Investigating Seismicity and Structure of the Pecos, Texas Region of the Delaware Basin Using a Temporary Nodal Network

Jenna Lynn Faith
University of Texas at El Paso

Follow this and additional works at: https://scholarworks.utep.edu/open_etd

Part of the Geophysics and Seismology Commons

Recommended Citation

https://scholarworks.utep.edu/open_etd/3969

This is brought to you for free and open access by ScholarWorks@UTEP. It has been accepted for inclusion in Open Access Theses & Dissertations by an authorized administrator of ScholarWorks@UTEP. For more information, please contact lweber@utep.edu.
INVESTIGATING SEISMICITY AND STRUCTURE OF THE PECOS, TEXAS REGION OF THE DELAWARE BASIN USING A TEMPORARY NODAL NETWORK

JENNA LYNN FAITH
Doctoral Program in Geological Sciences

APPROVED:

______________________________
Marianne S. Karplus, Ph.D., Chair

______________________________
Diane I. Doser, Ph.D. Co-Chair

______________________________
Richard Langford, Ph.D.

______________________________
Jason Ricketts, Ph.D.

______________________________
William Ellsworth, Ph.D.

______________________________
Stephen L. Crites, Jr., Ph.D.
Dean of the Graduate School
Dedication

To my grandparents, Frank and Ruth Tallyen, for always supporting my dreams and providing the opportunities to make them come true. I wish you both could be here to see this.

To my husband, Jordan Rigdon Caylor. Without your constant love, support, and coding help, I can honestly say with about 95% certainty that I would not have made it this far.
INVESTIGATING SEISMICITY AND STRUCTURE OF THE PECOS, TEXAS REGION OF THE DELAWARE BASIN USING A TEMPORARY NODAL NETWORK

by

JENNA LYNN FAITH, B.S.

DISSEPTION

Presented to the Faculty of the Graduate School of The University of Texas at El Paso in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Department of Earth, Environmental, and Resource Sciences THE UNIVERSITY OF TEXAS AT EL PASO

December 2023
Acknowledgements

There are too many people to thank for helping me accomplish this goal of a defending a PhD. At UTEP, I thank my graduate advisors Diane Doser and Marianne Karplus. I thank Diane for taking me on as a student, always offering support, being a great mentor and teacher, and being a role model as a female in geophysics. I thank Marianne for offering the TexNet project to me, requiring funds, sticking up for my mental health when I needed it with collaborators, and being a great mentor. I thank the (former) graduate advisor, Annette Veilleux, for always being there to help and offer guidance. Carlos Montana for always being there to help with any technology issues. I thank all of my other professors, mentors, and committee members; Aaron Velasco, Stephen Veitch, Laura Serpa, Galen Kaip, Rip Langford, Jason Ricketts, Julian Chaput, Bill Ellsworth, and all of the other staff and professors in the geology department at UTEP. Even if I didn’t have you all as a professor, everyone was kind, welcoming, and showed they cared. I thank the committee members who offered me the various scholarships while at UTEP: the Cearley Graduate Scholarship in Geological Sciences, Bruce Davidson Memorial Graduate Student Award, the McBride Scholarship, and a Graduate School Travel Grant. These scholarships helped me tremendously and took financial stress off of my shoulders.

I would like to thank TexNet grant UTA18-000882 for funding the research and field work done in Pecos, Texas. I appreciate and thank each field assistant who spent time assisting with the monthly field work in Pecos: Jordan Caylor, Mohan Pant, Tony Garcia, Julia Astromovich, Lucia Gonzalez, Solymar Ayala Cortez, Dr. Lin Ma, and Lourdes Moreu. I thank the Town of Pecos City, Reeves County, Texas Department of Transportation, Wise Water Properties, Crockett Middle School, and all other land owners for hosting seismic nodal stations.

I thank Michelle Scalise at Mission Support and Test Services for sharing codes to run the HypoInverse software and organize the output. I thank Cliff Thurber at the University of Wisconsin-Madison for looking at my tomography files and providing feedback when I just couldn’t get the program to run correctly. I thank my colleagues at Los Alamos National Laboratory, specifically Mike Begnaud, Jonathan MacCarthy, Ellen Syracuse, Richard Alfaro-Diaz, Cathy Snelson, and Ting Chen for their endless support, research input, and advice while being a graduate student there. I thank my colleagues at NA-22 for all of the career and life advice I received while being a fellow there, and for choosing me to be one of your fellows. The experience and connections I made during the fellowship forever changed my career as a
scientist. I thank the Stanford professors and students who hosted me for two weeks in 2019, and helping me learn to code better and start one of my projects. I thank TexNet again for hosting me at their main office in Austin to teach me how to use their database so I could complete my projects.

At Juniata College, I thank my undergraduate professors Ryan Mathur, Matt Powell, and Larry Mutti for engraining me with the knowledge I needed for graduate school, reassuring me I was good enough to make it, the confidence to succeed, and the shoulder to cry on when I did not think my GRE scores were good enough to get in anywhere. A special shoutout to Pat Rea and Larry for recommending I apply to UTEP because attending this school has changed my life in ways I never would have imagined.

I thank my husband, Jordan Caylor, for so many things. When it came to school, he helped me so much with coding, got SimulPS working on my computer, gave me codes to run it, and helped with some bugs. He also contributed to a good portion of the field work. He was a field assistant many times, and put a lot of physical labor into digging holes for the nodes. As my husband, he has been my biggest fan and supporter throughout this journey. He has made me feel confident when I couldn’t do it myself, talked me into not quitting too many times, and was always there to listen to me complain. I do not know if I would have gotten this far without his constant support.

I cannot write this dissertation without thanking all of the friends I made at UTEP that helped me through and kept me going: Derek Scott, David Lankford-Bravo, Brandon Lajoie, Alex Eddy, Alex Garcia, Alan Vennemann, Tony Garcia, Sebastian Lopez, Juan Ochoa, Kate Grisi, Hanah and Dakota Draper, Erik Urquidi, Garrett Goff, Marc Ornelas, David Guenaga, and Lucia Gonzalez.

I am eternally grateful for my family and friends back home who have constantly supported me throughout my entire life. The biggest ones being my mom, Debbie, my brother Derek, Aunt Cheryl, Aunt Cathy, cousin Megan, sister-in-law, Kayla, and my best friends: Jenna, Regina, Veronica, Samantha, Cory, Ryan, Molly, Nakita, and Nate. Without their constant love and support, I would not be here today. I also am grateful for the support from Jordan’s family while obtaining this degree: Kim, Jim, Tim, Justin, McKenna, Jaxson, Gigi, Gump, Pop, and the Schraer’s.
Lastly, I cannot go without acknowledging the emotional support from my dog, Megan, and bearded dragon, Spike. They both have offered so much love, and endless cuddles, kisses, and smiles, especially on the hard days when I thought about quitting.
Abstract

With increasing earthquakes in the Delaware Basin since 2009, earthquake studies, including accurate hypocenters, are critically needed in the Delaware Basin to identify the structures producing earthquakes, and to determine if they are related to unconventional petroleum development and production. In 2018, with funding from the Texas Seismological Network (TexNet), the University of Texas at El Paso deployed and maintained a nodal network of 25 Magseis Fairfield Z-Land Generation 2 5-Hz seismic nodes in the Pecos, Texas region of the Delaware Basin, known as The Pecos Array. The network was deployed from November 2018 until the beginning of January 2020, with an additional two months of data recorded in September and October 2020. The network collected continuous 3-component data with a 1000-Hz sampling rate. The spacing of the nodes varied from ~2 km in town to ~10 km farther away from the city center. The primary goal of this network was to improve estimation of event hypocenters, which will help to determine why there has been an increase in earthquakes over the past several years. We summarize the scientific motivation, deployment details, and data quality of this network. Data quality statistics show that we successfully collected continuous data with signal-to-noise ratios that allow us to detect and locate events, hundreds of them being estimated at $M_L < 0.50$. This unique dataset is contributing to new seismotectonic studies in the Delaware Basin.

Using the Pecos Array, we analyzed five months (January – May 2019) of data using the machine learning algorithm, called PhaseNet, to assess how well it picked and located events compared to manual picks and earthquake locations from TexNet. We found that P-phase picks differed by an average of 0.0780 seconds, and S-phase picks differed by an average of 0.1829 seconds compared to the analysts. We used these phases to associate and relocate earthquake
events using HypoInverse and compared the resulting two catalogs. We found that events from both catalogs align with shallow normal faults and maximum horizontal compressive stresses observed in the area. Ninety-eight percent of all events occur above 5 km in depth, with an average error of < 1.5 km. Recent research suggests wastewater disposal and fracking to be the cause of the earthquakes. While we did not specially study the source of the earthquakes, our events fall at the depths where these procedures are occurring. We conclude that PhaseNet is a good alternative compared to standard automatic pickers, and that our network was able to successfully constrain hypocenter locations in the Pecos, TX region.

The Pecos Array was used to conduct a local 3-D tomography study in the Pecos, Texas region using the SIMULPS algorithm. Previous tomography studies have been conducted to update the velocity model of the entire Permian basin, but these studies covered a broader area, and their best velocity results were in the upper 35 km of the crust. With our denser array, localized earthquakes, and dense ray paths in the Pecos, Texas area, we have created a detailed 3-D velocity model for the top 8 km of this region. We observe velocity values that are consistent with reported geologies in the basin, and slower velocities on faults compared to areas without faults. We found that our new model is 28% slower compared to the DB1D velocity model that TexNet uses.
# Table of Contents

Dedication..........................................................................................................................iii  
Acknowledgements ...............................................................................................................v  
Abstract.................................................................................................................................viii 
Table of Contents ..................................................................................................................x 
List of Tables .........................................................................................................................xii 
List of Figures .......................................................................................................................xiii 

Chapter 1: The Pecos Array: A Temporary Nodal Seismic Experiment in the Pecos, Texas Region of the Delaware Basin .........................................................................................1  
   Introduction .........................................................................................................................1  
   Instrument and Deployment Details ..................................................................................4  
   Overall Data Quality and Availability .............................................................................7  
   Initial Observations ..........................................................................................................9  
   Summary ............................................................................................................................11  
   Figures ...............................................................................................................................12  

Chapter 2: Using PhaseNet, a Deep-Neural-Network, to Locate Earthquakes in the Pecos, Texas Region of the Delaware Basin ........................................................................23  
   Introduction .......................................................................................................................23  
   Data Collection ...............................................................................................................24  
   PhaseNet Workflow ........................................................................................................25  
   TexNet Analyst Workflow ...............................................................................................26  
   Comparison of Phase Picks ...........................................................................................27  
   Comparison of Earthquake Locations ..........................................................................28  
   Discussion .........................................................................................................................30  
   Conclusion .........................................................................................................................34  
   Tables ...............................................................................................................................36  
   Figures ...............................................................................................................................38  

Chapter 3: A 3-Dimensional Tomography Model of the Pecos, Texas Region of the Delaware Basin ..................................................................................................................46  
   Introduction .......................................................................................................................46
Seismic Data ................................................................................................................. 48
Initial Velocity Model ................................................................................................. 49
Tomography Method .................................................................................................. 50
Velocity Model Results ............................................................................................. 51
Discussion ................................................................................................................... 53
Conclusions ................................................................................................................ 55
Figures ....................................................................................................................... 57

Chapter 4: Comparing Earthquake Hypocenters from the TexNet, PhaseNet, and SIMULPS Catalogs Using the Pecos Array ................................................................. 71
Introduction ............................................................................................................... 71
Seismic Used .............................................................................................................. 71
Comparison of Seismicity Locations ......................................................................... 72
Discussion .................................................................................................................. 72
Conclusion .................................................................................................................. 73
Figures ....................................................................................................................... 74

References .................................................................................................................. 77
Appendix ...................................................................................................................... 85
Supplemental Figures ................................................................................................. 85

Vita 89
List of Tables

Table 2.1: Stations used, their locations, and months of operations during the timeframe of this study, January to May 2019.................................................................36
Table 2.2: Velocity model used for event relocation (Savvaidis et al., 2019). .................................37
List of Figures

Figure 1.1: Map of the Permian Basin with its geologic features, including the Delaware Basin to the west, the Midland Basin to the east, and the Central Basin Platform in between them. Our study area is labeled by the dashed red line. The location of Pecos, Texas is identified by the green triangle. The study area is located in Reeves County. ..................................................12

Figure 1.2: A map of the TexNet station network coverage between January 2017 to October 2018, a month before the deployment of the Pecos Array. The orange circles represent the earthquakes located by TexNet within a 50 km radius of Pecos, TX between January 2017 to October 2018. Pecos is located in the same area as station PB02. ..................................................13

Figure 1.3: Maps showing monthly changes in network configuration. (a) A map of every site occupied during the entire 14-month period. (b) A map of the network in the months of November and December 2018 with only 18 nodes. (c) A map of the network during the months of January through April 2019. The three circled stations were only deployed during January 2019 and were never reoccupied. (d) The final network configuration from May 2019 to the beginning of January 2020. ..........................................................14

Figure 1.4: Maps showing the additional deployments during Fall 2020. (a) Network configuration during September 2020. A total of 47 nodes were deployed. (b) Network configuration during October 2020. Twenty-six nodes were deployed. ........................................15

Figure 1.5: Photographs of the nodal deployment process for a Magseis Fairfield Z-Land Generation 2 3-Component 5-Hz seismic node. (a) The handheld computer (yellow and black device) used to activate or deactivate the node. (b) Photograph of a partially deployed node. Each node was buried such that the top was ~3 – 5 cm below the surface. A bubble level was used to level it, and a compass was used to orient the node to north. The node is approximately 30 cm long (including the spike that couples to the ground), and about 12.5 cm in diameter. .................16

Figure 1.6: An illustration of the availability and data return for each station during the 14-month deployment. The larger black boxes represent times when the node was deployed, but not actively collecting data. The small, black vertical lines represent the short time each month when we would replace a node with a new one. The blue line represents that time period in which a station was deployed and actively collecting data. (Casey et al., 2018)..................................................16

Figure 1.7: An illustration of the availability and data return for each station during the September to October 2020 deployment. The larger black boxes represent times when the node was deployed, but not actively collecting data. The small, black vertical line represents the short time when we serviced the node. The blue line represents that time period in which a station was deployed and actively collecting data. ..........................................................16

Figure 1.8: (a) Node median sample_rms values (in counts) for each station over the 14-month deployment. Warmer colors represent higher noise, and cooler colors represent lower noise. Higher noise values tend to occur along busy highways, while lower noise values are associated with stations that are located in more remote areas. The labeled stations show an example of this. (b) The total number of events each station recorded during the 14-month deployment. Not all stations were used for event picking, and are omitted from the figure. Cooler colors represent more events picked, while gray and warmer colors represent stations with less events picked on them. We must take into consideration how long each station was deployed when analyzing these numbers. ...........................................................................17

Figure 1.9: A sample of the Probability Distribution Function (PDF) plots in the network. These plots are another way to analyze noise, and show data for each component (top plots are east,
middle plots north, and bottom plots are vertical components). The colors (mostly violet to cyan) represent probability that the specific power level will occur at a certain frequency. Column 2 represents the station (PC53) with the lowest RMS value. Column 3 represents the station (PC45) with the highest RMS values (Casey et al., 2018). The red line represents the minimum power at each period, the black line represents mode, and the blue line represents that maximum power. The brown lines represent a low and high noise model for the entire network.

