values of 0.25%g. Over periods of time, varying from minutes to days, individual stations may experience elevated deviations from this background behavior (Figure ~~5~~6) (Please explain units on vertical axis. Not clear if it is a histogram, as the text says). . The causes of these transient periods of elevated background noise are varied and likely related to temporally changing site-specific noise such as local construction projects, changing traffic patterns, or persistent regional storms.

Comment [GS6]:

3.2 Latency Budget

Understanding and documenting the latency budget, along with assessing alert validity, is a critical aspect of EEW system operation. We define ASTUTI system latency, $\delta t_{latency}$, as the sum of three components $\delta t_{latency} = \delta t_{data} + \delta t_{detect} + \delta t_{alert}$ where δt_{data} is the time it takes for data to be transmitted from the phones and received by the hub, δt_{detect} is the time it takes for a given processing algorithm to detect an event, and δt_{alert} is the time it takes for a message to be received by a user after it has been sent by the processing algorithm. Of these, δt_{data} and δt_{alert} are only dependent on telecommunications factors and they are independent of the specifics of a given earthquake event. For different earthquakes, δt_{detect} depends on multiple factors including magnitude, rise-time, hypocentral location, near-source network geometry, and local site acceleration response. These factors will control the time it takes for seismic waves to travel from an earthquake to be sensed by a phone, the time it takes for a site to exceed detection thresholds, and the time it takes a particular algorithm to issue a detection.

From six months of continuously-streamed data, we find that δt_{data} varies from 0.35 to 0.45 seconds depending on time of day; peak internet usage (early evening when people arrive home from work) also correlates with higher δt_{data} (Figure ~~6a~~7a). For comparison, published δt_{data} for the Italian [*Satriano et al.*, 2011], the Chinese [*Zhang et al.*, 2016], and ShakeAlert [~~*E S Cochran et al.*, 2018~~][~~*Elizabeth S*~~ *Cochran et al.*, 2018] EEW networks are 0.9, 2, and 1-3 secs, respectively. (These are surprisingly low transmission latency times. It would be useful to know how they are calculated and show some examples of it).

In order to robustly and simultaneously measure the real-time distribution of δt_{detect} and δt_{alert}, as well as to test the full operation of the ASTUTI EEW system, we used smartphones' programmable vibration feature to cause the phones to shake at expected relative arrival times for a scenario earthquake. We use the M_W 7.6 2012 Nicoya earthquake as a scenario event and program S-wave arrival time for each site calculating hypocentral distance (see methods) using a fixed move-out value for V_s of 3.2 *km/s* **(**Figure ~~6b~~7b-d**)** (Figure 7a seems out of place here and is not referenced in the text)**.** Each phone vibrated for ~10 seconds. Although they certainly do not reproduce the frequency nor amplitude content of real seismic waves interacting with built structures, smartphone vibrations are ~~decent~~ reasonable proxies for local earthquake ~~accelerations in that they~~as they exceed the accelerationthreshold values ~~we~~ used for earthquake ~~detection at specified times consistent with the earthquake rupture evolution and seismic wave propagation~~ (Figure ~~6c~~7c)**.** Of course, the model of seismic wave propagation for the event could be more sophisticated but the constant V_s value is sufficient for the purpose of estimating average values of δt_{detect} and δt_{alert}. Note here we use S-wave arrival times to be conservative, but it is possible that trigger thresholds could be exceeded in the P-wave (see System Performance below).

Over a period of 3 days, we repeated the test, vibrating the phones on the Nicoya peninsula three times an hour at the M7.6 scenario relative times, resulting in a total of 216 simulations (Figure ~~6d~~7d**)**. For this scenario, we find δt_{detect} ~~for the PGAN-4 method~~ has a mean value of ~12-13 seconds. We measured δt_{alert} by sending text message alerts to a set of 15 people with cell phones in and around the greater San Jose region**.** We find mean value for δt_{alert} is ~4 seconds (Figure ~~6d~~7d). Because this is from a small number of phones that were relatively close together, it is not

Comment [GS7]:

Formatted: Subscript

Comment [GS8]:

clear how representative this metric is of latencies that would occur when sending many thousands of alerts across a wider geographic region. For comparison we note, that in a recent test of ShakeAlert in San Diego county using the U.S. federal Wireless Emergency Alert (WEA) system ~~[*Sarah E. Minson et al.*, 2020]~~[*Minson et al.*, 2020], median δt_{alert} was ~13 seconds. To the best of our knowledge no other smartphone EEW projects have reported alerting latency data [*Finazzi*, 2016; *Kong et al.*, 2016; *Finazzi and Fassò*, 2017; *Kong et al.*, 2018; *Kong et al.*, 2019a; *Kong et al.*, 2019b; *Finazzi*, 2020; *Kong et al.*, 2020].

4 System Performance

4.1 Event Accuracy and Timeliness

We evaluate ASTUTI's performance in both detecting earthquakes and delivering alerts to people who will potentially experience shaking for a given earthquake, based on off-line playback from Dec 2019 to June 2020. ~~Because of Covid-19 related travel restrictions limiting field work and equipment maintenance, and because of algorithm tuning and development of our alert messaging algorithm, the entire system was not operational until August 2020.~~

We use Did You Feel It (DYFI) data to provide both a self-consistent metric and the most natural ground-truth for whether people in a given region experienced shaking [*Atkinson and Wald*, 2007]. ~~If, alternatively, our objective were to evaluate performance for non-human 'users' such as critical facilities, lifeline, and/or structures, then DYFI reports would be less appropriate than instrumental strong-motion measurements such as ShakeMaps [*David J Wald*, 2000].~~

During the period of assessment, ~~ASTUTI using the PGAN-4 algorithm~~we detected 5 of the 13 events (M_w 4-5.3) that were accompanied by DYFI reports of shaking ~~somewhere~~ in Costa Rica (Figure ~~7~~8, Figure S1, Table S1). The majority of these events (9 of 13) occurred outside of the network, either offshore in the MAT or in ~~within neighboring~~in Panama~~'s international borders~~ (Figure S1). Four of the detected events had thrust mechanisms and one had a strike-slip mechanism and ~~they~~ ranged in magnitude from from M_w 4.8 ~~to M_w~~ to 5.3. Only one of the detected events (2020-03-07, M_w 5.2) had an epicenter entirely within the network (Figure ~~7d~~8d, 8d). ~~Generally, the ASTUTI-detected events were accompanied by stronger shaking that was felt by much larger percentages of the population as defined by interpolated DYFI reports.~~ The detected events had a median MMI of 4.3 and a max MMI of 6 with ~17% to 73% (41% median) of the population experiencing felt shaking of at least MMI 2-3. In contrast, the non-detected events had median MMI of 2.9 and max MMI of 3.8 with 0 to 19% (0.002% median) of the population experiencing felt shaking of at least MMI 2-3.

