

Tracking Groundwater with the Ambient Seismic Field

a dissertation presented

by

Timothy Clements

to

The Department of Earth and Planetary Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Seismology

Harvard University

Cambridge, Massachusetts

July 2021

©2021 – Timothy Clements
all rights reserved.

Tracking Groundwater with the Ambient Seismic Field

ABSTRACT

Aquifers are vital groundwater reservoirs for residential, agricultural, and industrial activities worldwide. Tracking their state with high temporal and spatial resolution is critical for water resource management at the regional scale yet is rarely achieved from a single dataset. Here, we show that variations in groundwater levels can be mapped in space and time using perturbations in seismic velocity (dv/v). We employ the computing language Julia and cloud computing resources to make daily measurements of dv/v across California from the cross-correlation and auto-correlation of the ambient seismic field. We compare our results to estimates of groundwater from groundwater wells, precipitation and satellite measurements. dv/v reproduces the groundwater level changes that are marked by the multi-year depletions and rapid recharges typical of California's cycles of droughts and floods. dv/v correlates spatially with vertical surface displacements and deformation measured with GPS and the GRACE satellite.

Contents

1	INTRODUCTION	1
1	Groundwater monitoring using hydrology and geodesy	2
2	Measuring the seismic velocity change, dv/v	3
3	Using nonlinear elasticity to turn dv/v into a strainmeter	7
4	dv/v turned into a measure of hydraulic head	10
5	Monitoring Attenuation	12
6	Synopsis	13
2	TRACKING GROUNDWATER LEVELS USING THE AMBIENT SEISMIC FIELD	14
1	Data and Methods	16
2	Results and Discussion	18
3	AMBIENT SEISMIC NOISE CROSS-CORRELATION ON THE CPU AND GPU IN JULIA	22
1	Introduction	23
2	SeisNoise.jl Structure and Workflow	25
3	SeisNoise.jl Parallelization	29
4	Examples	34
5	Future Development	42
4	SINGLE SEISMIC STATION GROUNDWATER MONITORING	44
1	Introduction	45
2	Data	48
3	Methods	53
4	Results and Discussion	63
5	Conclusion	70
5	THE SEISMIC SIGNATURE OF CALIFORNIA'S EXTREME EVENTS	71
1	Introduction	73
2	Data and Methods	74
3	Results	75

4	Conclusions	79
6	EARTHQUAKE DETECTION AT THE EDGE	80
1	Earthquake Detection	81
2	TinyML: Machine Learning at the Edge	83
3	Data	84
4	Models for Tiny Earthquake Detection	86
5	Results	89
6	Discussion	91
7	Conclusions and Future Directions	94
7	CONCLUSIONS	96
	REFERENCES	137

Author List

Portions of Chapter 1 will be published in a forthcoming paper with Marine Denolle, Kurama Okubo and Congcong Yuan.

Chapter 2 was published in Geophysical Research Letters (*Clements and Denolle*⁵⁸).

Chapter 3 was published in Seismological Research Letters (*Clements and Denolle*⁵⁹).

Portions of Chapter 4 will be published in a forthcoming paper with Marine Denolle.

Portions of Chapter 5 will be published in a forthcoming paper with Marine Denolle, Kurama Okubo and Congcong Yuan.

Chapter 6 will be submitted as a single-author paper to Seismological Research Letters. T.C. would like to thank Prof. Vijay Reddi, Colby Banbury, and Mostafa Mousavi for discussions in developing this project.

Listing of figures

1.1	dv/v Workflow	6
2.1	Groundwater level change in San Gabriel Valley, CA during most recent drought (Fall 2012 - Fall 2016).	15
2.2	Relating seismic wavespeed temporal perturbation to groundwater levels.	17
2.3	dv/v and GPS measurements after the 2005 Rain Event.	19
2.4	dv/v and GPS measurements after the 2012-2016 Drought.	20
3.1	Data flow through objects in SeisNoise.jl. a) Ambient noise data read into SeisIO SeisData object for preprocessing and instrument response removal b) Windowed ambient noise data processed in a RawData object c) Fourier transforms (shown as spectrograms) stored in a FFTData object d) Cross-correlations stored in a CorrData object.	26
3.2	Median run times for common ambient-noise processing steps using 1 year of LHZ cross-correlations between stations BK.CMB.00.LHZ and US.WVOR.00.LHZ. Times for Julia/GPU, Julia/CPU and Python/GPU are colored white, grey and black, respectively. Median run time for conversion from SeisData to RawData in Python is not included.	33
3.3	Example output of SeisNoise.jl code example. (Top) 30-minute cross-correlations from 2019/2/3 between stations CI.SDD and CI.PER, filtered between 0.1-0.2 Hz. (Below) Stacked daily correlation for CI.SDD - CI.PER.	36
3.4	Location of 1,825 stations used from the LArge-n Seismic Survey in Oklahoma deployment operating on May 1, 2016. Average station spacing is 400m. Location of array in Northern Oklahoma is shown on inset map.	39
3.5	Stack of 1,664,400 cross-correlations from the LASSO array filtered in the frequency band 1-2 Hz and stacked in 50 m bins.	41