Figure 1.10: A moveout plot and corresponding signal to noise ratios (SNR) of stations for five random earthquakes during our 14-month deployment. In the SNR plots (right), orange circles represent Pecos Array stations, and purple squares represent TexNet stations. In the moveout plots, Pecos Array stations are labeled with “PC”, and TexNet stations are labeled with “PB” or called “PECS”. (a) This event occurred on December 24th, 2019 at 14:43:18.018. Location was 31.35776º latitude, -103.42906º longitude, 4.306 km depth, and M_L 0.89. (b) This event occurred on February 28th, 2019 at 23:22:55.169. Location was 31.43297º latitude, -103.52104º longitude, 2.841 km depth, and M_L 0.92. (c) This event occurred on June 19th, 2019 at 10:17:41.351. Location was 31.37733º latitude, -103.47277º longitude, 2.378 km depth, and M_L 0.74. (d) This event occurred on May 5th, 2019 at 08:11:17.746. Location was 31.38286º latitude, -103.4886º longitude, 2.378 km depth, and M_L 1.11. (e) This event occurred on October 14th, 2019 at 08:41:05.040. Location was 31.42746º latitude, -103.57742 longitude, 3.458 km depth, and M_L 0.88. Note that only TexNet stations within 75 km were plotted in the moveout plot, but all TexNet stations that had a P-phase pick were used in the SNR analysis.

Figure 1.11: A map of the locations of the five earthquakes used to determine signal to noise ratios in Figure 1.10. The purple stars indicate the locations of each earthquake, and are labeled accordingly based on which subfigure they are in Figure 1.10. The orange triangles are Pecos Array nodal stations.

Figure 1.12: (a) A histogram of magnitudes calculated for the ~2,500 events located in the polygon of interest (inset in upper right) for the 14-month deployment. These are M_L estimates provided by TexNet (Kavoura et al., 2020) (b) A bar graph showing how many earthquakes were located per month in the polygon of interest (in 10(a)) during the 14-month deployment. This information was provided by TexNet and includes the Pecos Array.

Figure 2.1: A map of the field area and seismic stations with Texas Roadways labeled. The blue triangles represent TexNet Stations (labeled PB** or PECS) while the red triangles represent the UTEP Pecos Array (labeled PC**). The black rectangle indicates the “polygon of interest” where we analyzed well-constrained earthquake locations. Pecos city center is in the middle of the figure where I-20, Hwy-17, and US-285 intersect. This figure shows every node that was deployed during the 14-month deployment, but the network only consisted of ~23 nodes at any given time (Table 2.1).

Figure 2.2: (a) A histogram of P-Phase time residuals. The residual is the result of the PhaseNet pick time subtracted from the analyst pick time. The delay of 0.043 seconds due to filtering and decimating has been accounted for. (b) Distribution of P-Phases for each individual event. The first histogram bar indicates P-Phases picked by both PhaseNet and an analyst. The middle histogram bar indicates phases detected only by PhaseNet. The last histogram bar indicates phases detected only by an analyst (i.e., missed by PhaseNet).

Figure 2.3: (a) A histogram of S-Phase time residuals. The residual is the result of the PhaseNet pick time subtracted from the analyst pick time. The delay of 0.043 seconds due to filtering and decimating has not been accounted for the S-picks due to the analysers using various filters to pick the S-phase. (b) Distribution of S-Phases for each individual event. The first histogram bar
indicates S-Phases picked by both PhaseNet and an analyst. The middle histogram bar indicates phases detected only by PhaseNet. The last histogram bar indicates phases detected only by an analyst (i.e., missed by PhaseNet).

Figure 2.4: Locations of all earthquakes between January and May 2019. The purple diamonds are PhaseNet events, and the green circles are TexNet analyst events. The “polygon of interest” is the thick black line. These locations were obtained using HypoInverse (Klein, 2014).

Figure 2.5: The orange lines indicate the locations of the shallow normal faults mapped by Hennings et al., (2021). The green dots are locations from TexNet analyst picks, and the purple diamonds are locations from PhaseNet picks. Cross section A-A’ is oriented along the strike of the faults and includes event within the entire polygon of interest. It is 38 km long and 17 km wide. Cross section B-B’ (17 km long) is perpendicular to A-A’ and includes events within 2.5 km of the cross section line. Cross section C-C’ (3 km long) is perpendicular to A-A’ and includes events within 1.25 km of the cross-section line.

Figure 2.6: Cross sections of seismicity (January to May 2019) in the Pecos area (see Figure 2.5 for details). All depths are relative to sea level. (a) A-A’ is oriented northwest-southeast and shows all events in the polygon of interest. (b) B-B’ is oriented northeast-southwest and shows events within 2.5 km of the cross section line. Ovals indicate two clusters of events. (c) C-C’ is also oriented northeast-southwest and shows events located within 1.25 km of the cross section line. Ovals highlight smaller clusters at ~1.5 km depth and ~3.5 km depth.

Figure 2.7: The depth distribution of seismicity for the TexNet analyst catalog and the PhaseNet catalog. The gray blocks indicate the depth of seismicity for the PhaseNet catalog (relocated using HypoInverse). The solid red line indicates the depth of seismicity for the TexNet analyst catalog (relocated using HypoInverse).

Figure 2.8: (a) Cross section figure from Horne et al., (2022) of a survey line ~25 km northwest of Pecos, TX. The shallow normal faults are mapped in the top 2 km of this survey. (b) A stratigraphic column of geologic formations whose depths were derived from sonic logs. The log is ~40 km southeast of Pecos, TX. The depths are from Sheng et al. (2022).

Figure 3.1: Map of the Permian Basin with its geologic features, including the Delaware Basin to the west, the Midland Basin to the east, and the Central Basin Platform in between them (See Data and Resources). Our study area is labeled by the dashed red line. The location of Pecos, Texas is identified by the green triangle. The study area is located in Reeves County.

Figure 3.2: A stratigraphic column of the Delaware Basin geology. Depths of formations are from Sheng et al., (2022).

Figure 3.3: (a) A map of the stations and earthquakes used in the tomography. The red triangles are Pecos Array stations, and blue triangles are TexNet stations. The yellow squares are the grid nodes. (b) A map of the grid nodes used for the tomography. The black polygon is the “area of interest” and earthquakes are only located within this polygon. The grid nodes are more densely spaced where the earthquakes and more stations are located.

Figure 3.4: A map of the 658 high quality events used for the VELEST inversion (Kissling et al., 1994) to determine the starting velocity model. The color of the dots represents the event depth.

Figure 3.5: (a) The blue line is the DB1D TexNet (TN) velocity model (Savvaidis et al., 2019), the green line is the velocity model used in Sheng et al. (2022), the red line is the smoothed model of Sheng et al. (2022) (SS), and the black line is the final starting velocity model used in the tomography. (b) The results of the VELEST (Kissling et al., 1994) inversions using the TexNet and Smoothed Stanford starting models for various depth spacings. (c) The orange
dashed line is the average of the VELEST runs. The black solid line is the final starting velocity model used in SIMULPS. The top 4 km of this model were adjusted so velocity continuously increases with depth. .................................................................61

Figure 3.6: The results of the L-Curve test to determine our damping parameter. The dark brown dot labeled “Damping = 40” is where we observed the “elbow” of the graph. .................................62

Figure 3.7: The Vp perturbations for depth slices 0.5 km to 3.0 km. Red colors represent areas decreasing in velocity, while blue areas represent areas increasing in velocity. The purple dotted line represents the densely gridded area. .................................................................63

Figure 3.8: The shallow normal faults (Hennings et al., 2021) mapped at the depth slices that they may occur at (1 – 3 km). They are overlain with the velocity (warm colors represent slower velocities, cool colors represent faster velocities), and Vp perturbations (red colors represent areas decreasing in velocity, blue colors represent areas increasing in velocity). The purple dotted line represents the densely gridded area. .................................................................64

Figure 3.9: (a) Earthquake locations from the original input file (blue dots) versus the relocations results from SIMULPS (red dots). (b) Earthquake depth comparisons of the original input file (blue dots) versus the earthquake depth results from SIMULPS (red dots). .................................65

Figure 3.10: The seismicity at depth slices 2 – 5 km overlain on Vp perturbations. Each slice shows earthquakes in a ± 0.5 km range of the depth slice layer name. For example, for “Depth: 2.0 km”, the figure is showing seismicity ranging from 1.5 – 2.5 km. .................................................................66

Figure 3.11: The ray density or derivative weight sum (DWS) for depth slices 0.5 – 8.0 km. The lighter colors represent areas with little ray density coverage. The darker colors represent areas with high ray density coverage. .................................................................67

Figure 3.12: The diagonal resolution element (DRE) for depth slices 0.5 – 8 km. Lighter colors represent areas with low DRE values. Darker colors represent areas with high DRE values........68

Figure 3.13: (a) The blue dotted line represents the starting velocity model used in SIMULPS. The pink line is the velocity results for the center of our study area. (b) The pink line represents the velocity within a non-faulted area (and center of our study area), located at the pink square in (d). The navy line represents the velocity of an area within a shallow normal fault zone, located as the navy square in (d). (c) Comparisons of our final velocity model (pink) to the TexNet DB1D velocity model (blue), and the Smoothed Sheng et al. (2022) model (green). (d) The locations of the pink “Final Model” in (a and c), pink “Model at Center” in (b), and the navy “Model on Fault” in (b). .................................................................69

Figure 3.14: (a) The locations of the tomography cross sections. Black dots are grid nodes and green triangles are station locations. (b) Tomography cross section of A – A’. The location is the red polygon in (a). (c) Tomography cross section of B – B’. The location is the blue polygon in (a). (d) Tomography cross section of C – C’. The location is the yellow polygon in (a). Each cross section (b) – (d) represents the grid nodes that are inside of the polygon. The pink square in each figure is the location where the three cross sections intersect. The black dots represent earthquake hypocenters. White triangles represent stations that are within each polygon. They are located under the maximum elevation (1.1 km above sea level) because that value is determined by the station with the highest elevation. If a station elevation is lower than that maximum value, it will map beneath it. ..................................................................................70

Figure 4.1: (a) Comparison of TexNet epicenters in blue, and PhaseNet epicenters in yellow. (b) Comparison of TexNet epicenters in blue, and SIMULPS epicenters in red. (c) Comparison of
SIMULPS epicenters in red, and PhaseNet epicenters in yellow. (d) Comparison of all three catalogs in one map. TexNet epicenters are blue, SIMULPS red, and PhaseNet are yellow.

Figure 4.2: (a) Comparison of TexNet earthquake depths in blue and PhaseNet depths in yellow. (b) Comparison of TexNet earthquake depths in blue and SIMULPS depths in red. (c) Comparison of SIMULPS earthquake depths in red and PhaseNet depths in yellow. (d) Comparison of earthquake depths for all three catalogs. TexNet depths are blue, SIMULPS red, and PhaseNet are yellow.

Figure 4.3: A histogram comparing earthquake depths of the three catalogs. PhaseNet depths are the yellow bars, TexNet depths is the blue solid line, and SIMULPS depths is the red solid line.

Figure A2.1: Seismograms (a)-(d) represent examples where PhaseNet and the TexNet analysts are in agreement with their phase picks. The slight pick difference is due to the analyst picking phases on unprocessed data, while PhaseNet phase picks were performed on filtered and decimated data. Seismogram (e) is an example of a seismogram where PhaseNet and the analysts did not agree, and the processing effects do not make up for the pick differences. Seismogram (f) shows an example where PhaseNet picked a P- and S- phase, but the TexNet analysts did not. Seismogram (g) shows an example where the TexNet analysts picked a P- and S- phase, but PhaseNet did not. The seismograms shown are the filtered and decimated waveforms.

Figure A3.1: The Vp perturbations for depth slices 0.5 – 8 km. Red colors represent areas decreasing in velocity, while blue areas represent areas increasing in velocity. The purple dotted line represents the densely gridded area.

Figure A3.2: Velocity values for depth slices 0.5 – 8 km. Warmer colors represent slower velocities. Cooler colors represent faster velocities.