For each detected event, in addition to the associated DYFI data we plot the estimated position of the S-wave front (assumed to be the front of peak shaking) at the time that the alert was issued (solid magenta circles in Figure ~~7~~8). Detection times, δt_{detect}, ~~for PGAN-4~~ ranged from 11-30 secs (median 22 secs, Table 1, Figure 8~~7). (In the Supplemental Material section we present results using the PGAN-3 configuration exhibiting median δt_{detect} of 16.8 seconds).~~8).. The 11-30 second range of detection times compares well with other scientific grade systems constructed and operated in similar seismic hazard settings, such as Cascadia and Mexico. For Cascadia, in 2019-2020 Shake Alert issued 4 alerts with δt_{detect} timing ranging from 8.3-13.9 seconds

[*ShakeAlertEventPage*, 2020]. Published ~~examples from the SASMEX system in Mexico show~~ δt_{detect} values from Mexico for 3 events ~~ranging~~range from 12-18 seconds [*Cuéllar et al.*, 2018] and, although not explicitly presented, we estimate from data presented in ~~the~~a recent ~~SASMEX~~ performance summary [*Cuéllar et al.*, 2017] that minimum δt_{detect} is ~ 8 seconds. However, in the Mexican system there is essentially no data transmission nor data distribution delay, as the alert is defined locally by the sensing stations and sent to users via radio links.

For each detected event we plot record sections ~~for each station's recorded~~of P messages and of the expected range of P- and S-wave velocities of 6 to 3 km/s, respectively (Figure ~~8~~9). We find that 2 of the 5 events (2019-12-08 and 2020-03-07, Figure ~~8a~~9a,b and Figure ~~8g~~9g,h) triggered on the P wave, ~~which can be seen~~as shown by the magenta vertical line ~~occurring~~ on the left side of the record section plots. For each of these events, all four of the triggering ~~PGAN~~ stations exceeded the threshold (0.6 g for the first and 0.55 %g for subsequent stations;shown in dashed green lines on the right panel ~~in~~ of each figure). The three remaining events triggered with the S-wave. For all ~~of the~~ events, except ~~aside from that on~~on the one on 2020-03-07, ~ 10 seconds elapsed between the first and fourth station detections, surpassing the detection threshold. For the 2020-03-07 event, only 2 seconds elapsed, because the 2020-03-07 event was entirely within the network ~~and so spreading seismic waves from the sub-surface event arrived much more synchronously at the sensors than for events with offshore hypocenters~~.

Formatted: Not Highlight

Formatted: Not Highlight

4.2 Shaking

In addition to reporting the detection times we also assess the likely warning outcomes for population percentages ~~for the case where the warning precinct for a detected event is the entire Costa Rican population (Figure 7). We use the WorldPop 1 km gridded database~~ [*Tatem*, 2017] for ~~our population data.~~the case where the warning precinct for a detected event is the entire Costa Rican population (Figure 8). We interpolate the spatial distribution of DYFI reports to define an area of felt shaking, reasoning that it is likely ground-shaking occurred between sites of reported shaking [*Kodera et al.*, 2018; ~~*E*~~~~*Elizabeth S*~~ *Cochran et al.*, 2019]. Although DYFI data contains uncertainty from humans' perceptional subjectivity [*Goltz et al.*, 2020], we minimize this by applying a binary "shaking" or "no-shaking" classification to DYFI reports. We contend this is a conservative approach in that, for a given earthquake, the area of felt shaking represented by DFYI data is an under- rather than over-estimation: it seems more likely that people who felt shaking would not report it rather than that people would report felt shaking when none was felt. We categorize outcomes as True Positive-Shaking (TP-S): the system correctly detects an event and a user receives an alert prior to felt shaking; True Positive-No Shaking (TP-NS): the system correctly detects an event, a user receives an alert but does not feel shaking; No Alert (NA): the system correctly detects an event, but a user does not receive an alert prior to shaking; False Alert (FA): the system incorrectly detects an event and sends an alert, no shaking occurs anywhere; and Missed Alert (MA): a felt event occurs but the system does not issue an alert. For each of these scenarios, to permit conservative evaluation of when users could expect to receive warning messages, in the plots we add to δt_{detect} 5 seconds, ~1 second more than the median value of δt_{alert} from the vibration tests (see section 3.2).

Formatted: Font: Italic, Do not check spelling or grammar

~~For PGAN-4~~Over the 6 months of evaluation, 15.5-71.0% (median 38.5%) of the population would have received TP-S outcomes, 27.1-83% (median 52.8 %) would have TP-NS outcomes,

Formatted: Not Highlight

and 4.9-8.8% (median 2.5%) would have NA outcomes (Table 1~~, in Table S2 we display metrics for PGAN-3 configuration).~~). Outcomes can be grouped further according to onshore (in-network) and offshore (out-of-network) source categories. For the onshore events (12/8/19, 3/7/20, and 3/13/20) more time after warning is received is generally available because δt_{detect} is, on average, 10 seconds faster than the two offshore events (1/21/20a,b) and because the events on either the Nicoya or Osa peninsulas are far enough away from population concentrations (Figure ~~7a~~8a,d,e). Interestingly, however, the two off-shore events yield the largest TP-S outcomes because Costa Rica's population is concentrated in and around the capital city, San Jose, near the center of the country and more than 100 km from these events (Table 1).

4.3 Missed & False Alerts

In order to analyze a self-consistent dataset of felt shaking, we further classify MAs to be events for which the system did not produce an alert but that produced enough shaking for people to file DYFI reports. This is a minimum number as there may be other, smaller events that were felt but for which no DYFI reports were filed. During the period of evaluation there were 8 such MAs with a range of M_w 4-5 (Table S1, Figure S1). Six of these MAs were located well outside of the network (Figure S1). Of the remaining two, only one (12-31-2020, M_w 4) was located entirely within the network, the other (4-17-2020, M_w 4) was located on the Pacific coast at the edge of the network. For these events, less than 0.5% of the population reported shaking and the reported shaking never exceeded MMI III (Figure S1).