4.1	Location of seismometer CILJR (gold triangle) and groundwater proxies. The chartreuse rectangle denotes the 4km x 4km precipitation grid cell from PRISM dataset. The black line indicates the path of Interstate 5 through the Tejon pass. The dashed blue circles approximate limit of spatial sensitivity of CILJR autocorrelation at lag times of $\tau = 2$ and 8 seconds, respectively. The filled blue dots indicate position of nearest groundwater wells to CILJR. The orange rectangle denotes 0.25° x 0.25° grid cell from CSR GRACE/GRACE-FO RL06 version 2 Liquid Water Equivalent (LWE) dataset.	50
4.2	Cumulative annual (Oct 1 - Sep 30) precipitation levels near seismic station CILJR for wettest 10% of days (red) and remaining 90% of days (blue) with rainfall.	51
4.3	Groundwater levels in the Castac Lake Valley Basin. The monitoring and pumping wells are shown in light blue and red, respectively. The elevation of Castac Lake level shown in black.	53
4.4	North - vertical single-station cross-correlation for station CILJR. Top: Daily NZ cross-correlations from 2003-2021 for lag times $\tau \in [2, 10]$ seconds in the 2-4 Hz frequency band from CILJR with amplitude scaled by τ . Bottom: Daily power spectral density for station CILJR. White regions indicate data gaps or instrument failures.	54
4.5	dv/v in the 2-4 Hz frequency band for CILJR. Top: Average dv/v for station CILJR using equation 4.4, where the colorbar indicates c_{ik} value. Bottom: Correlation matrix for all SC component dv/v times series.	56
4.6	Cumulative deviation from the moving mean (CDMk) of precipitation using 1, 6, 24, and 92 month moving means.	60
4.7	Surface wave sensitivity for CILJR. Left: Velocity model for CILJR from SCEC Community Velocity Model. Right: Surface wave sensitivity kernels ($\frac{\partial c}{\partial V_s}$ = partials of the phase velocity with shear velocity for fixed frequency and density) for CILJR computed with Computer Programs in Seismology ¹⁰²	62
4.8	Groundwater level changes measured at monitoring well MW1 compared to modeled groundwater levels using equation 1.15.	63
4.9	dv/v compared to CDMk of precipitation, GRACE LWE and groundwater level near station CILJR. Left axis: dv/v (red dots colored by correlation coefficient to stack), scaled CDMk of precipitation (chartreuse) with correlation coefficient = -0.97, scaling factor = -42.1, and scaled GRACE LWE (orange) with correlation coefficient = -0.88, scaling factor = -20.7. Right axis: Groundwater level w.r.t surface (blue with cyan dots) for a well 10 km away from CILJR with correlation coefficient = -0.90, scaling factor = -79.3. Bottom left axis: Cumulative annual water year precipitation (Oct 1 - June 1) for PRISM grid cell containing station CILJR.	64