Figure A3.3: Seismicity at depth slices 0.5 – 8 km overlain on Vp perturbations. Each slice shows earthquakes in a ± 0.5 km range of the depth slice layer name. For example, for “Depth: 4.0 km”, the figure is showing seismicity ranging from 3.5 – 4.5 km.
Chapter 1: The Pecos Array: A Temporary Nodal Seismic Experiment in the Pecos, Texas Region of the Delaware Basin

INTRODUCTION

The Town of Pecos City, Texas (TX) (hereafter referred to as Pecos) is located in Reeves County within the Delaware Basin, the western sub-basin of the larger Permian Basin (Figure 1.1). The Delaware Basin covers an area of approximately 33,500 km$^2$ throughout west Texas and southeastern New Mexico, and has a volume of approximately 170,000 km$^3$ of sedimentary strata (Hills, 1984). The Delaware Basin is bounded by the Central Basin Platform to the east, the Diablo Platform to the west, the Marathon Shelf and Ouachita-Marathon orogenic belt to the south, and the Northwestern Shelf to the north-northwest (Figure 1.1).

Exploration for crude oil and natural gas in the Delaware Basin began in the 1920s, contributing to the success of the Permian Basin over the last 100+ years. After its first major oil boom in the 1970s, it was thought the basin had dried up, but advances in hydraulic fracturing and horizontal drilling allowed the basin to thrive again (Federal Reserve Bank of Dallas, 2022). A 10-station seismic array, deployed in 2000, located ~240 km south of Pecos, TX, indicates that seismicity started in the Pecos region in 2007, when an earthquake was recorded (Frohlich et al., 2019). In 2009, a few more earthquakes were recorded. This number continued to increase with more than 1,400 events being recorded in 2017, correlating with oil and gas production returning to the basin (Frohlich et al., 2019). To better locate and understand seismicity in Texas, the Texas Seismological Network (TexNet) began deploying broadband seismometers in January of 2017 and has continuously been expanding their network (Savvaidis et al., 2019) (Figure 1.2). In 2021 alone, over 400 events of magnitude $> 1.5$ were recorded by TexNet within a 50 km radius of Pecos, TX (see Data and Resources).
In response to the large increase of earthquakes in the Pecos, TX region, TexNet funded the University of Texas at El Paso (UTEP) to deploy a temporary nodal array, called the Pecos Array, in 2018. The Pecos Array was designed to improve ray coverage and better constrain event hypocenters near Pecos, TX. The array consisted of 38 Magseis Fairfield Z-Land Generation 2 5-Hz seismic nodes deployed for 14 months from November 2018 to the beginning of January 2020, each month deploying 23 (on average) nodes (Figure 1.3 (a)). The starting network spanned a region of approximately 70 km by 65 km (Figure 1.3 (b)), while the final network was confined to an area of 50 km by 40 km (Figure 1.3 (d)). The Pecos Array project officially ended in January 2020, but additional TexNet funding for a different project allowed us to deploy nodes during the months of September and October 2020. Passive source data was collected and complements the original Pecos Array (Figure 1.4).

In other areas of West Texas, events are related to tectonic (Doser et al., 1991) and human-induced processes (Frohlich et al., 2016), so it is important to have accurate hypocenters to assess if these events are related to petroleum production activities. Recent research associates the increase in seismicity in the Delaware Basin to shallow wastewater disposal procedures, and a small percent due to hydraulic fracturing (Skoumal et al., 2020; Zhai et al., 2021; Savvaidis et al., 2020; Grigoratos et al., 2022). Regional studies by Lomax and Savvaidis (2019) give hypocentral depths ranging from -1 km to 12 km with an error up to 5 km, while Sheng et al. (2022) give depths ranging from 0.8 km to 3 km. Such ambiguity in the depth causes challenges in determining what mechanisms are causing the seismicity.

The Delaware Basin and Central Basin Platform are defined structurally by a complex network of basement-rooted faults (Horne et al., 2021), and shallow normal faults (Hennings et al., 2021). Both reverse and normal faults are found throughout the basin, and have the potential
to slip with elevated pore pressures (Hennings et al., 2021). Shallow normal faults are mostly observed in Reeves County in the Delaware Mountain Group and Bone Springs Formation (Hennings et al., 2021) with a northwest-southeast orientation of maximum horizontal stress (Snee et al., 2018; Horne et al., 2021). Hennings et al. (2021) calculated deterministic fault-slip potential (DFSP) for the entirety of the Delaware Basin, and found that Reeves County has a low DFSP of $\leq 2.5$ MPa., indicating that the shallow normal faults are prone to reactivation with an increase in pore pressure. The basement-rooted faults show stable DFSP; therefore, they have not been correlated to earthquake events in the Pecos area (Hennings et al., 2021). Our dense network will specifically help to constrain the event depths around the Pecos area, and provide insights into the source of the seismicity.

With the population of Pecos increasing almost 50% since 2010 (U.S. Census Bureau, 2021), mitigating the potential of oil and gas activities inducing damaging earthquakes is critical to the safety of the local population. The safety of infrastructure is also important considering Reeves County produced nearly 80 million barrels of oil and 886 million mcf (thousand cubic feet) of gas in 2021 (Texas Railroad Commission, 2022).

The Pecos Array will contribute better-constrained hypocenter locations, structural and tectonic understanding, improved velocity models, tomography studies, noise characterization, updated hazard maps, and overall, more detailed studies of this specific area of the Delaware Basin. This data set provides station spacing that varies between 2-10 km, compared to the previous TexNet spacing of $>20$ km in 2018 (Savvaidis et al., 2019) (Figure 1.2), 14 months of continuous data collection, plus two additional months of data. In this paper, we describe the data collection of the Pecos Array, including network geometry, individual station details, field challenges, data availability, and quality.
**INSTRUMENT AND DEPLOYMENT DETAILS**

From November 2018 through January 2020, the UTEP seismology team made monthly visits to service an average of 23 Magseis Fairfield Z-Land Generation 2 5-Hz 3-component nodes deployed in the Pecos, TX region. (Figure 1.3). The spacing of the nodes varied from ~2 km in town and increased to ~10 km in regions farther away from the city center. The nodes collected continuous three component data with a sampling rate of 1000-Hz, an internal corner frequency of 5-Hz, with gain set at 12 dB.

We visited each site every ~35 days, the approximate maximum battery life, to swap the nodes. We buried each node such that the top was located ~3-5 cm from the surface, was aligned with true north using a handheld compass (considering the local declination of 8°), and leveled with a bubble level (Figure 1.5). After each month’s service visit, we brought the nodes back to UTEP and offloaded the data. Data were offloaded in Fairfield continuous (fcnt) format for each month and then converted to PH5 format using the Incorporated Research Institutions for Seismology (IRIS) Portable Array Seismic Studies of the Continental Lithosphere (PASSCAL) PH5 tools (see Data and Resources). Then SEED format files were extracted from the PH5 archives for each month, and the data were archived in SEED at the IRIS Data Management Center (DMC) (Karplus et al., 2018).

The configuration of the network was based on the need to improve earthquake locations for seismicity occurring near Pecos, TX, and constrained by challenges seeking landowner permissions to host seismic nodes. The nodes were only placed in areas where we had landowner permission. The nodes were located at various sites, including along Texas Department of Transportation (TXDOT) highways, county roads, Pecos City water well fields, city and county
infrastructure sites, private businesses, and Crockett Middle School, where a node was co-
located with a TexNet broadband seismic station. The noise varied from site to site.

The number of occupied sites varied month to month due to road construction, safety,
movement of nodes from noisy sites to more optimum locations, filling in seismicity gaps, and
landowner permissions (Figure 1.6). Only 18 nodes were deployed at the beginning of network
operation in November 2018 (Figure 1.3 (b)). Our plan was to deploy some nodes in areas where
we had secured landowner permission while waiting for permission to occupy other potential
sites. In January 2019, we were given permission by TXDOT to place nodes along certain
highways (Figure 1.3 (c)). This allowed us to locate nodes closer to the city center with smaller
node spacing, the ability to fill in seismic network gaps, and an overall decrease in the area of
our network. The network mostly maintained this configuration until May 2019, when we
obtained new landowner permissions (Figure 1.3 (d)). These permissions allowed us to remove
nodes from near the highway in the northwest corner of our deployment, and move them to a less
trafficked area.

Excluding the first two months of the deployment and May 2019, the network
consistently had 20-25 nodes deployed, with an average of 23 nodes per month. The month of
May only had 19 nodes due to challenges faced in the field. One node, PC41, was thought to be
lost/stolen, so a new node was not put in its place, but the node was found during the next
deployment in June. We found site PC45 under mounds of dirt due to road construction, so it
was unsafe to reoccupy that location.

These types of situations would occur often in our field area, and we would have to adapt.
We would sometimes find our node trapped by road construction or find it was in an unsafe
space, so we would pull the node and not replace it with a fresh one. We would sometimes have
a new location to put the fresh node, but other times we would have to wait until the next month to redeploy that node in a different location. This allowed us to maintain flexibility in deploying nodes to ensure adequate station coverage of the region.

There were two areas where we had to move the node < 50 meters, but gave them new names since their GPS points would be different. The first example is the area where stations PC06, PC10, and PC26 were located. PC06 was our first site. A bee’s nest invaded the site, so we moved it to an area about 25 meters away. This became site PC10, but we subsequently had to move the node again since it was located within a washout zone when it rained. The final location was PC26, which was about 20 meters away from both of the previous sites. The other sites that needed to be moved slightly were PC13 and PC44. PC13 was first placed under a bridge, and then was moved about 15 meters away from the bridge so it would be less noisy and not be affected by possible road construction. It became PC44.

In September 2020, another group at UTEP received funding from TexNet for a different project in the Pecos area. The nodes were redeployed for about six weeks that recorded passive source data that compliments that original Pecos Array. In September, 47 nodes were deployed (Figure 1.4 (a)). In October, 26 nodes were deployed (Figure 1.4 (b)). In many cases, we reoccupied sites from the November 2018 to January 2020 deployment, but also expanded the network to cover a larger area (Figure 1.7). To find previously occupied sites, we used GPS points and photographs. The GPS has an accuracy of 3 meters. Due to lack of rain and rocky soil, the holes dug for the original deployment were often still there. Softer soil was present adjacent to hard soil, making it fairly easy to access the same holes. If we were not positive where the original node was located, the node site was given a new GPS coordinate and name.
September’s deployment covered an area of 85 km by 80 km, while October covered an area of 70 km by 60 km. The node spacing for these two months varied from ~2 km in town and increased to ~10 km in regions farther away from the city center. The nodes collected continuous three component data at a sampling rate of 1000-Hz for the two months. Once we took the nodes back to UTEP, the off-loading and processing procedures were the same as described for the previous phases.

**Overall Data Quality and Availability**

The Pecos Array dataset will be available at the IRIS DMC in SEED format under network code 7P (2018–2020) (Karplus et al., 2018). The site locations that were occupied for the entire 14-month deployment returned at least 95% of the data. Figure 1.6 shows the availability for each station during the deployment. Larger black boxes represent times when the node was deployed, but not collecting data. This could have been due to the battery dying, not being able to replace the node, or the recovered node data was not useable. The small, black vertical lines represent the short times each month when we would swap the node with a fresh one. The horizontal blue line represents the time periods in which a site was deployed and actively collecting data. Throughout the 14 months, we were able to capture continuous seismic data across many stations.

An analysis of data quality was conducted using the IRIS PIQQA (PI’s Quick Quality Assessment) Data Quality Report and MUSTANG (Modular Utility for Statistical Knowledge Gathering) tools (Casey et al., 2018). The sample_rms metric was computed to determine average noise levels at each station. The sample_rms calculates the RMS variance of trace amplitudes within a 24-hour window. We calculated the sample_rms value for each day, for each station between November 10, 2018 to January 10, 2020 for each of the three components. For
each station, we found the median sample_rms value for each component, and averaged the values between the three components. This resulted in a single sample_rms value per station for the entire deployment.

Sites along busy roads were much noisier compared to sites that were located away from major highways (Figure 1.8 (a)). We observe this especially in the northwest corner of the deployment. Stations PC39, PC34, and PC35 were placed several meters off of a busy road, and their average sample_rms values were 7710 counts. We obtained new land permissions and were able to remove those stations off the road. We deployed stations PC50, PC51, PC53, and PC54 in a more remote area, and they produced an average sample_rms value of 432 counts.

We calculated the number of events each station recorded during the 14-month deployment (Figure 1.8 (b)). While some stations were deployed longer than others, each station recorded events while they were out. We see those stations with lower sample_rms values are the ones that recorded more events. While analyzing this data, we must consider how long each station was deployed. The earthquake locations that TexNet provided did not use every station we deployed for event picking, but this was not due to lack of good quality data.

Probability density function (PDF) plots were created through MUSTANG’s noise-pdf public web service as another way to analyze noise (Casey et al., 2018). The first column in Figure 1.9 is a Network Composite PDF for each component. It is created by using power spectral density (PSD) values for all network stations. Columns two and three in Figure 1.9 analyze stations with the lowest (PC53) and highest (PC45) sensitivity-scaled sample_rms values for each component. For frequencies above 1-Hz, PC45 is showing much higher power compared to PC53.
INITIAL OBSERVATIONS

The Pecos Array data show that the nodes have clearly recorded earthquakes (Faith et al., 2021). Figure 1.10 shows five moveout plots and corresponding signal to noise ratios (SNR) for stations of five random earthquakes throughout our 14-month deployment. The noise was calculated by averaging the absolute value of the signal 45 seconds before the estimated P-arrival. Signal was calculated by averaging the absolute value of the signal in a 5 second window beginning at the estimated P-arrival. The calculated signal divided by the calculated noise resulted in our ratio. Only stations with P-phase picks were used. Both the Pecos Array nodes and the TexNet stations were filtered using a bandpass filter between 1-25 Hz before the SNR analysis was calculated. TexNet provided the pick, location, and magnitude information.