False alerts (FAs) can occur because of system glitches or non-seismic local or regional accelerations. From our time period of analysis we identified two types of ~~system glitch~~ sources that affected three or more neighboring stations nearly simulateously. The first is unexpected phone vibration because of user error when programming the vibration tests. The second is regional power-grid instability causing AC electricity fluctuations that, in turn, cause multiple phones to unexpectedly vibrate simultaneously when they re-start as the regional AC power cycles off and on. Once they are identified, these are relatively straightforward FA sources to mitigate in real-time by better accounting of programmed vibrations and by identifying and excluding sites experiencing regional power cycling. ~~Real, non-earthquake accelerations from sources such as vehicles passing on the street, thunder shaking a structure, or people moving furniture near the phone installation are more difficult to mitigate. As the number of vertices in the PGAN triggering criteria is increased, the probability of regionally simultaneous non-seismic accelerations decreases. In our early tests we found that it was necessary to include more than three stations in the PGAN detection algorithm to mitigate the effect of these types of unwanted accelerations on alert generation.~~ (In my opinion this discussion is unnecessary because you require at least four sensors in the polygon to report exceedance of the PGA. Traffic, human motion or the other effects you mention would not take place simultaneously)

Comment [GS9]:

~~In our analysis of six months of ASTUTI data using PGAN-4~~During the evaluation period, once system glitches were identified and noisy stations removed, we eliminated all FAs yielding a FA rate of 0%. This is similar performance to the result from offline runs ~~of the PLUM algorithm~~ on West Coast U.S. data ~~that had a 0% FA rate~~of the ground-motion based algorithm from which ours is derived (Cochran, *et al.* 2019). For further comparison, published FA rates for science-grade EEW networks range from 2.5% in Taiwan [*Xu et al.*, 2017] to 8% for Shake Alert [*Kohler* *et al.*, 2020]. The Japanese system reported many FAs after the 2011 *M* 9 Tohoku earthquake but to the best our knowledge has not published a FA rate [*Hoshiba*, 2014]. For smartphone EEW, the Earthquake Network Project sets detection thresholds so that they do not exceed 1 FA per year, per country [*Finazzi*, 2020]. This approach is a way to balance the often competing objectives of minimizing both MAs and FAs. The MyShake project has not reported FA rate, stating that "the false positive events are not so well quantified, because we do not have a large number of false positive samples from the system" [*Kong et al.*, 2020].

5 Discussion

Our investigation demonstrates for the first time that off-the-shelf smartphones deployed in a fixed network can reliably and effectively provide operational EEW for a country with Costa Rica's size and population. Performance metrics for the ASTUTI network compare favorably with scientific-grade EEW, at least during the six months of evaluation. Data latency is lower than some science grade networks and detection latency is ~~similar albeit~~ slightly higher due to the ground-motion based detection method triggering usually on S-waves. The spatial distribution of earthquakes that occurred during the evaluation period appears to be representative of the spatial distribution of the earthquakes expected to do the most damage in Costa Rica (with the exception of the southern Caribbean region) and so it is likely that similar ASTUTI performance could be expected over longer time intervals.

There is a marked cost differential between the fixed network smartphone approach and science grade networks. Capital costs for the entire ASTUTI network were ~USD 22,000 and annual operating and maintenance costs are ~USD 20,000~~ (not ~ 1-2 full-time equivalent salaries)~~. For comparison, ShakeAlert capital costs are close to USD 100 million and annual operating costs are ~ USD 39 million [*Given et al.*, 2018]. When normalized by population for Costa Rica (~5 M) and combined populations of California, Oregon, and Washington (~52 M), ASTUTI's capital costs are 450 times less and annual operating costs are 192 times less than Shake Alert's. When normalized by area for the same regions (~51,100 km^2 and ~863,000 km^2, respectively), ASTUTI's capital costs are 279 times less and annual operating costs are 119 times less than Shake Alert's. ~~Despite this roughly two order of magnitude difference in capital and annual costs, we find, generally, that performance metrics for our fixed-network smartphone EEW approach compare favorably with scientific-grade EEW, at least during our six months of evaluation. Data latency, δt_{data}, for the ASTUTI network is lower than some science grade networks. Alerting latency, δt_{alert}, is the same regardless of the type of sensing network. Detection latency, δt_{detect}, is similar, albeit slightly higher.~~ Increasing station density by ~ 30% such that four stations would cover the area that three stations currently cover (at a cost of less than ~USD 10,000) ~~would~~could probably decrease δt_{detect} by 3-5 seconds~~. Futhermore, the spatial distribution of earthquakes that occurred during the first six months of ASTUTI network operation appears to be representative of the spatial distribution of the earthquakes expected to do the most damage in Costa Rica (with the exception of the southern Caribbean region) and so it is likely that similar ASTUTI performance could be expected over longer time intervals~~.

~~Faster detection times for EEW systems using higher-grade sensors is due primarily to P-wave based detection [*Cuéllar et al.*, 2018; *Chung et al.*, 2019]. We identify P-waves in the acceleration records of individual stations for all 5 detected events (Figure 8) and, notably, two~~

~~of the five events were detected as the result of triggers in the P-waves. Accordingly, although ground-motion-based algorithms may be more likely to trigger on S-waves for moderate-sized events, there is no *a priori* requirement of this. Faster warnings could occur for larger events, such as the M_w 5.2 2020-03-07 event where P arrivals triggered the event and are seen in the records of many stations (Figure 8d). Additionally, as future generation MEMs accelerometers become more sensitive, detection on the P-wave could become more likely, although trigger thresholds have to be carefully considered given the noise levels at station locations. A better understanding of the relationship between expected P-wave amplitudes and their spatial coherence for a variety of events would allow for improved predictions of the performance of accelerometer-based EEW.~~

The ASTUTI strategy of warning the entire country upon detection of any event would permit warnings to reach large percentages of the population (15-70% TP outcome) for ~~the~~ earthquakes that ~~"mattered", that is, earthquakes~~ that caused enough shaking ~~for people to submit DYFI reports~~. The detected events had median and maximum MMI levels of 4.3 and 6 with ~17% to 73% (41% median) of the population experiencing felt shaking. In contrast, the non-detected events had median and max MMI levels of 2.9 and 3.8 with 0 to 19% (0.002% median) of the population experiencing felt shaking. For smaller and sufficiently out-of-network events that did not trigger alerts, less than 0.5% of the population reported shaking and the reported shaking never exceeded MMI ~~III~~3. These population outcomes are a function of the spatial distribution of event locations, station density, and population spatial distribution. For example, for events that nucleate on the Nicoya or Osa peninsulas, the detecting stations are far enough away from population concentrations to afford more people time to receive warnings and undertake mitigating actions. Consistent with the results of the vibration test for the 2012 M7.6 Nicoya earthquake scenario, this suggests that the approach could work well for the case of subduction zone events of significant size (>M7). ~~For events in the central Pacific coastal embayment between the Nicoya and Osa peninsulas, however, the events (2020-01-21a,b) occur farther away from the coastal sensors, and closer to higher population centers, leading to shorter detection and warning times. This is an unavoidable consequence of Costa Rica's geography and is common to any EEW system reliant on land-based sensors.~~