4.10	dv/v compared to CDMk of precipitation (chartreuse), elastic model of Talwani et al. ²⁵⁷ (orange) and model of Sens-Schönfelder and Wegler ²³⁰ (SSW06 - blue) at station CI.LJR. All models have been scaled by fitting coefficients.	66
4.11	Hysteresis curve for dv/v and CDMk. (a) Hysteresis between 2008-2016. Transition period starts around 2012-2013 when dv/v increases above previous high. (b) Hysteresis between 2014-2021. Slight recovery in groundwater levels suggests recoverable deformation.	67
4.12	Cumulative deviation from the moving mean of precipitation and average $\dot{v}/v \sim \Delta Q_C^{-1}$ measured on EN and NZ components of CI.LJR for lag times $\tau \in [5, 10]$ seconds in the 2-4 Hz frequency band.	68
5.1	Location of all seismometers (≈ 700) used in this study. Time of observations is 1999-2021.	75
5.2	Changes in GRACE Liquid Water Equivalent and dv/v across California in 2-4 Hz frequency band. (a) Change in dv/v measured between January 2005 and May 2005 (scatter points) and GRACE LWL (shading) over the same time frame. (b) Change in dv/v and GRACE LWL between September 2011 and September 2016.	77
5.3	Representative dv/v time series. (a) dv/v at CI.NJQ compared with surface level of nearby Alisal Reservoir and stream flow from Alisal creek. (b) dv/v at CI.RXH and elevation of the Salton Sea. Dashed vertical lines are timing of nearby earthquakes. (c) dv/v and surface temperature change at CI.SAL. Dashed line indicates timing of 2010 Mw 7.2 Baja Earthquake. (d) dv/v at CI.WES, groundwater level at USGS well 324603115480501 and root-mean-square E-N displacement at GPS station P494 after 2010 M_w 7.2 Baja Earthquake.	78
6.1	Velocity time-series of seismic event recorded on the East-West (E), North-South (N) and Vertical (Z) components of a three-component high-gain broadband seismometer. Red and blue bars and vertical lines denote the two-second window and exact arrival time of P-wave and S-wave, respectively.	82
6.2	Locations of sensors used in this study. Blue triangles represent sensors providing earthquake and/or noise waveforms, whereas red triangles represent sensors providing only noise data.	84
6.3	Representative two-second windows of the vertical component of noise, P-wave, and S-wave, respectively, after pre-processing.	86
6.4	Model architecture for Deep Neural Network with two fully-connected layers of varying size and one output layer.	87
6.5	Model architecture for Convolutional Neural Network with two convolutional layers of varying number of channels.	88

6.6	Model architecture for Depthwise-Separable Convolutional Neural Network with two DS-CNN layers of varying size.	89
6.7	Accuracy vs. latency trade-off curve for suite of fully-connected (DNN), convolutional (CNN) and depthwise-separable convolutional (DS-CNN) models. Gray dashed line is cutoff for acceptable models, shown at latency of 0.1 seconds or 10 inferences per second.	90
6.8	Accuracy vs. number of trainable parameters trade-off curve for suite of fully-connected (DNN), convolutional (CNN) and depthwise-separable convolutional (DS-CNN) models.	91
6.9	Precision vs recall trade-off curve for highest performing convolutional model on test data set. Colorbar denotes probability threshold above which test data is classified as P-wave or S-wave. Horizontal dashed line at precision = 0.25 is threshold for which network has no skill, as P-waves and S-waves both represent 25% of the test set.	92
6.10	Normalized confusion matrix for highest performing convolutional model on test data set (N = 62, 748). Total number of P-wave, S-wave, and noise samples are 15,687, 15,687 and 31,374, respectively. Upper-right value has been rounded from 0.004 to 0. for display purposes.	93

THIS THESIS IS DEDICATED TO THE TAX PAYERS OF SARATOGA SPRINGS, NEW YORK, THE STATE OF NEW YORK, AND THE UNITED STATES FOR FUNDING MY EDUCATION AND SCIENCE IN GENERAL.