These five events range in magnitude from $M_L$ 0.74 to 1.11, and are all located in the polygon of interest in Figure 1.12 (a) (Figure 1.11). The Pecos nodes have an average SNR of 7.73 (for these five events), while the TexNet stations’ average SNR is 3.39. The only TexNet stations that are located in the polygon of interest is a shallow vault station PB02 (available until July 2020), and PCOS (a borehole station deployed in April 2019 at 100 m), both co-located with our PC01 station. We would expect these stations to show higher SNR compared to surface nodes. The events in Figure 1.10 (b) and (c) have only been detected because our array provided enough coverage and quality P-phase arrivals to locate them. Figure 1.10 (b) earthquake was only detected on two TexNet stations and Figure 1.10 (c) was only detected on three. While we do not find phase arrivals on all of the Pecos Array nodes, the picks we do have contribute to more earthquakes being detected, and better constrained hypocenters.

Currently, we are locating events using the dataset in Faith et al., 2021, and confining our area to the polygon of interest in Figure 1.12 (a). TexNet provided the location and magnitude
information, which includes our Pecos Array nodes. TexNet’s public catalog does not include our nodal array, and only shows magnitudes > 0. In the polygon of interest, ~2,500 events have been located over 14 months, ranging in magnitude from -0.54 to 2.54 (Figure 1.12 (a)). The median magnitude value is 0.79, and are $M_L$ estimates (Kavoura et al., 2020). Figure 1.12 (b) shows the number of earthquakes located per month during the 14-month deployment. February 2019 had the highest number of earthquakes at 262, while November 2019 had the lowest at 83. The average number of events per month is 176. The association and event locations were generated using SeisComp3 software, and then refined using the location algorithm, NonLinLoc (Lomax et al., 2000; 2009). It uses the 1-D velocity model, DB1D in Savvaidis et al., 2019.

These locations will then be used to create tomographic models to help us update the velocity layering of the area and analyze the underlying structure. These data are currently being used to validate a machine learning algorithm, PhaseNet (Zhu et al., 2018), to determine how well it picks phases and locates events in Texas using different instruments than what it was trained on in California (Faith et al., 2021). A high-resolution receiver function analysis is also being conducted using the data (Veitch et al., 2022). Lastly, waveforms recorded by the nodes are being used for moment tensor analyses (Savvaidis and Roselli, 2021).

Once earthquake phase data and accurate locations are obtained, many research projects can be derived from this dataset. For example, we know that there is velocity anisotropy in the area (Doser et al., 1992), so this can be further studied and incorporated into a new generation of velocity models. Noise recorded by the network can be used to conduct ambient noise tomography studies. If there appear to be multiple sources for earthquakes (e.g. tectonism, water injection, hydraulic fracturing), further signal analysis can be completed to determine if the
different source waveforms produce specific characteristics enabling a rapid means to differentiate between them.

**SUMMARY**

We have been able to successfully deploy a temporary nodal-network in the Pecos, TX region that provided continuous data for 14 months from November 2018 to the beginning of January 2020 with two additional months of data in September and October 2020. Our network provided denser station spacing and ray coverage in the Pecos, TX area, which will allow for more precise earthquake locations, tectonic and structural understanding, and improved modeling. This deployment was able to detect thousands of events in an area that is of high importance to the community given the sudden increase in earthquakes over the past several years. We hope researchers use this data set to expand the research being done in this area of west Texas.
Figure 1.1: Map of the Permian Basin with its geologic features, including the Delaware Basin to the west, the Midland Basin to the east, and the Central Basin Platform in between them. Our study area is labeled by the dashed red line. The location of Pecos, Texas is identified by the green triangle. The study area is located in Reeves County.
Figure 1.2: A map of the TexNet station network coverage between January 2017 to October 2018, a month before the deployment of the Pecos Array. The orange circles represent the earthquakes located by TexNet within a 50 km radius of Pecos, TX between January 2017 to October 2018. Pecos is located in the same area as station PB02.
Figure 1.3: Maps showing monthly changes in network configuration. (a) A map of every site occupied during the entire 14-month period. (b) A map of the network in the months of November and December 2018 with only 18 nodes. (c) A map of the network during the months of January through April 2019. The three circled stations were only deployed during January 2019 and were never reoccupied. (d) The final network configuration from May 2019 to the beginning of January 2020.
Figure 1.4: Maps showing the additional deployments during Fall 2020. (a) Network configuration during September 2020. A total of 47 nodes were deployed. (b) Network configuration during October 2020. Twenty-six nodes were deployed.
Figure 1.5: Photographs of the nodal deployment process for a Magseis Fairfield Z-Land Generation 2 3-Component 5-Hz seismic node. (a) The handheld computer (yellow and black device) used to activate or deactivate the node. (b) Photograph of a partially deployed node. Each node was buried such that the top was ~3 – 5 cm below the surface. A bubble level was used to level it, and a compass was used to orient the node to north. The node is approximately 30 cm long (including the spike that couples to the ground), and about 12.5 cm in diameter.

Figure 1.6: An illustration of the availability and data return for each station during the 14-month deployment. The larger black boxes represent times when the node was deployed, but not actively collecting data. The small, black vertical lines represent the short time each month when we would replace a node with a new one. The blue line represents that time period in which a station was deployed and actively collecting data. (Casey et al., 2018)
Figure 1.7: An illustration of the availability and data return for each station during the September to October 2020 deployment. The larger black boxes represent times when the node was deployed, but not actively collecting data. The small, black vertical line represents the short time when we serviced the node. The blue line represents that time period in which a station was deployed and actively collecting data.
Figure 1.8: (a) Node median sample_rms values (in counts) for each station over the 14-month deployment. Warmer colors represent higher noise, and cooler colors represent lower noise. Higher noise values tend to occur along busy highways, while lower noise values are associated with stations that are located in more remote areas. The labeled stations show an example of this. (b) The total number of events each station recorded during the 14-month deployment. Not all stations were used for event picking, and are omitted from the figure. Cooler colors represent more events picked, while gray and warmer colors represent stations with less events picked on them. We must take into consideration how long each station was deployed when analyzing these numbers.
Figure 1.9: A sample of the Probability Distribution Function (PDF) plots in the network. These plots are another way to analyze noise, and show data for each component (top plots are east, middle plots north, and bottom plots are vertical components). The colors (mostly violet to cyan) represent probability that the specific power level will occur at a certain frequency. Column 2 represents the station (PC53) with the lowest RMS value. Column 3 represents the station (PC45) with the highest RMS values (Casey et al., 2018). The red line represents the minimum power at each period, the black line represents mode, and the blue line represents that maximum power. The brown lines represent a low and high noise model for the entire network.
Figure 1.10: A moveout plot and corresponding signal to noise ratios (SNR) of stations for five random earthquakes during our 14-month deployment. In the SNR plots (right), orange circles represent Pecos Array stations, and purple squares represent TexNet stations. In the moveout plots, Pecos Array stations are labeled with “PC”, and TexNet stations are labeled with “PB” or called “PECS”. (a) This event occurred on December 24th, 2019 at 14:43:18.018. Location was 31.35776º latitude, -103.42906º longitude, 4.306 km depth, and M_L 0.89. (b) This event occurred on February 28th, 2019 at 23:22:55.169. Location was 31.43297º latitude, -103.52104º longitude, 2.841 km depth, and M_L 0.92. (c) This event occurred on June 19th, 2019 at 10:17:41.351. Location was 31.37733º latitude, -103.47277º longitude, 2.378 km depth, and M_L 0.74. (d) This event occurred on May 5th, 2019 at 08:11:17.746. Location was 31.38286º latitude, -103.48886º longitude, 2.378 km depth, and M_L 1.11. (e) This event occurred on October 14th, 2019 at 08:41:05.040. Location was 31.42746º latitude, -103.57742 longitude, 3.458 km depth, and M_L 0.88. Note that only TexNet stations within 75 km were plotted in the moveout plot, but all TexNet stations that had a P-phase pick were used in the SNR analysis.

Figure 1.11: A map of the locations of the five earthquakes used to determine signal to noise ratios in Figure 1.10. The purple stars indicate the locations of each earthquake, and are labeled accordingly based on which subfigure they are in Figure 1.10. The orange triangles are Pecos Array nodal stations.
Figure 1.12: (a) A histogram of magnitudes calculated for the ~2,500 events located in the polygon of interest (inset in upper right) for the 14-month deployment. These are $M_L$ estimates provided by TexNet (Kavoura et al., 2020) (b) A bar graph showing how many earthquakes were located per month in the polygon of interest (in 10(a)) during the 14-month deployment. This information was provided by TexNet and includes the Pecos Array.
Chapter 2: Using PhaseNet, a Deep-Neural-Network, to Locate Earthquakes in the Pecos, Texas Region of the Delaware Basin

Introduction

Since 2009 there has been an increase in seismicity in the town of Pecos City, TX (hereafter referred to as Pecos) region (Figure 2.1) from several earthquakes detected in 2009 to more than 2000 events detected in 2017, correlating with an increase in oil and gas production in the area (Frohlich et al., 2019). To better understand the increase in seismicity, the Texas Seismological Network (TexNet) began deploying broadband seismometers throughout Texas in 2017, and continues to update their network (Savvidis et al., 2019). From January 2017 to December 2021, TexNet recorded approximately 2,100 earthquakes of magnitude > 1.5 within a 50 km radius of Pecos, Texas, along with 57 M$_{L}$ 3+ earthquakes, and approximately 990 M$_{L}$ 2.0-2.9 earthquakes (TexNet Earthquake Catalog, 2022, Last accessed 08/29/2022). The sparse station spacing of TexNet results in hypocentral depths uncertainties of up to 4-5 km (Lomax and Savvidis, 2019). To better constrain these hypocenters, TexNet funded the University of Texas at El Paso (UTEP) to design a temporary network that included the deployment of ~23 three-component Magseis Fairfield Z-land 5-Hz nodes in and around the Pecos, TX area. The nodes recorded data from November 2018 to the beginning of January 2020.

Our 14-month network deployment created a very large dataset, and as our data set grew, it became difficult for analysts to pick phase arrivals by hand in a timely matter. The accuracy and subjectivity of picks strongly affects the calculated earthquake locations and depths. Analyst picking can be inconsistent because they choose phases based on years of experience and personal judgment, and not every analyst picks waveforms the same way (Leonard, 2000). Consistent picking is crucial in this area of West Texas because events are related to both
tectonic (Doser et al., 1991) and human-induced processes (Frohlich et al., 2016, Savvaidis et al., 2022), so it is important to have accurate hypocenters when making observations about the source of the earthquakes.

We wanted to test whether a machine learning algorithm could obtain faster, accurate, and more consistent phase picks compared to humans, so we tested five months of data using a deep-neural-network-based arrival-time picking method “PhaseNet” (Zhu and Beroza, 2018), that picks both P and S phases. We wanted to determine how well an algorithm trained on California earthquakes would perform in the Pecos, TX region of the Delaware Basin. Research in other settings, such as the Raton Basin in New Mexico/Colorado (Wang et al., 2020) and in Oklahoma and southern Kansas (Park et al., 2022), have been able to accurately identify and pick P and S phases using PhaseNet.

For this paper, we performed an in-depth statistical analysis of phase pick differences between PhaseNet and TexNet’s data, picked by analysts. We then used the pick catalogs to associate and locate the events to compare how pick detection and timing differences affect hypocentral locations.

**DATA COLLECTION**

Seismic data for this study were obtained from two networks, TexNet and the Pecos Array, a temporary nodal network operated by the University of Texas at El Paso (Figure 2.1). TexNet started deploying broadband seismometers throughout Texas in January 2017 (Savvaidis et al., 2019). In 2018, TexNet funded a UTEP seismology team to deploy the Pecos Array, a network of ~23 Magseis Fairfield Z-Land Generation 2 5-Hz, 3-component nodes (Karplus et al., 2018; Faith et al., 2023). This dense network provided better ray coverage for the Pecos region. Data for this study cover the months of January – May 2019.
The spacing of the nodes is as little as ~2 km in Pecos and increased to ~10 km in regions farther from the city center. Occupation of node sites varied by month due to road construction, movement of nodes from noisy sites to more optimum locations, safety, and landowner permissions (Table 2.1). The nodes collected data from November 2018 to the beginning of January 2020 at a sampling rate of 1000-Hz with an internal corner frequency of 5-Hz. Data from the Pecos Array is available through the Incorporated Research Institutions for Seismology (IRIS) Data Management Center (DMC) under network code 7P (2018–2020) (Faith et al., 2023).

**PHASENET WORKFLOW**

PhaseNet is a deep-neural-network that picks P- and S- phase arrivals. It was originally trained by “learning” signal patterns of labeled signal and noise using data from the Northern California Seismic Network. PhaseNet analyzes 3-component (3-C) seismograms at 100-Hz, and outputs the probability that each sample is either a P-phase, S- phase, or noise (Zhu and Beroza, 2018). To start this process, 3-C seismic waveforms are input into PhaseNet. These waveforms undergo four down-sampling stages of convolution. Once the algorithm reaches the final stage, the waveforms undergo four up-sampling stages of deconvolution. This converts the information to the output, the probability distributions of P- and S- phase arrival times and noise. For a more detailed analysis of how PhaseNet works, please refer to Zhu and Beroza (2018).

The peaks of the probability distribution for the P and S phases are taken to be the phase arrival times. The chosen probability threshold will trigger PhaseNet to call a signal a “pick”. It is on a scale of 0 – 1, where 0 is low probability of signal being a real arrival time and 1 is a very high probability. We used a probability threshold of 0.3 because it best allowed us to extract accurate arrival times and to obtain enough phases to locate an event.
The following preprocessing procedures were applied before applying PhaseNet. We filtered the data using the ObsPy butterworth-bandpass, minimum phase filter from 1-25 Hz (Beyreuther et al., 2020), and then decimated the Pecos Array data from 1000-Hz to 100-Hz since PhaseNet was trained on waveforms sampled at 100-Hz. The ObsPy decimate function was used, which only uses the decimation-factor sample, while the rest are thrown away. This function applies a lowpass filter before decimating to avoid aliasing artifacts (The ObsPy Development Team, 2022). The TexNet stations were also filtered between 1-25 Hz, and decimated from 200-Hz to 100-Hz before being processed by PhaseNet.