~~For outcomes where an alert is received and no shaking experienced, we believe it is important to differentiate between event-based FA and shaking-based TP-NS outcomes. The ASTUTI~~ASTUTI's 0% FA rate is lower than all other EEW systems' reported FA rates aside from retrospective testing of ground-motin based detection with West Coast U.S. data ~~[*E S Cochran et al.*, 2019].~~[~~*Elizabeth S*~~ *Cochran et al.*, 2019]. Our examination was only over six months, however, and so more operational time is required for a longer-term, more representative FA rate. As expected, a low detection threshold criterion combined with a country-wide alerting region also leads to significant population percentages (20-80%) ~~with~~ where an alert was received and no was shaking experienced (TP-NS outcomes~~.~~). TP-NS percentage could be reduced by modifying our alerting philosophy to include shaking estimation from source-parameter characterization; however, we suggest that the benefit of attempting a more refined warning may be outweighed by the added uncertainty associated with EEW parameter estimation and ground-motion prediction ~~[*Sarah E. Minson et al.*, 2018; *S.E. Minson et al.*, 2019]~~[*Minson et al.*, 2018; *Minson et al.*, 2019]. It is not clear what penalty there may be in terms of user engagement if an EEW system provides alerts without felt shaking, especially if users were to receive a rapid follow-on message stating that an event had been correctly detected although they did not feel shaking. In fact, the general population may appreciate receiving an alert whenever an earthquake occured, even if no one felt it [*Nakayachi et al.*, 2019]. Indeed, we are at the very early stages of studying the nuanced relationships between EEW system performance and human sentiment. Given this evolving understanding, a clear high-priority must be EEW pre- and post-event education and messaging [*S McBride et al.*, 2020].

Given the concentration of Costa Rica's population in San Jose (> 90%), it is useful to ~~further consider outcomes~~assess DCHO feasibility for ~~its~~the greater metro region. S-wave arrival in San Jose for the three onshore or near-shore events (2019-12-08, 2020-03-07, 2020-03-13) was between ~30 and 50 seconds after median receipt of alert time (δt_{detect} + 5 seconds) (Figure ~~7a~~8a,d,e) . For the offshore events (2020-01-21a, 2020-01-21b) the S-wave arrival was much sooner, a few seconds after median receipt of alert time (Figure ~~7b~~8b,c). Thus, for three of the five detected events, our results suggest that DCHO for a large percentage of San Jose residents is an achievable objective. For the two offshore events DCHO may not be a widely achieveable objective. Because there is a paucity of studies that have directly evaluated DCHO, however, there is some uncertainty in assessing EEW users' likely response. For instance, for the HayWired exercise [*Porter and Jones*, 2018], DCHO took 5-15 seconds and for a New Zealand ShakeOut exercise DCHO timing varied from 10-30 seconds (64% of participants took less than 10 seconds and 34% took 11-30 secs) [*S K McBride et al.*, 2019]. For two actual shaking events, however, *Nakayashi et al* (2019) report that taking no action was the most common outcome for people who received EEW alerts. This uncertainty highlights the often counter-intuitive nature of humans' interaction with automated warning systems and stresses the importance of further mitigation education and social science studies targeting DCHO activities.

~~Our investigation began by questioning whether off-the-shelf smartphones deployed in a fixed network could provide EEW in a reliable and effective enough manner so that significant populations could receive warnings in time for DCHO protective actions to be undertaken. The answer is unequivocally affirmative.~~ Faster detection times for EEW systems using higher-grade sensors is due primarily to P-wave based detection [*Cuéllar et al.*, 2018; *Chung et al.*, 2019]. We identify P-waves in the acceleration records of individual stations for all 5 detected events (Figure 9) and, notably, two of the five events were detected as the result of triggers in the P-waves. Accordingly, although ground-motion-based algorithms may be more likely to trigger on S-waves for moderate-sized events, there is no *a priori* requirement of this. Faster warnings could occur for larger events, such as the M_w 5.2 2020-03-07 event where P arrivals triggered the event and are seen in the records of many stations (Figure 9d). Additionally, as future generation MEMs accelerometers become more sensitive, detection on the P wave could become more likely, although trigger thresholds have to be carefully considered given the noise levels at station locations. A better understanding of the relationship between expected P wave amplitudes and their spatial coherence for a variety of events would allow for improved predictions of the performance of accelerometer-based EEW.

As pointed out by others [~~*David J*~~ *Wald*, 2020b], ~~however,~~ updating construction codes and practices as well as promoting more widespread and thorough acceptance of simple mitigation actions such as DCHO deserve to be considered ~~as~~ top priorities for earthquake-related resource allocation. The ASTUTI fixed network EEW approach ~~we describe here~~ is so inexpensive, however, that it need not interfere with other budgetary priorities for ~~resource-limited~~ countries that do not have the resources to maintain advanced seismic networks but that could benefit from some form of EEW. Myriad implementation scenarios ~~can be envisioned ranging from stand-alone smartphone networks to augmentations of preexisiting scientific-grade instrumentation. Basic stand-up for a network such as ASTUTI requires purchasing and installation of phones and SIM cards and spinning up instances of the processing algorithms on a cloud-based server. Similarly, the approach is so inexpensive that the benefits of a fixed-network smartphone approach (station coverage for expected earthquake source regions and network reliability) outweigh the even lower costs associated with crowd-sourcing.~~ could be envisioned ranging from stand-alone smartphone networks to hybrid augmentations of preexisiting scientific-grade instrumentation. Futhermore, the approach is so inexpensive that the sensing and detection benefits of judiciously placed sensors outweighs the even lower costs associated with crowd-sourcing. We note that in no way do we discount crowd-sourcing's value, especially the potential contribution to ground-motion and structural seismic response studies of massive numbers of crowd-source acceleration records [*Kong et al.*, 2018]. Moreover, as digital home assistants become more ubiquitous, crowd-sourcing efforts (users permitting access to their personal devices) will move to fixed network modes. Especially for resource-limited populations, however, the day is still years away when user-contributed data will be spatially complete enough and temporally reliable enough to provide robust EEW capability for a larger population. With smartphone-based EEW in a nascent phase of rapid growth, we stress the need for continued scientific investigation and validation of these methodologies by the global community. Especially given the need for transparency when issues of public safety are concerned, we encourage open-access data policies for smartphone-based EEW systems.

6 Conclusions

We have demonstrated that a fixed-network of smartphones using a ground-motion-based detection algorithm can provide operational EEW at a cost that is generally 1-2 orders of magnitude less than scientific-grade networks. The ASTUTI combination of detection capability with a country-wide alerting strategy demonstrates an effective EEW strategy for the common earthquake hazard setting where large population centers are close enough to be affected by large earthquakes but far enough away from them so that alerts may arrive in time to permit simple protective actions such as DCHO. More social science research is needed to ascertain whether populations will be tolerant of a system that correctly detects events and that sends alerts to users who do not necessarily feel shaking.