Acknowledgments

First and foremost, I would like to thank my advisor, Marine Denolle, who has been a constant source of encouragement to explore big, challenging scientific problems. I thank my committee members Jim Rice, Brendan Meade and Kaighin McColl for shaping the way I think about hydrology, earthquakes and computing. I am very grateful to the academic and non-academic staff at EPS for making life easier for us graduate students - Sarah Colgan, Aimee Smith, Danielle Da Cruz, Maryorie Grande, Summer Smith, Annika Quick, and Esther James. I would not be studying seismology without the mentors and excellent teachers I had at Cornell - Rowena Lohman, Matt Pritchard, Katie Keranen, Geoff Abers, Terry Jordan, Rick Allmendinger, Sue Kay, Dana Peterson, Diego Quiros, and Muawia Barazangi. Thank you to Grant Garven for teaching me field hydrology and assisting in our project that sadly did not come to fruition. I am grateful to my fellow graduate students in the Denolle group - Jiuxun Yin, Seth Olinger, Natasha Toghramadjian, Congcong Yuan, for their support, questions and discussions outside of science and also the post-docs - Loic Viens, Chris Van Houtte, Zhitu Ma, Chengxin Jiang, Xiaotao Yang, Kurama Okubo, and Laura Ermert, who have been both a great resource to discuss science and enjoy time with at conferences. Many long hours were had in Geomuseum 200B - thank you to my office mates Harriet Lau, Blake Hodgkin, Tamara Pico, Sophie Coulson, Marisa Borreggine, Jocelyn Fuentes and Ben Thompson for many great conversations and after work drinks. I am in debt to the HGWISE board for making a space for thorns and roses. The DIB has made EPS a more inclusive and resilient place to work and study. The best part of grad school has been my G1 squad - James Mueller, Naomi Wagner, Aleyda Trevino, Franklin Wolfe, Yaya Ku, Ana Camilla Gonzalez-Nayeck, Sophie Coulson, Tia Scarpelli and Colleen Golja, and living with my best friends Franklin Wolfe, Judy Pu, Peter Davis, Brenna Andreozzi, and Colin Keil. Finally, I thank my family - my mom, my dad, my sister and Julia, for their constant support and encouragement throughout grad school.

1

Introduction

I frequently hear music in the very heart of noise.

George Gershwin

Groundwater is the largest reservoir of fresh water globally, supplying one third of the water used for residential, agricultural, and industrial use^{4,74}. In the last half century, over-withdrawal of groundwater has led to subsidence and a loss of groundwater storage in numerous aquifers across

the world^{134,90}. This is especially problematic for more than two billion people worldwide that live farther than 5 km from a source of surface fresh water¹³⁷. Compounding past and present declines in groundwater is the role of climate change, which will lead to more frequent and pronounced dry and hot years, along with stronger extreme precipitation events in places like California^{251,259,70}.

Here, I propose two seismic proxies, the relative change in seismic velocity, dv/v , and the change in coda attenuation, ΔQ_c^{-1} , obtained from the cross-correlation of ambient seismic noise, to monitor groundwater levels, specifically hydraulic heads. These new proxies allow for groundwater monitoring at a range of spatial and temporal scales, from meters and minutes to hundreds of kilometers and decades. This thesis builds on fundamental advances in theory, laboratory measurements, seismic observations, and computing power that have allowed ambient seismic monitoring to become possible in the last two decades¹⁷⁵. To do so, I develop relations between our seismic measurements and hydraulic heads, using both an established theoretical framework and an empirical approach. To scale the analysis from the single station to the network scale, I introduce new methods for applying high-performance and cloud computing to seismic data. This thesis is organized into seven chapters, including this introduction, where I develop the theoretical foundation to relate seismic and hydrological observations, and a short conclusion.

I GROUNDWATER MONITORING USING HYDROLOGY AND GEODESY

Groundwater monitoring is difficult due to a lack of knowledge of the subsurface and a sparsity of groundwater measurements¹³⁴. Conventionally, hydrologists map groundwater recharge, flow into the saturated zone, across an aquifer using a network of monitoring wells¹⁵⁷ and the groundwater fluctuation method¹⁰⁰. The method converts changes in hydraulic head $h = \frac{P}{\rho_0 g} + z$, where P is pressure, ρ_0 is the density of water, g is the gravitational acceleration at the surface, and z is a reference datum, to recharge using the specific yield, S_y , which relates the volume of water released

from storage per unit decline of head Δb ¹⁰⁰. The recharge, R , at each point in an aquifer is then, $R = S_y \Delta b / \Delta t$, where t is time. The groundwater fluctuation method is only as precise as the sampling frequency of measurements and the spatial density of groundwater wells¹⁰⁰.

A second more sophisticated approach is the numerical simulation of Darcy's law⁹² in a poroelastic medium, which requires knowledge of the hydraulic conductivity K , hydraulic gradient $\Delta b / \Delta l$, and the area upstream of the measuring site¹⁰⁰, and gives a more accurate model of groundwater flow. Numerical models are limited by the imprecise knowledge of K ²²³, which can vary over 14 orders of magnitude over short spatial scales⁸⁸.