**TexNet Analyst Workflow**

We compared the automatic picks from PhaseNet with picks derived from TexNet’s typical phase picking workflow (Savvaidis et al., 2019). Both workflows picked phases on the Pecos Array stations and various TexNet stations (Table 2.1). TexNet’s workflow involves the following steps: 1) automatic picking using a short-term average over long-term average (STA/LTA) calculation, 2) application of an enhanced S-picker module (Grigoli et al., 2018), and 3) manual checking and correction of picks by human analysts (Savvaidis et al., 2019).

We found that there is a slight time delay in the PhaseNet phases due to down-sampling and filtering. The analyst picked P-phases on the Pecos Array nodes using unfiltered data at a 1000-Hz sampling rate, while PhaseNet picked events on filtered data with a 100-Hz sampling rate. During this process, the PhaseNet phase is delayed by 0.043 seconds. The analysts did use various filters to pick the S-phase, so determining a consistent delay was not possible due to each event being filtered differently. While still using a 1000-Hz sampling rate, the range of filters used to determine the S-phases were 1 – 5 Hz, 2 – 8 Hz, and 3 – 12 Hz.
**COMPARISON OF PHASE PICKS**

In this study, we compared PhaseNet’s automatic P- and S-phase arrival picks with those obtained from the TexNet workflow for 5 months from January to May 2019. We compared 915 events. In total, 8,659 P-phases were found in common between the TexNet analysts (hereafter termed “analyst”) and PhaseNet. A PhaseNet phase was considered “in common” with the analyst if it was within +/- 1.50 seconds of the analyst phase pick time. The same station/phase pair were compared for each event. These common phases were then used in our analyses.

Statistics are based on phases having time residuals between -1.500 and +1.500 seconds of the analyst pick. We calculated the time residual by subtracting the PhaseNet time arrival from the analyst time arrival. The mean P-phase time residual was 0.0780 seconds with a standard deviation of 0.1748 seconds. The absolute value was used to calculate the mean since both positive and negative residuals are generated, and averaging these would produce a mean around zero. The histogram in Figure 2.2 (a) shows that the time residuals are mostly falling between -0.1 and +0.1 seconds, indicating that PhaseNet and the analyst are in agreement with most of the P-phase picks. The 0.043 second delay due to filtering and decimating has been accounted for.

Figure 2.2 (b) shows how all 14,247 P-phases were distributed for each individual event. The first histogram bar shows how many P-phases were shared between the analyst and PhaseNet (i.e., the same event and the same station), the second histogram bar indicates how many P-phases were just found by PhaseNet (missed by the analyst), and the third histogram bar shows how many phases PhaseNet missed (only detected by the analyst). Overall, PhaseNet and the analysts shared 61% of the P-phases, PhaseNet found 26% of the phases that were not picked by analysts, but missed 13% of them (Figure A2.1).
For the S-Phase, 8,625 phases were found in common between the analysts and PhaseNet. Analyses were conducted for time residuals between -1.50 and +1.50 seconds. Using the same criteria as the P-phases, the mean time residual was 0.1829 seconds (using the absolute value) with a standard deviation of 0.2678 seconds. The histogram in Figure 2.3 (a) shows the time residuals are more skewed to the left, indicating PhaseNet is picking S-phases at a later time compared to the analyst and has more phases with larger time residuals compared to the P-phases. Unlike the P-phases, there was no consistent delay that we could account for with the S-phases due to decimating and filtering the data. This is because the analysts used various filters, and we do not know which filters were used on each event. Not being able to account for this may be a reason why the time residuals are skewed to the left.

Using the same criteria as the P-phases, Figure 2.3 (b) shows how all 16,346 S-phases were distributed for each individual event. Overall, PhaseNet and the analysts shared 55% of the S-phases, PhaseNet found 14% of the phases that were not picked by analysts, but missed 31% of them (Figure A2.1).

Correlation coefficients were calculated to determine if time differences were related to either magnitude or depth. In both cases, the correlation coefficients were nearly 0, suggesting that these two parameters do not systematically affect the time differences. There was also no correlation between time residual and station distance from the event location. The time residuals appear random, but we did not examine possible correlations to other factors such as time of day, weather, or source.

**Comparison of Earthquake Locations**

After the statistical analyses of PhaseNet versus analyst picks, earthquake hypocenters were compared using the two data sources. The analysis was confined to the “polygon of
interest” around Pecos, TX (Figure 2.1). Initial association and preliminary event locations were generated using SeisComp3 software (Helmholtz Centre Potsdam, 2023), and then refined using a global search-location algorithm, NonLinLoc (Lomax et al., 2000; 2009). These initial locations and phase picks were then used to perform relocations using HypoInverse (Klein, 2014). We relocated the events using HypoInverse because we had better control over the input variables. This allowed us to confirm we were conducting the one-to-one comparison study.

Both workflows use the 1-D velocity model, DB1D (Savvaidis et al., 2019) (Table 2.2). We used this simple, 1-D velocity model because TexNet uses this in their original locations.

To obtain accurate hypocenters, a station distance weight was applied. If a station was within 15 km of an event, it was fully weighted at 1.0, and a cosine taper was applied to stations that were between 15 km to 45 km. If a station was beyond 45 km, it was weighted at zero. This was done because stations further away were often residual outliers and affected the depth. Events were removed if the phase data met one of the following criteria: all stations used were at distances >10 km, the second closest station was >10 km, or if there were 2 or less S-phases used in the relocation.

Between January 2019 and May 2019, the analysts located 849 events within the “polygon of interest”, and PhaseNet located 1,067. Figure 2.4 shows the analyst locations as green dots, and the PhaseNet events as purple diamonds. The PhaseNet events had an average horizontal error of 0.721 km, while the analyst average horizontal error was 0.477 km. Error is determined by creating an error ellipsoid that is a 3 x 3 spatial covariance matrix. The matrix is rotated into the principal solution, and the errors are calculated by taking the square root of the eigenvalues (Klein, 2014). The events are very similar in trend and location. We observe three subparallel northwest-southeast trends. This direction corresponds with the orientation of
maximum horizontal compressive stress ($S_{H_{\text{max}}}$) in the area (Snee et al., 2018). The earthquake epicenters are generally coincident with the shallow normal faults mapped by Hennings et al. (2021) and Horne et al. (2022) (Figure 2.5). The only group of seismicity that does not have a fault near it is the cluster in the southeast corner of the polygon.

Various cross-sections were constructed to determine if any subsurface patterns were visible. The locations of the cross sections can be found in Figure 2.5. We observe that depths from both catalogs are in agreement, with 98% of events occurring shallower than 5 km (relative to sea level) (Figure 2.6 (a)). The PhaseNet events range from 0 km to 9 km, with 53% of events locating within the top 2 km of the basin, with an average vertical error of 1.244 km. The analyst events range from 0 km to 7 km, with 56% of events locating between 2 km and 4 km with an average vertical error of 0.721 km (Figure 2.7).

In Figure 2.6 (a), (cross-section along the strike of the seismicity in the region of interest) we observe seismicity increasing in depth towards the northwest. In Figure 2.6 (b) (cross section perpendicular to the strike of seismicity in the area of interest) we observe two parallel clusters that dip slightly to the southwest. In Figure 2.6 (c) (cross-section through the central cluster) we observe a layer of seismicity in both catalogs at ~3.5 km depth, and the PhaseNet catalog having another layer at ~1.5 km depth. We would need to include all of the locations (beyond the 5 months we analyzed) to feel confident in determining potential structures.

**DISCUSSION**

In Texas, earthquakes have been induced by oil and gas operations since 1925 (Frohlich et al., 2016), but also occur naturally (Doser et al., 1991). Recent research suggests that the current seismicity is induced by human activities (Frohlich et al., 2016 and 2019, Savvaidis et al., 2020). Most researchers associate the increase in seismicity in the Delaware Basin with shallow
wastewater disposal procedures and a small part with hydraulic fracturing (Skoumal et al., 2020; Zhai et al, 2021; Savvaidis et al, 2020; Sheng et al., 2022). Wastewater disposal occurs between 0 km – 4 km depth, with the highest volumes between 1 km and 2 km, while hydraulic fracturing occurs between 1.5 km – 4 km, with the highest volumes between 2 km and 3 km (Skoumal et al., 2020).

The PhaseNet catalog places most of its events within the top 4 km, with over 50% located in the top 2 km. This would place them in the Delaware Mountain Group, the main disposal formation (Sheng et al., 2022). The analyst catalog located over 50% of its events between 2 km and 4 km, which places them in the lower Delaware Mountain Group, Bone Springs Formation, and the upper Wolfcamp Formation (Savvaidis et al., 2022; Sheng et al., 2022). In Figure 2.6 (c), we observe a defined layer of seismicity in both catalogs at ~3.5 km depth. Waste water injection and fracking may be contributing to these events.

Our locations align with the shallow normal faults mapped in the area (Hennings et al., 2021 and Horne et al., 2022) (Figure 2.5), and match the focal mechanisms indicating northwest-southeast striking normal faulting obtained southeast of our study area (Huang, et al., 2022; Sheng et al. 2022). The shallow normal faults are between 3 – 80 million years in age, and are located throughout Reeves County (Horne et al., 2022). The faults that are in our study area are located with “moderate confidence” since their locations were informed from Interferometric Synthetic Aperture Radar (InSAR) surface deformation images or earthquake locations rather than 3-D seismic studies (Horne et at., 2022). Faults to the northwest of Pecos (~25 km) with the same orientation have high mapping confidence. These faults have a mean throw of 29 m, an average length of 3,250 m, and a dip orientation of 78° (Horne et al., 2022). These faults appear between depths 0.25 – 2 km in the Ochoan Salt, Delaware Mountain Group, and Bone Springs
Formation (Horne et al., 2022) (Figure 2.8 (a)). According to a sonic well log southeast of our study area (~40 km) (Sheng et al., 2022), these formations are slightly deeper compared to the Horne et al., (2022) survey completed northwest of our study area (Figure 2.8 (b)). If we assume the faults occur in the same formations, the depths of the faults could range between 0.25 – 3.2 km in the Pecos region. Most of our events are between 2 - 4 km in depth, so we possibly have events occurring in this faulted region.

Hennings et al. (2021) calculated deterministic fault-slip potential (DFSP) for the entirety of the Delaware Basin, and found that Reeves County (where Pecos is located) has a low DFSP of ≤ 2.5 MPa., indicating that the shallow normal faults are prone to reactivation with small increases in pore pressure. While we did not analyze wastewater disposal and hydraulic fracturing locations, depths, and other statistics to directly compare to our earthquake events to these operations, our hypocenters do occur in the basin formations where these operations are occurring (Savvaidis et al., 2020; Skoumal et al., 2020; Grigoratos et al., 2022). It is possible that these events were caused by human activities, but further investigation must be conducted.

We realize that using a 1-D velocity model may not be providing us with the best picture of what is happening in the Pecos region. This basin is very complicated, with many wave propagation and velocity issues to solve. For example, it is known from sonic log analysis that there is up to a 20% difference in vertical and horizontal velocity in the Delaware Basin (Doser et al. 1992). This anisotropy can occur due to the thin layering of sedimentary rocks, presence of over-pressured zones, or the lithologic mineral grain orientation of shales (Wild and Crampin, 1991). Since anisotropy has been observed in other parts of the basin, we assume this phenomenon could be important in this part of the basin as well. Sheng et al. (2022) have suggested that there may be up to three low velocity zones that are affecting propagation and
travel times in this area. Even with these velocity model complexities, we do feel using the DB1D velocity model (Table 2.2) was sufficient to compare earthquake catalogs derived from PhaseNet and human analyst picks. We would need to incorporate these other velocity complexities for a more in-depth study of the sources causing these events.

Lomax and Savvaidis (2019) compared earthquake locations derived from using 1-D and 3-D velocity models. For events < 12 km from their PB02 station (co-located with our PC01 station), the average focal depth using the 3-D model was 2-4 km, while the 1-D model provided average focal depths between 3–6 km, both with an error of 5 km (Lomax and Savvaidis 2019). Our relocation results provided both catalogs recording 98% of depths above 5 km, with an error of < 1.5 km (Figure 2.7). While the catalogs provide similar depth ranges, the PhaseNet and analyst relocations are shallower compared to the Lomax and Savvaidis (2019) results, and provide a smaller vertical error.

While our PhaseNet and TexNet analyst catalogs are similar, there are some notable differences. Overall, the analysts picked more S-phases compared to PhaseNet. This could be because they used various filters to pick the S-phases, while PhaseNet only used the initial preprocessing workflow (Figure 2.3 (b)). More S-phases contribute to better constrained hypocenters, and smaller depth and horizontal error. The initial analyst locations were first run through an STA/LTA picker, and then double checked by humans, which allowed them to filter out false positives. The PhaseNet events were not double-checked by humans, so there are likely false positives in its phase list. For example, in the B-B’ cross section (Figure 2.6 (b)), there are many events located at the surface, which may be attributed to noise or using a velocity model with velocities that are too high.
While running PhaseNet, we set our phase probability threshold to 0.3, which may have been too low, leading to more less confident picks in the data set. PhaseNet has performed well in other basins, such as the Raton Basin in New Mexico/Colorado (Wang et al., 2020) and in Oklahoma and southern Kansas (Park et al., 2022). Both of these studies used a probability of 0.5. Lower probability picks may have introduced more error, leading to less certain hypocenter solutions which do not cluster as well as those determined using reviewed analyst picks. In the future, we would recommend increasing the probability threshold to eliminate error in location and association.