Acknowledgments, Samples, and Data

We gratefully acknowledge USAID Office of Foreign Disaster Assistance for funding and especially Gari Mayberry for her oversight of this work since its inception. We thank Jeff McGuire, Annemarie Baltay, and Evelyn Roeloffs for very helpful discussions and earlier reviews of this manuscript. We thank the journal reviewers for their comments and suggestions. Data supporting the conclusions in this paper may be obtained at: http://ds.iris.edu/ds/nodes/dmc/forms/assembled-data/. Assembled data set number: 21-004; short name: ASTUTI.

References

Allen, R. M., E. S. Cochran, T. J. Huggins, S. Miles, and D. Otegui (2018), Lessons from Mexico's earthquake early warning system, *Eos, Earth and Space Science News*, *99*.

Allen, R. M., and D. Melgar (2019), Earthquake early warning: Advances, scientific challenges, and societal needs, *Annual Review of Earth and Planetary Sciences*, *47*, 361-388.

~~Anthony, R. E., A. T. Ringler, D. C. Wilson, and E. Wolin (2019), Do low‐cost seismographs perform well enough for your network? An overview of laboratory tests and field observations of the OSOP Raspberry Shake 4D, *Seismological Research Letters*, *90*(1), 219-228.~~

Atkinson, G. M., and D. J. Wald (2007), "Did You Feel It?" intensity data: A surprisingly good measure of earthquake ground motion, *Seismological Research Letters*, *78*(3), 362-368.

Becker, J. S., S. H. Potter, L. J. Vinnell, K. Nakayachi, S. K. McBride, and D. M. Johnston (2020), Earthquake early warning in Aotearoa New Zealand: a survey of public perspectives to guide warning system development, *Humanities and Social Sciences Communications*, *7*(1), 1-12.

~~Böse, M., T. H. Heaton, and E. Hauksson (2012), Real-time finite fault rupture detector (FinDer) for large earthquakes, *Geophysical Journal International*, *191*(2), 803-812.~~

Chung, A. I., I. Henson, and R. M. Allen (2019), Optimizing earthquake early warning performance: ElarmS‐3, *Seismological Research Letters*, *90*(2A), 727-743.

Cochran, E~~., J. Lawrence, C. Christensen, and R. Jakka (2009), The Quake-Catcher Network: Citizen Science Expanding Seismic Horizons, *Seismological Research Letters*, *80*(1), 26~~. S. (2018), To catch a quake, *Nat Commun*, *9*(1), 2508, doi:10.1038/s41467-018-04790-9.

Cochran, E. S., J. Bunn, S. E. Minson, A. S. Baltay, D. L. Kilb, Y. Kodera, and M. Hoshiba (2019), Event Detection Performance of the PLUM Earthquake Early Warning Algorithm in Southern California, *Bulletin of the Seismological Society of America*, *109*(4), 1524-1541.

Cochran, E. S., M. D. Kohler, D. D. Given, S. Guiwits, J. Andrews, M. A. Meier, M. Ahmad, I. Henson, R. Hartog, and D. Smith (2018), Earthquake early warning ShakeAlert system: Testing and certification platform, *Seismological Research Letters*, *89*(1), 108-117.

Cuéllar, A., G. Suárez, and J. M. Espinosa‐Aranda (2017), Performance Evaluation of the Earthquake Detection and Classification Algorithm 2(tS‐tP) of the Seismic Alert System of Mexico (SASMEX), *Bulletin of the Seismological Society of America*, *107*(3), 1451-1463, doi:10.1785/0120150330.

Cuéllar, A., G. Suárez, and J. M. Espinosa‐Aranda (2018), A Fast Earthquake Early Warning Algorithm Based on the First 3 s of the P‐Wave Coda, *Bulletin of the Seismological Society of America*, *108*(4), 2068-2079, doi:10.1785/0120180079.

DeMets, C., R. G. Gordon, and D. F. Argus (2010), Geologically current plate motions, *Geophysical Journal International*, *181*(1), 1-80.

Espinosa-Aranda, J. M., A. Cuellar, A. Garcia, G. Ibarrola, R. Islas, S. Maldonado, and F. H. Rodriguez (2009), Evolution of the Mexican seismic alert system (SASMEX), *Seismological Research Letters*, *80*(5), 694-706 %@ 1938-2057.

Evans, J., R. M. Allen, A. Chung, E. Cochran, R. Guy, M. Hellweg, and J. Lawrence (2014), Performance of several low‐cost accelerometers, *Seismological Research Letters*, *85*(1), 147-158.

~~Evans, J. R., R. H. Hamstra Jr, C. Kündig, P. Camina, and J. A. Rogers (2005), TREMOR: A wireless MEMS accelerograph for dense arrays, *Earthquake spectra*, *21*(1), 91-124.~~

Finazzi, F. (2016), The earthquake network project: Toward a crowdsourced smartphone‐based earthquake early warning system, *Bulletin of the Seismological Society of America*, *106*(3), 1088-1099.

Finazzi, F. (2020), The Earthquake Network Project: A Platform for Earthquake Early Warning, Rapid Impact Assessment, and Search and Rescue, *Frontiers in Earth Science*, *8*, 243.

Finazzi, F., and A. Fassò (2017), A statistical approach to crowdsourced smartphone-based earthquake early warning systems, *Stochastic Environmental Research and Risk Assessment*, *31*(7), 1649-1658.

~~Fujinawa, Y., and Y. Noda (2013), Japan's earthquake early warning system on 11 March 2011: performance, shortcomings, and changes, *Earthquake Spectra*, *29*(S1), S341-S368.~~

~~Fujinawa, Y., K. Takahashi, Y. Noda, H. Iitaka, and S. Yazaki (2011), Remote detection of the electric field change induced at the seismic wave front from the start of fault rupturing, *Int. J. Geophys.*, *752193*, doi:10.1155/2011/752193.~~

Given, D. D., R. M. Allen, A. S. Baltay, P. Bodin, E. S. Cochran, K. Creager, R. M. de Groot, L. S. Gee, E. Hauksson, and T. H. Heaton (2018), Revised technical implementation plan for the ShakeAlert system—An earthquake early warning system for the West Coast of the United States*Rep. 2331-1258*, US Geological Survey.

Goldberg, D., D. Melgar, and Y. Bock (2019), Seismogeodetic P‐wave amplitude: No evidence for strong determinism, *Geophysical research letters*, *46*(20), 11118-11126.

Goltz, J. D., H. Park, G. Nakano, and K. Yamori (2020), Earthquake ground motion and human behavior: Using DYFI data to assess behavioral response to earthquakes, *Earthquake Spectra*, 8755293019899958.

Gregor, N., N. A. Abrahamson, G. M. Atkinson, D. M. Boore, Y. Bozorgnia, K. W. Campbell, B. S.-J. Chiou, I. Idriss, R. Kamai, and E. Seyhan (2014), Comparison of NGA-West2 GMPEs, *Earthquake Spectra*, *30*(3), 1179-1197.

Heaton, T. (1985), A Model for a Seismic Computerized Alert Network, *Science*, *228*, 987-990.

Hoshiba, M. (2014), Review of the nationwide earthquake early warning in Japan during its first five years, in *Earthquake Hazard, Risk and Disasters*, edited, pp. 505-529, Elsevier.

Hoshiba, M., K. Ohtake, K. Iwakiri, T. Aketagawa, H. Nakamura, and S. Yamamoto (2010), How precisely can we anticipate seismic intensities? A study of uncertainty of anticipated seismic intensities for the Earthquake Early Warning method in Japan, *Earth, planets and space*, *62*(8), 611-620.

Kamigaichi, O., M. Saito, K. Doi, T. Matsumori, S. y. Tsukada, K. Takeda, T. Shimoyama, K. Nakamura, M. Kiyomoto, and Y. Watanabe (2009), Earthquake early warning in Japan: Warning the general public and future prospects, *Seismological Research Letters*, *80*(5), 717-726.

~~Ketelaar, P. E., and M. Van Balen (2018), The smartphone as your follower: The role of smartphone literacy in the relation between privacy concerns, attitude and behaviour towards phone-embedded tracking, *Computers in Human Behavior*, *78*, 174-182.~~

Kobayashi, D., P. LaFemina, H. Geirsson, E. Chichaco, A. A. Abrego, H. Mora, and E. Camacho (2014), Kinematics of the western Caribbean: Collision of the Cocos Ridge and upper plate deformation, *Geochemistry, Geophysics, Geosystems*, *15*, 1671-1683.

Kodera, Y., Y. Yamada, K. Hirano, K. Tamaribuchi, S. Adachi, N. Hayashimoto, M. Morimoto, M. Nakamura, and M. Hoshiba (2018), The propagation of local undamped motion (PLUM) method: A simple and robust seismic wavefield estimation approach for earthquake early warning, *Bulletin of the Seismological Society of America*, *108*(2), 983-1003.

Kohler, M. D., D. E. Smith, J. Andrews, A. I. Chung, R. Hartog, I. Henson, D. D. Given, R. de Groot, and S. Guiwits (2020), Earthquake Early Warning ShakeAlert 2.0: Public Rollout, *Seismological Research Letters*, *91*(3), 1763-1775.

Kong, Q., R. M. Allen, M. D. Kohler, T. H. Heaton, and J. Bunn (2018), Structural Health Monitoring of Buildings Using Smartphone Sensors, *Seismological Research Letters*, *89*(2A), 594-602, doi:10.1785/0220170111.

Kong, Q., R. M. Allen, L. Schreier, and Y.-W. Kwon (2016), MyShake: A smartphone seismic network for earthquake early warning and beyond, *Science advances*, *2*(2), e1501055.

Kong, Q., A. Inbal, R. M. Allen, Q. Lv, and A. Puder (2019a), Machine learning aspects of the MyShake global smartphone seismic network, *Seismological Research Letters*, *90*(2A), 546-552.

Kong, Q., R. Martin‐Short, and R. M. Allen (2020), Toward Global Earthquake Early Warning with the MyShake Smartphone Seismic Network, Part 1: Simulation Platform and Detection Algorithm, *Seismological Research Letters*.

Kong, Q., S. Patel, A. Inbal, and R. M. Allen (2019b), Assessing the sensitivity and accuracy of the MyShake smartphone seismic network to detect and characterize earthquakes, *Seismological Research Letters*, *90*(5), 1937-1949.

~~Lawrence, J. F., E. S. Cochran, A. Chung, A. Kaiser, C. M. Christensen, R. Allen, J. W. Baker, and e. al. (2014), Rapid earthquake characterization using MEMS accelerometers and volunteer hosts following the M 7.2 Darfield, New Zealand, earthquake, *Bull. Seism. Soc. Am.*, *104*(1), 184-192.~~

McBride, S., A. Bostrom, J. Sutton, R. M. de Groot, A. S. Baltay, B. Terbush, P. Bodin, M. Dixon, E. Holland, and R. Arba (2020), Developing post-alert messaging for ShakeAlert, the earthquake early warning system for the west coast of the United States of America, *International Journal of Disaster Risk Reduction*, 101713.

McBride, S. K., J. S. Becker, and D. M. Johnston (2019), Exploring the barriers for people taking protective actions during the 2012 and 2015 New Zealand ShakeOut drills, *International journal of disaster risk reduction*, *37*, 101150.

Meier, M.-A., J. Ampuero, and T. H. Heaton (2017), The hidden simplicity of subduction megathrust earthquakes, *Science*, *357*(6357), 1277-1281.

Meier, M. A. (2017), How "good" are real - time ground motion predictions from earthquake early warning systems?, *Journal of Geophysical Research: Solid Earth*, *122*(7), 5561-5577.

~~Meier, M. A., Y. Kodera, M. Böse, A. Chung, M. Hoshiba, E. Cochran, S. Minson, E. Hauksson, and T. Heaton (2020), How Often Can Earthquake Early Warning Systems Alert Sites With High - Intensity Ground Motion?, *Journal of Geophysical Research: Solid Earth*, *125*(2), e2019JB017718.~~

Melgar, D., and G. P. Hayes (2019), Characterizing large earthquakes before rupture is complete, *Science Advances*, *5*(5), eaav2032.

Minson, S. E., A. S. Baltay, E. S. Cochran, T. C. Hanks, M. T. Page, S. K. McBride, K. R. Milner, and M.-A. Meier (2019), the Limits of earthquake early Warning Accuracy and Best Alerting strategy, *Scientific reports*, *9*(1), 2478 %@ 2045-2322.

Minson, S. E., B. A. Brooks, C. L. Glennie, J. R. Murray, J. O. Langbein, S. E. Owen, R. A. Iannucci, and D. L. Hauser (2015), Crowd-Sourced Global Earthquake Early Warning, *Science Advances*.

Minson, S. E., M.-A. Meier, A. S. Baltay, T. C. Hanks, and E. S. Cochran (2018), The limits of earthquake early warning: Timeliness of ground motion estimates, *Science advances*, *4*(3), eaaq0504 %@ 2375-2548.