**CONCLUSION**

The Pecos Array was able to successfully determine earthquake hypocenters in this region using both human analyst and PhaseNet phase picks. Focal depth uncertainty decreases from 5 km for the TexNet catalog to < 1.5 km (Lomax and Savvaidis 2019). The majority of events are shallow (< 4 km depth), locating at depths where both wastewater disposal and production occur. Epicenters form lineations that parallel mapped shallow normal faults (Hennings et al., 2021; Huang et al., 2022). While these trends match depths of oil and gas operations (Skoumal et al., 2020; Grigoratos et al., 2022; Savvaidis et al., 2022), further research must be conducted to identify potential mechanisms.

PhaseNet appears to be a good choice as a first-look phase picker that can later be checked by analysts. Each analyst event that was compared to PhaseNet was first run through a STA/LTA picker and then was checked by an analyst, so arrival times and phases were most likely adjusted compared to what the STA/LTA picker initially chose. While PhaseNet may not be ready to completely replace humans in the picking process, we believe it identifies phases well enough, even with the low probability thresholds employed here, to replace STA/LTA as
the initial picker in the processing workflow. PhaseNet was significantly faster compared to the STA/LTA and analyst checking process, and its results closely matched the analyst’s, but with a slightly higher error.
Table 2.1: Stations used, their locations, and months of operations during the timeframe of this study, January to May 2019.

<table>
<thead>
<tr>
<th>Station</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation (m)</th>
<th>Operation Dates (Year 2019)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PC01</td>
<td>31.4089</td>
<td>-103.5102</td>
<td>793</td>
<td>January - May</td>
</tr>
<tr>
<td>PC02</td>
<td>31.4014</td>
<td>-103.5430</td>
<td>801</td>
<td>January - May</td>
</tr>
<tr>
<td>PC03</td>
<td>31.4026</td>
<td>-103.5260</td>
<td>798</td>
<td>January - May</td>
</tr>
<tr>
<td>PC04</td>
<td>31.4124</td>
<td>-103.4845</td>
<td>786</td>
<td>January - May</td>
</tr>
<tr>
<td>PC05</td>
<td>31.4277</td>
<td>-103.5098</td>
<td>788</td>
<td>January - May</td>
</tr>
<tr>
<td>PC06</td>
<td>31.4307</td>
<td>-103.4693</td>
<td>784</td>
<td>January - April</td>
</tr>
<tr>
<td>PC07</td>
<td>31.3722</td>
<td>-103.5127</td>
<td>794</td>
<td>January - May</td>
</tr>
<tr>
<td>PC08</td>
<td>31.3844</td>
<td>-103.4955</td>
<td>791</td>
<td>January - May</td>
</tr>
<tr>
<td>PC09</td>
<td>31.3902</td>
<td>-103.4853</td>
<td>785</td>
<td>January - May</td>
</tr>
<tr>
<td>PC10</td>
<td>31.4309</td>
<td>-103.4685</td>
<td>782</td>
<td>May</td>
</tr>
<tr>
<td>PC12</td>
<td>31.3666</td>
<td>-103.4519</td>
<td>789</td>
<td>February - March</td>
</tr>
<tr>
<td>PC13</td>
<td>31.3235</td>
<td>-103.4278</td>
<td>782</td>
<td>February - May</td>
</tr>
<tr>
<td>PC14</td>
<td>31.2083</td>
<td>-103.6580</td>
<td>850</td>
<td>January - May</td>
</tr>
<tr>
<td>PC16</td>
<td>31.3552</td>
<td>-103.3341</td>
<td>811</td>
<td>January - May</td>
</tr>
<tr>
<td>PC17</td>
<td>31.3777</td>
<td>-103.2713</td>
<td>794</td>
<td>January</td>
</tr>
<tr>
<td>PC18</td>
<td>31.3371</td>
<td>-103.2994</td>
<td>813</td>
<td>January</td>
</tr>
<tr>
<td>PC19</td>
<td>31.3253</td>
<td>-103.2752</td>
<td>812</td>
<td>January - May</td>
</tr>
<tr>
<td>PC20</td>
<td>31.2804</td>
<td>-103.3227</td>
<td>829</td>
<td>January - May</td>
</tr>
<tr>
<td>PC22</td>
<td>31.2875</td>
<td>-103.2202</td>
<td>809</td>
<td>January</td>
</tr>
<tr>
<td>PC30</td>
<td>31.3419</td>
<td>-103.5947</td>
<td>822</td>
<td>May</td>
</tr>
<tr>
<td>PC31</td>
<td>31.3272</td>
<td>-103.6425</td>
<td>831</td>
<td>January - May</td>
</tr>
<tr>
<td>PC34</td>
<td>31.4542</td>
<td>-103.5952</td>
<td>815</td>
<td>January - April</td>
</tr>
<tr>
<td>PC35</td>
<td>31.4772</td>
<td>-103.6585</td>
<td>853</td>
<td>January - April</td>
</tr>
<tr>
<td>PC37</td>
<td>31.4898</td>
<td>-103.4216</td>
<td>790</td>
<td>January, March - May</td>
</tr>
<tr>
<td>PC39</td>
<td>31.4397</td>
<td>-103.5684</td>
<td>805</td>
<td>January - March</td>
</tr>
<tr>
<td>PC41</td>
<td>31.3728</td>
<td>-103.4072</td>
<td>788</td>
<td>January - April</td>
</tr>
<tr>
<td>PC42</td>
<td>31.3244</td>
<td>-103.5441</td>
<td>806</td>
<td>January - May</td>
</tr>
<tr>
<td>PC43</td>
<td>31.2648</td>
<td>-103.5589</td>
<td>811</td>
<td>January - May</td>
</tr>
<tr>
<td>PC45</td>
<td>31.2846</td>
<td>-103.3892</td>
<td>797</td>
<td>January – April</td>
</tr>
<tr>
<td>PC53</td>
<td>31.4616</td>
<td>-103.6937</td>
<td>850</td>
<td>May</td>
</tr>
<tr>
<td>PB01</td>
<td>30.9437</td>
<td>-103.7811</td>
<td>1010</td>
<td>January – May</td>
</tr>
<tr>
<td>PB02</td>
<td>31.4089</td>
<td>-103.5102</td>
<td>788</td>
<td>January - May</td>
</tr>
<tr>
<td>PB03</td>
<td>31.0838</td>
<td>-103.5139</td>
<td>871</td>
<td>January - May</td>
</tr>
<tr>
<td>PB04</td>
<td>31.1869</td>
<td>-103.2694</td>
<td>812</td>
<td>February - May</td>
</tr>
<tr>
<td>PB06</td>
<td>31.6472</td>
<td>-103.2183</td>
<td>831</td>
<td>January – May</td>
</tr>
<tr>
<td>PB07</td>
<td>31.5794</td>
<td>-103.6679</td>
<td>856</td>
<td>January - May</td>
</tr>
<tr>
<td>PB10</td>
<td>31.2836</td>
<td>-103.7546</td>
<td>858</td>
<td>January - May</td>
</tr>
<tr>
<td>PB12</td>
<td>31.2133</td>
<td>-103.9577</td>
<td>1012</td>
<td>January - May</td>
</tr>
<tr>
<td>PB13</td>
<td>31.5542</td>
<td>-103.8460</td>
<td>956</td>
<td>May</td>
</tr>
<tr>
<td>PB16</td>
<td>31.1249</td>
<td>-103.2520</td>
<td>848</td>
<td>February - May</td>
</tr>
</tbody>
</table>
Table: 2.2: Velocity model used for event relocation (Savvaidis et al., 2019).

<table>
<thead>
<tr>
<th>Depth (km)</th>
<th>Vp (km/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DTM</td>
<td>3.50</td>
</tr>
<tr>
<td>0</td>
<td>4.41</td>
</tr>
<tr>
<td>2</td>
<td>5.53</td>
</tr>
<tr>
<td>6</td>
<td>5.92</td>
</tr>
<tr>
<td>10</td>
<td>6.06</td>
</tr>
<tr>
<td>28</td>
<td>6.62</td>
</tr>
<tr>
<td>40</td>
<td>8.00</td>
</tr>
<tr>
<td>46</td>
<td>8.13</td>
</tr>
</tbody>
</table>
Figure 2.1: A map of the field area and seismic stations with Texas Roadways labeled. The blue triangles represent TexNet Stations (labeled PB** or PECS) while the red triangles represent the UTEP Pecos Array (labeled PC**). The black rectangle indicates the “polygon of interest” where we analyzed well-constrained earthquake locations. Pecos city center is in the middle of the figure where I-20, Hwy-17, and US-285 intersect. This figure shows every node that was deployed during the 14-month deployment, but the network only consisted of ~23 nodes at any given time (Table 2.1).
Figure 2.2: (a) A histogram of P-Phase time residuals. The residual is the result of the PhaseNet pick time subtracted from the analyst pick time. The delay of 0.043 seconds due to filtering and decimating has been accounted for. (b) Distribution of P-Phases for each individual event. The first histogram bar indicates P-Phases picked by both PhaseNet and an analyst. The middle histogram bar indicates phases detected only by PhaseNet. The last histogram bar indicates phases detected only by an analyst (i.e., missed by PhaseNet).
Figure 2.3: (a) A histogram of S-Phase time residuals. The residual is the result of the PhaseNet pick time subtracted from the analyst pick time. The delay of 0.043 seconds due to filtering and decimating has not been accounted for the S-picks due to the analysts using various filters to pick the S-phase. (b) Distribution of S-Phases for each individual event. The first histogram bar indicates S-Phases picked by both PhaseNet and an analyst. The middle histogram bar indicates phases detected only by PhaseNet. The last histogram bar indicates phases detected only by an analyst (i.e., missed by PhaseNet).
Figure 2.4: Locations of all earthquakes between January and May 2019. The purple diamonds are PhaseNet events, and the green circles are TexNet analyst events. The “polygon of interest” is the thick black line. These locations were obtained using HypoInverse (Klein, 2014).
Figure 2.5: The orange lines indicate the locations of the shallow normal faults mapped by Hennings et al., (2021). The green dots are locations from TexNet analyst picks, and the purple diamonds are locations from PhaseNet picks. Cross section A-A’ is oriented along the strike of the faults and includes event within the entire polygon of interest. It is 38 km long and 17 km wide. Cross section B-B’ (17 km long) is perpendicular to A-A’ and includes events within 2.5 km of the cross section line. Cross section C-C’ (3 km long) is perpendicular to A-A’ and includes events within 1.25 km of the cross-section line.
Figure 2.6: Cross sections of seismicity (January to May 2019) in the Pecos area (see Figure 2.5 for details). All depths are relative to sea level. (a) A-A’ is oriented northwest-southeast and shows all events in the polygon of interest. (b) B-B’ is oriented northeast-southwest and shows events within 2.5 km of the cross section line. Ovals indicate two clusters of events. (c) C-C’ is also oriented northeast-southwest and shows events located within 1.25 km of the cross section line. Ovals highlight smaller clusters at ~1.5 km depth and ~3.5 km depth.
Figure 2.7: The depth distribution of seismicity for the TexNet analyst catalog and the PhaseNet catalog. The gray blocks indicate the depth of seismicity for the PhaseNet catalog (relocated using HypoInverse). The solid red line indicates the depth of seismicity for the TexNet analyst catalog (relocated using HypoInverse).

Figure 2.8: (a) Cross section figure from Horne et al., (2022) of a survey line ~25 km northwest of Pecos, TX. The shallow normal faults are mapped in the top 2 km of this survey. (b) A stratigraphic column of geologic formations whose depths were derived from sonic logs. The log is ~40 km southeast of Pecos, TX. The depths are from Sheng et al. (2022).
Chapter 3: A 3-Dimensional Tomography Model of the Pecos, Texas Region of the Delaware Basin

INTRODUCTION

Since 2009, there has been an increase in earthquakes in the Pecos, Texas (TX) region of the Delaware Basin. Several earthquakes were detected in 2009, and in 2017 more than 2000 earthquakes were recorded in the region (Frohlich et al., 2019). Due to this increase in earthquake activity, the Texas Seismological Network (TexNet) funded the University of Texas at El Paso (UTEP) to deploy a temporary nodal network around the Pecos, TX area to supplement the TexNet stations from November 2018 to the beginning of January 2020 (Faith et al., 2023). The data from the nodal network can allow for more precise earthquake hypocenters, better tectonic and structural understanding, and improved velocity modeling.

Pecos, TX is located within the Delaware Basin, the western sub-basin of the larger Permian Basin (Figure 3.1). The Delaware Basin covers an area of approximately 33,500 km² throughout west Texas and southeastern New Mexico, and has a volume of approximately 170,000 km³ of sedimentary strata (Hills, 1984). The Delaware Basin formed during the Paleozoic Era (Adams, 1965). The region was originally known as the Tobosa Basin at the beginning of the Paleozoic Era, and sediments, such as the Ellenburger carbonates and Simpson Formation, began depositing (Hills, 1984). During the Mississippian, compression from the southwest reactivated Proterozoic zones of weakness, and caused the Tobosa Basin to split into the Delaware Basin, Central Basin Platform, and Midland Basin (Hills, 1984; Adams, 1965).

During the Early Permian, the Marathon orogeny thrusted rocks northward, providing the source for mud and fine sand deposition, along with intervals of deposition of thin limestones. The several thousand-meter thick Wolfcamp sediments were deposited during this time causing
rapid subsidence and an increase in compressional stress to uplift the Central Basin Platform several thousand meters (Adams, 1965). Deposition continued throughout the Middle Permian. By the Late Permian, Guadalupian sands and silts deposited in the basin. The seas started to recede, which caused evaporites to be deposited on Paleozoic sediments (Hills, 1984).