Minson, S. E., J. ~~R. Murray, J. O. Langbein, and J. S. Gomberg (2014), Real-time inversions for finite fault slip models and rupture geometry based on high-rate GPS data, *Journal of Geophysical Research: Solid Earth*, *119*(4), doi:10.1002/2013JB01062.~~

~~Minson, S. E., J.~~ K. Saunders, J. J. Bunn, E. S. Cochran, A. S. Baltay, D. L. Kilb, M. Hoshiba, and Y. Kodera (2020), Real-Time Performance of the PLUM Earthquake Early Warning Method during the 2019 M 6.4 and 7.1 Ridgecrest, California, Earthquakes, *Bulletin of the Seismological Society of America*, *110*(4), 1887-1903, doi:10.1785/0120200021.

Nakayachi, K., J. S. Becker, S. H. Potter, and M. Dixon (2019), Residents' Reactions to Earthquake Early Warnings in Japan, *Risk Analysis*, *0*(0 %@ 0272-4332), doi:10.1111/risa.13306 %U https://onlinelibrary.wiley.com/doi/abs/10.1111/risa.13306.

Olson, E. L., and R. M. Allen (2005), The deterministic nature of earthquake rupture, *Nature*, *438*(7065), 212-215.

Pagani, M., J. Garcia-Pelaez, R. Gee, K. Johnson, V. Poggi, V. Silva, M. Simionato, R. Styron, D. Viganò, and L. Danciu (2020), The 2018 version of the global earthquake model: hazard component, *Earthquake Spectra*, *36*(1_suppl), 226-251.

Porter, K., and J. Jones (2018), How many injuries can be avoided in the HayWired Scenario through earthquake early warning and drop, cover, and hold on, *The HayWired Earthquake Scenario: US Geological Survey Scientific Investigations Report—2017-5013. Reston, VA: USGS*, 401-429.

~~Press, A. (2020), https://www.usnews.com/news/best-states/california/articles/2020-08-11/california-quake-alerts-to-be-standard-on-android-phones, edited.~~

Protti, M., V. González, A. V. Newman, T. H. Dixon, S. Y. Schwartz, J. S. Marshall, L. Feng, J. I. Walter, R. Malservisi, and S. E. Owen (2014), Nicoya earthquake rupture anticipated by geodetic measurement of the locked plate interface, *Nature Geoscience*, *7*(2), 117-121.

Roulston, M. S., and L. A. Smith (2004), The boy who cried wolf revisited: The impact of false alarm intolerance on cost–loss scenarios, *Weather and Forecasting*, *19*(2), 391-397.

Satriano, C., L. Elia, C. Martino, M. Lancieri, A. Zollo, and G. Iannaccone (2011), PRESTo, the earthquake early warning system for southern Italy: Concepts, capabilities and future perspectives, *Soil Dynamics and Earthquake Engineering*, *31*(2), 137-153.

Saunders, J. K., B. T. Aagaard, A. S. Baltay, and S. E. Minson (2020), Optimizing Earthquake Early Warning Alert Distance Strategies Using the July 2019 Mw 6.4 and Mw 7.1 Ridgecrest, California, Earthquakes, *Bulletin of the Seismological Society of America*, *110*(4), 1872-1886, doi:10.1785/0120200022.

ShakeAlertEventPage (2020), edited, https://earthquake.usgs.gov/earthquakes/eventpage/us60007f42/executive; https://earthquake.usgs.gov/earthquakes/eventpage/uw61562126/executive; https://earthquake.usgs.gov/earthquakes/eventpage/nc73355700/executive; https://earthquake.usgs.gov/earthquakes/eventpage/nc73344735/executive.

Stogaitis, M., et al. (2020), Earthquakes at Google, *presented at 2020 Fall Meeting, American Geophysical Union, Abstract S044-08*.

~~Strauss, J. A., and R. M. Allen (2016), Benefits and costs of earthquake early warning, *Seismological Research Letters*, *87*(3), 765-772.~~

Suárez, G., J. M. Espinosa‐Aranda, A. Cuéllar, G. Ibarrola, A. García, M. Zavala, S. Maldonado, and R. Islas (2018), A dedicated seismic early warning network: The Mexican Seismic Alert System (SASMEX), *Seismological Research Letters*, *89*(2A), 382-391 %@ 0895-0695.
~~Tamaribuchi, K., M. Yamada, and S. Wu (2014), A new approach to identify multiple concurrent events for improvement of earthquake early warning, *Zisin*, *67*(2).~~
Tatem, A. J. (2017), WorldPop, open data for spatial demography, *Scientific data*, *4*(1), 1-4.
~~Thakoor, K., J. Andrews, E. Hauksson, and T. Heaton (2019), From earthquake source parameters to ground‐motion warnings near you: The ShakeAlert earthquake information to ground‐motion (eqInfo2GM) method, *Seismological Research Letters*, *90*(3), 1243-1257.~~
Trugman, D. T., M. T. Page, S. E. Minson, and E. S. Cochran (2019), Peak ground displacement saturates exactly when expected: Implications for earthquake early warning, *Journal of Geophysical Research: Solid Earth*, *124*(5), 4642-4653.
Wald, D. J. (2000), *" ShakeMaps": Instant Maps of Earthquake Shaking*, US Department of the Interior, US Geological Survey.
Wald, D. J. (2020a), Practical limitations of earthquake early warning, *Earthquake Spectra*, *36*(3), 1412-1447, doi:10.1177/8755293020911388.
Wald, D. J. (2020b), Practical limitations of earthquake early warning, *Earthquake Spectra*, 8755293020911388.
Xu, Y., J. P. Wang, Y.-M. Wu, and H. Kuo-Chen (2017), Reliability assessment on earthquake early warning: A case study from Taiwan, *Soil Dynamics and Earthquake Engineering*, *92*, 397-407, doi:10.1016/j.soildyn.2016.10.015.
Zhang, H., X. Jin, Y. Wei, J. Li, L. Kang, S. Wang, L. Huang, and P. Yu (2016), An earthquake early warning system in Fujian, China, *Bulletin of the Seismological Society of America*, *106*(2), 755-765.
~~1.~~

1. ~~**Figure 1**.~~

Figure 1. Cartoon depicting fixed-network **(a)** and crowd-sourced **(b)** EEW warning times for two population centers for a hypothetical earthquake (red star). The red circle indicates the time when the event is detected. With a fixed-network, the event is detected by nearby sensors located in zones of sparse population and both population centers could receive warning. With crowd-sourcing, the first population center's smartphones are used as the detectors and so it does not receive a warning. Moreover, the warning time for the 2nd population center is reduced from the what the fixed-network could have provided.