Today, the main formations comprising the Delaware Basin (from deepest to shallowest) are the Ellenburger carbonates, undifferentiated beds of shales and limestones, Wolfcamp shales, Bone Springs sands, Delaware Mountain Group sandstones, and Ochoan salts/anhydrites (Hills, 1984) (Figure 3.2). Wastewater injection is occurring in the Delaware Mountain Group, while the Wolfcamp is currently the main target production formation in this region (Skoumal et al., 2020).

Shallow normal faults, estimated to range from 3 to 80 million years in age (Horne et al., 2022) have been mapped throughout Reeves County (Figure 3.1) including in our study region, and strike parallel to earthquake locations (Hennings et al., 2021; Horne et al., 2022). These faults are consistent with focal mechanisms indicating northwest-southeast normal-faulting trends determined southeast of our study area (Sheng et al., 2022). The faults mapped within our study area have “moderate confidence” in their location because researchers used InSAR surface deformation images or earthquake locations, instead of 3-D seismic studies (Horne et al., 2022), to map them. Faults to the northwest of Pecos, TX (~25 km away) with the same orientation have high mapping confidence. Horne et al. (2022) found that these faults have a mean throw of 29 m, an average length of 3,250 m, and a dip orientation of 78º. The faults with high mapping confidence appear between depths 0.25 – 2 km in the Ochoan Salt, Delaware Mountain Group, and Bone Springs Formation (Horne et al., 2022). Hennings et al. (2021) have also mapped
basement-rooted faults (> 4 km deep) for the Delaware Basin, but the basement-rooted faults do not intersect our study area.

Huang et al. (2019) recently conducted a tomography study to update the velocity model for the entire Permian Basin that used local, regional, and teleseismic data that covered a much larger region than our study area. Since the data coverage was broader, they obtained a general velocity model that extended to depths as great as 80 km. Their best velocity results were in the top 35 km, but since their model covered such a large area with sparse station spacing, it cannot be used to determine more localized velocity changes within the Pecos region. TexNet developed their DB1D velocity model using this regional tomography study. With a denser network, localized earthquakes, and more ray paths, we hope to create a more detailed 3-D velocity model for the top 8 km in the Pecos region of the Delaware Basin using the SIMULPS2017 algorithm (Eberhart-Phillips et al., 2021; Thurber, 1993; Evans et al., 1994).

**SEISMIC DATA**

Seismic data for this study were obtained from two networks, TexNet and the Pecos Array, a temporary nodal network operated by the University of Texas at El Paso (Faith et al., 2023) (Figure 3.3 (a)). The Pecos Array was deployed from November 2018 to the beginning of January 2020. The station spacing varied from ~2 km and extended to ~10 km with increasing distance from the city center (Faith et al., 2023) (Figure 3.3 (a)). The tomography inversion used a total of 1,941 events, all occurring during the 14-month deployment, and constrained to a polygon of interest shown in Figure 3.3 (b). Events were deleted if they had less than 5 P-phase arrivals, or were located above the local maximum elevation (1.1 km above sea level). This led to the exclusion of 532 events.
The event catalog building and phase picking were completed by TexNet. TexNet’s phase picking workflow involves the following steps: 1) automatic picking using a short-term average over long-term average (STA/LTA) calculation, 2) application of an enhanced S-picker module (Grigoli et al., 2018), and 3) human analysts manually checking and correcting picks (Savvaidis et al., 2019). TexNet then used SeisComp3 software (Helmholtz Centre Potsdam, 2023) to conduct association and event location, and finally refined the locations using NonLinLoc (Lomax et al., 2000; 2009). They used the 1-D velocity model, DB1D of Savvaidis et al. (2019).

**Initial Velocity Model**

The initial minimum starting velocity model used for the tomography was calculated using VELEST (Kissling et al., 1994). VELEST could only use a maximum of 658 events in its inversion, so we carefully chose which events were used. Only the P-phases were used, and each event was required to have at least 8 phase picks with residuals of less than 0.5 s. The events ranged in depth from 0.5 km to 8.5 km, and were evenly spread out over the study area (Figure 3.4). The default VELEST damping parameters were used in the inversion.

We used two different velocity models as inputs to VELEST (Figure 3.5 (a)). One was a smoothed version of the velocity model from Sheng et al. (2022). The model had to be smoothed to remove low velocity zones due to unstable results from using the original model. Using the original Sheng et al. (2022) model produced low velocities at depths where the velocity should have been faster, and moved the hypocenters deeper on each iteration. The residuals were also higher compared to using the smoothed version. The other model was the DB1D velocity model used by TexNet (Savvaidis et al., 2019). We also ran VELEST to invert for velocity models using various depth increments of 0.5 km, 1 km, and 2 km. We conducted multiple runs using
the previous output’s velocity model until we found the model with the lowest average residual. Regardless of depth increments and starting model, each converged to similar values (Figure 3.5 (b)). These values were averaged together and influenced the starting model used in SIMULPS. The tomography program (SIMULPS) requires a 1-D continuously increasing initial velocity model, so we adjusted the averaged VELEST model to accommodate this. The final model is the black line shown in Figure 3.5 (c).

**Tomography Method**

The SIMULPS2017 inversion package simultaneously solves for earthquake relocations and the velocity structure by using P-arrival times, earthquake origin times, and locations (Eberhart-Phillips et al., 2021, Evans et al., 1994, Thurber, 1993). The inversion is formulated to be a linearized, overdetermined problem, and the velocity structure is represented as discrete points instead of blocks, that increase in velocity with depth (Thurber, 1993). SIMULPS uses approximate ray tracing. SIMULPS creates a large set of curves connecting the source to the receiver, and numerically calculates the travel time along each of these curves. The inversion iterates through these travel times to find the “fastest” curve (Evans et al., 1994), the true ray path.

For our problem, we used the P-phase arrival times of 1,941 events, weighted at 1.00 to conduct the tomography. The horizontal grid nodes have a 2 km spacing in the center where the majority of events are located; the spacing increases to 5 km, and then to 10 km spacing to the edge of the study area (Figure 3.3). For depth, grid nodes in the top 4 km have a spacing of 0.5 km, and then the depth spacing increases to 1 km at depths of 4 to 8 km. A node becomes fixed (not able to change velocity) only if it has a derivative weight sum (DWS) (a ray density value) of < 20. We weighted all stations < 75 km from an event at 1.00, and weighted at 0.00 if they
were > 85 km. For the time residual weighting, we weighted residuals < 0.75 s at 1.00, and anything > 2.00 s was weighted at 0.00. In between these values, SIMULPS applies a linear taper. We did not use S-phases because only 24% of the total S-phases had time residuals < 0.75, whereas 98% of the total P-phases had time residuals < 0.75 s. The S-phases were originally weighted at 75%, and then down-weighted again via the time residual weighting, so many S-phases were discarded in this process.

We ran a series of one-iteration inversions with a large range of damping parameters, ranging from 0 to 1000, in order to choose an appropriate damping parameter. We then plotted the data misfit versus model variance (Evans et al., 1994). The trade-off curve returned a damping parameter of 40, the value we used in this study (Figure 3.6).

**Velocity Model Results**

We were able to resolve our inversion in three iterations. Figure A3.1 and A3.2 show the Vp perturbations and velocity results of each layer. The top 3.0 km had the most changes compared to the starting velocity model, so those will be the slices discussed most in our results. Between depths of 1 to 2 km, we observe an area with positive Vp perturbations that dissipate at greater depth. The velocities range from 4.15 km/s to 4.25 km/s. At a depth of 3.0 km, velocities become more homogeneous with a value of around 4.20 km/s (Figure 3.7).

The shallow normal faults are shown on the depth slices in which they can possibly exist (Hennings et al., 2021; Horne et al., 2022) (Figure 3.8). At depth slices of 1.0 - 2 km, we observe the faults coinciding with areas where velocities are increasing, whereas at depth slices of 2.5 km and 3 km, the faults coincide with regions where the velocities are decreasing. The velocity at each slice is in the range of 4.20 km/s to 4.25 km/s.
The SIMULPS hypocenter locations did not significantly change from the original locations (Figure 3.9 (a)). Most of the seismicity is occurring shallower than 5 km, with a majority of the events between 2 to 4 km (Figure 3.9 (b)). Between depth slices of 0 km to 1.5 km and > 6.0 km, we do not observe a correlation between Vp perturbations and earthquake locations (Figure A3.3). Between depth slices of 2.0 km to 5 km, we observe that most events are located in areas with negative Vp perturbations, or they are on the edge between positive and negative (Figure 3.10).

Figure 3.11 shows the derivative weight sum (DWS) or ray density for each layer. DWS is the total number of ray paths that touch a grid node. In the shallow layers, the ray densities are greatest where the polygon of interest constrains the earthquake locations (Figure 3.11). The deeper layers show a wider area of high DWS areas due to the rays turning within these layers to reach the receivers outside of the polygon of interest.

Figure 3.12 shows the diagonal resolution element (DRE) for each grid node. The DRE can be used to determine the resolution of our model. If the DRE is higher, this means that grid node is more resolved. We observe that the areas with the highest DRE correspond to the same areas with high DWS, and follow the north-west/south-east trend of the polygon of interest. For the top 6 km, each layer has their highest DRE value in the center of the study area. At 7 km and 8 km, we observe their highest DRE values are to the south and west of the densely gridded node square; this is due to the ray paths turning within these layers as the wave moves to the receivers near the outside of the study area. The layers above 3 km and the 8 km layer all have a max DRE value > 0.14, while the remaining layers all have DRE values ≤ 0.10.
DISCUSSION

Our model is best resolved in the middle of the densely gridded node space. This area has the highest DWS, highest DRE, and has the most crossing ray paths. The velocity model for the centermost grid node (Figure 3.13 (d)) is shown as the pink line in Figure 3.13 (a). We observe the most changes from the starting model in the top 3 km. The velocity is very similar from 0.5 km to 4 km, ranging from 4.23 to 4.34 km/s. This range matches the geology of the area, which includes sandstones, shales, and interbedded limestones (Figure 3.2). At 6 km, the velocity increases to 5.2 km/s. We expect the Ellenburger carbonates to be in this depth range, and this velocity is in the range of carbonate rocks (Musset et al., 2016).

In Figure 3.13 (a), we observe a very small high velocity value of 4.31 km/s at depths 1 km – 1.5 km. Other velocity models such as Sheng et al. (2022) and Doser et al. (1992) show a high velocity zone of ~ 6 km/s around this depth, accounting for the Ochoan salt/anhydrite formation. While some sources (Musset et al., 2016; Jones, 2015) suggest the lower end of anhydrite and salt velocities is ~4.5 km/s, our model is not able to resolve this layer well. SIMULPS only allows a starting model where velocity increases with depth, so we were not able to use the original Sheng et al. (2022) or Doser et al. (1992) velocity models for this inversion. If we wanted to include the Ochoan Formation, we would have to use a different tomography program.

Hennings et al. (2021) and Horne et al. (2022) indicate the presence of shallow normal faults in the Ochoan evaporites, Delaware Mountain Group, and Bone Springs Formations ranging in depth from 0.25 – 2 km. According to a well log southeast of our study area (~50 km away from the high resolution shallow normal faults) (Sheng et al., 2022), these formations are deeper, ranging in depth from 1 – 3.2 km (Figure 3.2). If we assume the faults are in similar
formations throughout the basin, then the faults could range between 0.25 – 3.2 km in the Pecos, TX region. While our model does not show any obvious velocity patterns relating to the faults, we do observe slower velocities in areas that are directly on the shallow normal faults (Figure 3.13 (b)). The average velocity in the top 3 km is ~3% slower on a fault compared to not being on a fault. Faults with fractured rock generally have slower velocities due to fractured rock scattering energy and delaying seismic waves (Li et al., 1990). Our model at the center point was not on a fault, which may explain why it is faster closer to the surface (Figure 3.13 (b)). These shallow normal faults on average have a throw of 29 m (Horne et al., 2022), so our model may be interpreting them as a homogeneous layer.

Our final model closely resembles the “Smoothed Sheng et al. (2022)” model that was used as a starting model in VELEST (green line in Figure 3.13 (c)). They both have similar values for depths between 0.5 km to 5.0 km ranging from 4.23 km/s to 4.40 km/s. At 6 km, the Smoothed Sheng et al. (2022) model becomes faster compared to our new model. When compared to the TexNet DB1D model (navy line in Figure 3.13 (c)), we observe that only depths between 0.5 km to 2 km are similar. Velocities at depths between 2 km to 5 km are 28% slower, and depths 6 km to 8 km are on average 5% slower compared to the TexNet DB1D model. Due to these differences, the model developed by TexNet may overestimate too fast of velocities for this region of the Delaware Basin. This could become problematic because using faster velocities can affect hypocentral locations, especially depths.

Tomographic cross sections (Figure 3.14) do not show many differences compared to the 1-D models. In cross section A-A’, oriented north-south (Figure 3.14 (b)), we observe most velocity changes in the top 4 km, near the center of the densely gridded area. We observe velocities around 4.25 km/s. In cross sections B-B’, oriented east-west (Figure 10 (c)) and C-C’,
oriented northwest-southeast (Figure 3.10 (d)), we observe a zone at ~3.5 km depth that is ~12% faster in velocity compared to the area around it. At a velocity of 4.7 km/s, the area does have a high DWS and a reasonable DRE, so we expect this area to be well resolved. Since the geology here is interbedded, this area could potentially have a thicker layer of limestone compared to other areas, causing the increase in velocity.