Figure 2. Costa Rica seismotectonic framework, population distribution, ASTUTI network, and the M7.6 2012 Nicoya peninsula earthquake. Basemap, WorldMap 1 km gridded population (copper colormap) and shaded topography and bathymetry (blue and gray colormaps). **(a)** Colored circles, earthquakes in the NEIC catalog from August 15, 2000 to June 3, 2020, colored by magnitude. White triangles, ASTUTI smartphone locations. MAT, Middle America Trench. SJ, San Jose city. **(b)** Beachball, focal mechanism of the 2012 Nicoya earthquake. Colored grid cells, Did You Feel It (DYFI) Modified Mercalli Intensity (MMI) values from the 2012 Nicoya earthquake.

Figure ~~2~~3. A typical ASTUTI station. **(a)** Photo showing the smartphone enclosure affixed to a wall. **(b)** Photo inside the enclosure of the smartphone screen showing the QED software application display.

Figure ~~3~~4. Schematic diagram of the ASTUTI approach. The general flow of alerting data and information is clockwise starts in the Device Networks box and end in the Alert System box. QED, 'Quick Event Detection' smartphone software. NTP, network time protocol. UDP, user datagram protocol . MQTT, message queuing telemetry transport. API, application programming interface. MySQL, My structure queried language.

Figure ~~4~~5. Examples of raw accelerometer and PGA data types from one station. Black, raw accelerometer data. V, vertical axis; H1, horizontal axis; H2, horizontal axis. Blue, P, processed PGA data message. **(a)** Example of how the PGA data type preserves the signal of 3 small earthquakes at 0, ~75 and ~210 seconds. **(b)** Example of how the PGA data type suppresses noise spikes at ~ 0 seconds.

Figure ~~5~~6. Histogram of PGA values for one station for one day (gray) and one hour (red). The current triggering threshold is 0.6%g.

Figure ~~6~~7. ASTUTI system latency and M7.6 Nicoya scenario vibration test. **(a)** Hourly distribution of data latency, δt_{data}, for the 6 month observation period. Each column represents the median δt_{data} probability distribution function for an hour of the day. **(b)** Location map (as in Figure ~~1a~~2a) showing the M7.6 Nicoya scenario vibration test. Red star, epicenter. Magenta circle, S-wave position at the time, δt_{detect}**,** the time the event was detected. Green triangles, the stations that triggered the alert. **(c)** P-message record section from the vibration test. Stations are ordered with distance from the epicenter. Hypocentral distance shown in blue text. **(d)** Histograms of time to detection, δt_{detect} (top) and time to receipt of alert, δt_{alert}**,** (bottom) for all 216 vibration tests.

Figure ~~7~~8. ASTUTI results from the 5 detected earthquakes~~ for the PGAN-4 algorithm~~. Each plot has the same symbology and nomenclature. Title with white box, origin time and magnitude of the event from OVSICORI. Magenta-box title, δt_{detect}. Magenta solid circle, estimated position of S-wave (3.2km/s) at time δt_{detect}. Magenta dashed circle, estimated position of S-wave 20 seconds after time δt_{detect}. 20 seconds represents ~5 seconds for alerting time,δt_{alert} (see Figure ~~5d~~6d) and 15 seconds for protective action such as Drop-Cover-Hold-On (DCHO). Green filled triangles, four stations that triggered the alert. White solid triangles, active stations at time of event. White empty triangles, inactive stations at time of event. Colored grid squares, DYFI cells. Hot population colormap, population % True-positive (TP) outcomes. Gray population colormap, % No-alert (NA). Copper population colormap, % True-positive no-shaking (TP-NS). **(a)** 12 December 2019 M_w 4.8. **(b)** 21 January 2020a M_w 5.3. **(c)** 21 January 2020b M_w 5.2. **(d)** 7 March 2020 M_w 5.2. **(e)** 13 March 2020 M_w 5.0.

Figure ~~8.~~9. PGA, P-message, record section for the 5 detected earthquakes~~ for the PGAN-4 algorithm~~. Each plot has the same symbology, nomenclature, and scale. The left column shows all stations for a particular event, the right column shows a zoomed image of only the triggering stations. Stations are ordered with distance from the epicenter. Hypocentral distance shown in blue along the vertical axis. Triggering stations, thick black. Other stations, black. Triggering time, magenta line. Pink shading, ranges of 6 and 3 km/s for P- and S-wave arrival times, respectively. **(a,b)** 12 December 2019 M_w 4.8. **(c,d)** 21 January 2020a M_w 5.3. **(e,f)** 21 January 2020b M_w 5.2. **(g,h)** 7 March 2020 M_w 5.2. **(i)** 13 March 2020 M_w 5.0.

Table 1. Alerting outcomes for each detected event~~ for PGAN-4 algorithms.~~. See also Figures 7 and 8. Time to detection, δt_{detect}. TP%, percent of population with True-positive outcomes. TP-NS%, percent of population with True-positive no-shaking outcomes. NA%, percent of population with No-Alert outcomes.

Reviewer #3

The authors are developing a low-cost observation network system over a wide area for earthquake early warnings using smartphones. In addition, the delay time is considered using a large number of seismic records, and it can be evaluated as a highly practical study.
Although it is not related to the essence, I think that it will be a better paper if the following points are corrected, so please consider the correction.

p3 LINE 56
It says "70% and 15% of the population", but according to the description in the text at LINE 494, it would be "15-75%".
p.6 LINE 183
It seems to be an explanation of Figure 2a. Therefore, it is desirable to change the reference from Figure 2 to Figure 2a.
p.10 LIME 363
The word "from" continues twice in a sentence.
p.10 LINE 364
Figure 8d is duplicated, one should be deleted.
p.11 LINE 405
DFYI is misspelled. DYFI is correct.
p.16 LINE 633
The quotes contain unknown characters such as "% @".
p.17 LINE 698,702
The quote contains unknown characters such as "% @".
p.17 LINE 708
The quote contains unknown characters such as "0% @" and "% U".
Also, the volume number is missing. (Risk Analysis, Vol. 39, No. 8, 2019.)
p.18 LINE 736
The quote contains unknown characters such as "% @".
p.18 LINE 751
The unknown character string "1." should be deleted.
Figure 3
Labels (a),(b) are not written on the figure.
Figure 6
The 1-hour histogram is 0.4 to 0.6 more frequently on the horizontal axis than the 1-day histogram. Is the vertical axis common? The vertical axis label should be shown.
Figure 7C
It seems that the terminal with a distance of 30.3km is shaking earlier than the terminal with a distance of 9.9km from the epicenter. I think it is necessary to explain this.
Also, it looks like there are three terminals that aren't shaking at all. Isn't it necessary to explain this?
Table 1
I think that TP-NS is correct when combined with the description in the text.