Additional earthquake events could be added to the current catalog to improve the velocity models. First, including earthquake events that are outside of the polygon of interest, but still within our grid node area would expand the extent of data used in the tomographic inversion. Second, including earthquakes that were recorded on a nodal network operated in September and October 2020 (Faith et al., 2023) could further improve the velocity models. This network was also deployed by UTEP and compliments the original array. It includes 47 nodes deployed in September and 26 nodes in October. This would fill in ray path gaps to the north, northeast, southwest, and southeast. Lastly, the September and October 2020 network was deployed to collect eleven active source shots surrounding the region. Incorporating these shot data, especially having exact locations of the sources, could greatly improve the velocity modeling of the area, and provide ray paths at more azimuths. We could also potentially include the S-phases at 100% weight to determine if they improve our results.

CONCLUSIONS

We were able to derive P-wave tomographic velocity models for the upper 8 km of the Pecos, Texas region. We observe velocity values that are consistent with reported geology in the basin and slower velocities in regions with faults compared to areas without faults. For the area we could resolve, our new 1-D velocity model is 28% slower compared to the regional DB1D model TexNet currently uses for their earthquake locations. The velocity differences would
impact hypocentral locations. Having the most accurate locations would help determine what is causing the earthquakes.

Our results could be improved by expanding the study area to include earthquakes outside of the polygon of interest and include active source shots conducted near the region. This would provide more data, and increase the resolution for a larger area. We could also include the S-phases in the future. Including these phases would allow us to see 3-D structure more easily and may indicate areas that are more saturated than others.
Figure 3.1: Map of the Permian Basin with its geologic features, including the Delaware Basin to the west, the Midland Basin to the east, and the Central Basin Platform in between them (See Data and Resources). Our study area is labeled by the dashed red line. The location of Pecos, Texas is identified by the green triangle. The study area is located in Reeves County.
Figure 3.2: A stratigraphic column of the Delaware Basin geology. Depths of formations are from Sheng et al., (2022).
Figure 3.3: (a) A map of the stations and earthquakes used in the tomography. The red triangles are Pecos Array stations, and blue triangles are TexNet stations. The yellow squares are the grid nodes. (b) A map of the grid nodes used for the tomography. The black polygon is the “area of interest” and earthquakes are only located within this polygon. The grid nodes are more densely spaced where the earthquakes and more stations are located.
Figure 3.4: A map of the 658 high quality events used for the VELEST inversion (Kissling et al., 1994) to determine the starting velocity model. The color of the dots represents the event depth.
Figure 3.5: (a) The blue line is the DB1D TexNet (TN) velocity model (Savvaidis et al., 2019), the green line is the velocity model used in Sheng et al. (2022), the red line is the smoothed model of Sheng et al. (2022) (SS), and the black line is the final starting velocity model used in the tomography. (b) The results of the VELEST (Kissling et al., 1994) inversions using the TexNet and Smoothed Stanford starting models for various depth spacings. (c) The orange dashed line is the average of the VELEST runs. The black solid line is the final starting velocity model used in SIMULPS. The top 4 km of this model were adjusted so velocity continuously increases with depth.
Figure 3.6: The results of the L-Curve test to determine our damping parameter. The dark brown dot labeled “Damping = 40” is where we observed the “elbow” of the graph.
Figure 3.7: The Vp perturbations for depth slices 0.5 km to 3.0 km. Red colors represent areas decreasing in velocity, while blue areas represent areas increasing in velocity. The purple dotted line represents the densely gridded area.
Figure 3.8: The shallow normal faults (Hennings et al., 2021) mapped at the depth slices that they may occur at (1 – 3 km). They are overlain with the velocity (warm colors represent slower velocities, cool colors represent faster velocities), and Vp perturbations (red colors represent areas decreasing in velocity, blue colors represent areas increasing in velocity). The purple dotted line represents the densely gridded area.
Figure 3.9: (a) Earthquake locations from the original input file (blue dots) versus the relocations results from SIMULPS (red dots). (b) Earthquake depth comparisons of the original input file (blue dots) versus the earthquake depth results from SIMULPS (red dots).
Figure 3.10: The seismicity at depth slices 2 – 5 km overlain on Vp perturbations. Each slice shows earthquakes in a ± 0.5 km range of the depth slice layer name. For example, for “Depth: 2.0 km”, the figure is showing seismicity ranging from 1.5 – 2.5 km.
Figure 3.11: The ray density or derivative weight sum (DWS) for depth slices 0.5 – 8.0 km. The lighter colors represent areas with little ray density coverage. The darker colors represent areas with high ray density coverage.
Figure 3.12: The diagonal resolution element (DRE) for depth slices 0.5 – 8 km. Lighter colors represent areas with low DRE values. Darker colors represent areas with high DRE values.
Figure 3.13: (a) The blue dotted line represents the starting velocity model used in SIMULPS. The pink line is the velocity results for the center of our study area. (b) The pink line represents the velocity within a non-faulted area (and center of our study area), located at the pink square in (d). The navy line represents the velocity of an area within a shallow normal fault zone, located as the navy square in (d). (c) Comparisons of our final velocity model (pink) to the TexNet DB1D velocity model (blue), and the Smoothed Sheng et al. (2022) model (green). (d) The locations of the pink “Final Model” in (a and c), pink “Model at Center” in (b), and the navy “Model on Fault” in (b).
Figure 3.14: (a) The locations of the tomography cross sections. Black dots are grid nodes and green triangles are station locations. (b) Tomography cross section of A – A’. The location is the red polygon in (a). (c) Tomography cross section of B – B’. The location is the blue polygon in (a). (d) Tomography cross section of C – C’. The location is the yellow polygon in (a). Each cross section (b) – (d) represents the grid nodes that are inside of the polygon. The pink square in each figure is the location where the three cross sections intersect. The black dots represent earthquake hypocenters. White triangles represent stations that are within each polygon. They are located under the maximum elevation (1.1 km above sea level) because that value is determined by the station with the highest elevation. If a station elevation is lower than that maximum value, it will map beneath it.
Chapter 4: Comparing Earthquake Hypocenters from the TexNet, PhaseNet, and SIMULPS Catalogs Using the Pecos Array

INTRODUCTION

Earthquake locations can differ based on the algorithm used, the amount of P- and S-phases used to associate events, quality of phase pick times, and starting velocity model. In this chapter, I will conduct a short overview that compares the three catalogs I used in my dissertation, and to determine how the earthquake hypocenters differ.

SEISMIC USED

I will be comparing three catalogs that derived hypocenters using the Pecos Array: 1) TexNet, 2) PhaseNet, and 3) SIMULPS. The TexNet and SIMULPS catalogs will include data from all 14 months of the deployment (November 2018 – January 2020), while the PhaseNet catalog will only include data from January 2019 – May 2019 (5 months). The TexNet catalog is comprised of 2473 events, PhaseNet 1067, and SIMULPS has 1941 events.

The stations used for the TexNet and SIMULPS catalogs can be found in Figure 1.6, and the stations used in the PhaseNet catalog can be found in Table 2.1. All events were constrained to the polygon of interest in Figure 2.1. The velocity model used for the TexNet and PhaseNet locations was the DB1D model (Savvaidis et al., 2019) (Table 2.2), and the velocity model used for the SIMULPS locations was the pink line in Figure 3.13 (a).

The TexNet catalog was originally associated using SeisComp3 (Helmholtz Centre Potsdam, 2023), and then located using NonLinLoc (Lomax et al., 2000; 2009) using P- and S-phases. The PhaseNet catalog was also associated using SeisComp3, located using NonLinLoc, and then relocated using HypoInverse (Klein, 2014) using P- and S- phases. The SIMULPS catalog was relocated within the program itself using only P-phases, but used the TexNet
hypocenters and P-phase data as original input (Eberhart-Phillips et al., 2021; Thurber, 1993; Evans et al., 1994).

**Comparison of Seismicity Locations**

The epicenters of the three catalogs all show three northwest-southeast subparallel trends (Figure 4.1). They do not differ much when looking at the earthquake events in map view. There are, however, differences when looking at them in cross section (Figure 4.2). We observe that the PhaseNet catalog contains many events locating at 0 km, while the other two catalogs do not have this many events locating at this depth (Figure 4.2 (a)). The SIMULPS catalog has several events that are locating above 0 km, and the TexNet catalog only has two events locating above sea level (Figure 4.2 (b)).

For all three catalogs, a majority of all events are shallower than 6 km, with most of them falling between depths 2 – 4 km (Figure 4.3). The SIMULPS catalog has the most events locating deeper than 8 km.

**Discussion**

The epicenters for all three catalogs show three bands of seismicity striking in the northwest-southeast direction (Figure 4.1). They all show these three bands plus some diffused earthquakes at both the northwest and southeast corners. The southwest band of the PhaseNet catalog is locating closer towards the center of the area compared to the TexNet and SIMULPS catalogs (Figure 4.1 (d)). This could be due to using slightly different phase picks. We feel confident that these are real features since all three catalogs are similar.

PhaseNet hypocenters are locating 1 - 2 km shallower than the other two catalogs near the surface (Figure 4.2). This could be due to S-phase pick time differences, lack of S-phases in
the PhaseNet catalog, or the velocity model used could have been too fast. The TexNet hypocenters are also less compact compared to the other two catalogs.

In the histogram in Figure 4.3, we observe 98% of the events occurring above 6 km, indicating that most events are not occurring in the basement, but in sediments that lie above it. We observe 72% of events are between 1.5 - 4 km, placing them in the Wolfcamp, Bone Springs, and Delaware Mountain Group formations. These are the main formations where wastewater disposal and hydraulic fracturing are occurring (Skoumal et al., 2020). These are also the depths where shallow normal faults may occur in the region (Horne et al., 2022; Hennings et al., 2021).

**CONCLUSION**

In conclusion, this short comparison gives us insight into how different location algorithms, various inputs within these algorithms, quality of phase picks, using different velocity models, and using both P- and S- phases can affect hypocenter results, especially depth. Even though the inputs of these three catalogs did differ, we do see agreement. Each catalog trends in the northwest-southeast direction and aligns with orientations of shallow normal faults in the area (Hennings et al., 2021; Horne et al., 2022). We also observe many of our events between 1.5 – 4 km, the depths where wastewater disposal and hydraulic fracturing are occurring (Skoumal et al. 2020).
Figure 4.1: (a) Comparison of TexNet epicenters in blue, and PhaseNet epicenters in yellow. (b) Comparison of TexNet epicenters in blue, and SIMULPS epicenters in red. (c) Comparison of SIMULPS epicenters in red, and PhaseNet epicenters in yellow. (d) Comparison of all three catalogs in one map. TexNet epicenters are blue, SIMULPS red, and PhaseNet are yellow.
Figure 4.2: (a) Comparison of TexNet earthquake depths in blue and PhaseNet depths in yellow. (b) Comparison of TexNet earthquake depths in blue and SIMULPS depths in red. (c) Comparison of SIMULPS earthquake depths in red and PhaseNet depths in yellow. (d) Comparison of earthquake depths for all three catalogs. TexNet depths are blue, SIMULPS red, and PhaseNet are yellow.
Figure 4.3: A histogram comparing earthquake depths of the three catalogs. PhaseNet depths are the yellow bars, TexNet depths is the blue solid line, and SIMULPS depths is the red solid line.
References

Chapter 1:


Chapter 2:


Chapter 3:


Chapter 4:


Appendix

SUPPLEMENTAL FIGURES

Figure A2.1: Seismograms (a)-(d) represent examples where PhaseNet and the TexNet analysts are in agreement with their phase picks. The slight pick difference is due to the analyst picking phases on unprocessed data, while PhaseNet phase picks were performed on filtered and decimated data. Seismogram (e) is an example of a seismogram where PhaseNet and the analysts did not agree, and the processing effects do not make up for the pick differences. Seismogram (f) shows an example where PhaseNet picked a P- and S-phase, but the TexNet analysts did not. Seismogram (g) shows an example where the TexNet analysts picked a P- and S-phase, but PhaseNet did not. The seismograms shown are the filtered and decimated waveforms.
Figure A3.1: The Vp perturbations for depth slices 0.5 – 8 km. Red colors represent areas decreasing in velocity, while blue areas represent areas increasing in velocity. The purple dotted line represents the densely gridded area.
Figure A3.2: Velocity values for depth slices 0.5 – 8 km. Warmer colors represent slower velocities. Cooler colors represent faster velocities.
Figure A3.3: Seismicity at depth slices 0.5 – 8 km overlain on Vp perturbations. Each slice shows earthquakes in a ± 0.5 km range of the depth slice layer name. For example, for “Depth: 4.0 km”, the figure is showing seismicity ranging from 3.5 – 4.5 km.
Vita

Jenna Faith earned her Bachelor of Science in geology from Juniata College in Huntingdon, Pennsylvania in 2017. She joined UTEP’s doctoral program in geological sciences in fall of 2017. While at UTEP, she also completed a graduate certificate in Applied and Computational Mathematics.

While at UTEP, Jenna was the recipient of the Cearley Graduate Scholarship in Geological Sciences, Bruce Davidson Memorial Graduate Student Award, the McBride Scholarship, and a Graduate School Travel Grant. Jenna has presented her research at various meetings including the 2019 and 2021 Seismological Society of America annual meeting, the 2019 Geological Society of America conference, and the 2019 American Geophysical Union conference. In 2020, she received 1st place oral presentation at the department’s annual colloquium. She was also a “Three Minute Thesis” finalist in 2020.

Between 2017 and 2021, Jenna worked as a teaching assistant and research assistant for the Department of Earth, Environmental, and Resource Sciences. She was a lab instructor for Physical Geology, Geology for Engineers, and a Research-Based Physical Geology lab. In 2021, Jenna received a fellowship with the National Nuclear Security Administration. She spent a year working remotely with the Defense Nuclear Nonproliferation Research and Development (NA-22) office. She then received a Graduate Research Assistant position with the Earth and Environmental Sciences Division at Los Alamos National Laboratory, where she finished her degree.

Contact Information: jenna.faith21@yahoo.com