Groundwater extraction induces surface elevation changes that are sometimes dramatic. In particular, a majority of areas experiencing rapid subsidence in the United States are correlated with excessive groundwater extraction⁸⁹. When groundwater levels are lowered, pore pressure declines, and the pore space compacts due to a higher effective stress in the sediment matrix²⁶⁴. The manifestation of this compaction at depth is subsidence at the surface. In recent years, geodetic methods, such as Global Positioning System (GPS) and Interferometric Synthetic-Aperture Radar (InSAR), have shown promise to infer groundwater levels by monitoring land subsidence⁹⁰. Surface displacements measured by GPS provide high temporal but sparse spatial resolution of groundwater level changes^{18,130,118}, while those measured by InSAR provide high spatial resolution but limited temporal resolution^{130,91,47}. Similarly, gravity measurements from the GRACE satellite are sensitive to water mass changes, but only at large wavelengths, and suffer non-uniqueness between water mass and aquifer depths^{210,292}.

2 MEASURING THE SEISMIC VELOCITY CHANGE, DV/V

Seismic travel-time measurements provide indirect measurements of subsurface structure and mechanical properties, such as density, elastic constants, or fluid saturation. Travel-time tomography

using earthquake body waves (P and S -waves)⁸⁰ or noise correlation functions²³³ can retrieve lateral or vertical *static* variations in subsurface properties, while repeated or continuous travel-time measurements can recover *temporal* changes in subsurface structure, properties, or fluids^{62,198,293}. This thesis will consider temporal changes to the subsurface, with particular emphasis on fluids. When attempting temporal monitoring, we must consider both the seismic source and seismic phase with which we want to explore the subsurface. Seismic sources can be divided into active, such as earthquakes¹⁹³, air guns¹⁹⁸, electric pulses²⁹³, explosions¹⁷⁷ or oscillators⁶², or passive sources, such as ocean waves^{285,104,16}, wind¹⁰⁵ or background anthropogenic activity that generates emergent waves^{200,72}. Travel-time measurements can be made with impulsive, ballistic waves or multiply-scattered, coda waves. Ballistic waves, usually from earthquakes, have high signal to noise ratio, but uncertainties in the origin time and relative infrequency of the earthquakes hamper their use for travel-time measurements. Coda waves take a more circuitous path between the source and receiver than ballistic waves, scattering off heterogeneities in the Earth¹⁰. Coda waves thus sample a broader volume than direct, ballistic waves. Scattering reduces the sensitivity of coda waves to the original seismic source, which increases their sensitivity near the receiver⁷³.

If Earth material property changes, a travel-time difference accumulates in late coda waves, with phase shifts dt increasing proportionally with phase lags t . Waves in the late coda are thus more sensitive to small perturbations in velocity than ballistic sources. With the assumption that there is a homogeneous velocity change in the sampling medium, the relative time delay in the coda, dt/t , is related to the relative change in seismic velocity, dv/v , by $dt/t = -dv/v$. Recent work has shown that this relation holds for many realistic scenarios of velocity perturbation^{182,184,298}.

*Poupinet et al.*¹⁹³ measured seismic velocity variations near the Calaveras Fault, CA by comparing the relative delay of coda waves between doublet events, repeating earthquakes with similar epicenters and mechanisms. The “doublet technique”, also called Moving Window Cross Spectrum (MWCS), has been used to monitor variations in seismic velocity in fault zones^{193,73} and volca-

noes^{197,109}. Using earthquakes as seismic sources for continuous monitoring is limited by the need for repeated earthquakes, which only occur at major fault systems¹⁷³, volcanoes²³⁴ and sites of fluid injection²². Active sources can measure changes in crustal velocity with precision of $10^{-3} - 10^{-4}$ from travel-time measurements^{299,293,237}. These methods have not seen wide adoption due to their limited observation time ($t < 1$ year) and short source-receiver baseline ($L \leq 1 - 2\text{km}$).

Ambient seismic waves continuously sample the mechanical properties of the near-surface. While *Aki*⁶ hinted that information about the subsurface could be determined through the spatial autocorrelation of the ambient field, it was not until laboratory measurements in the field of acoustics²⁸² and theoretical studies^{67,278} showed that the cross-correlation of a diffuse field yields the elastodynamic Green's function of the solid material. *Campillo and Paul*⁴³ and *Shapiro and Campillo*²³² were the first to extract fundamental surface wave arrivals (Green's functions) from the cross-correlation of coda waves and ambient noise, respectively, recorded at pairs of seismometers. In the real Earth, where ocean storms and waves^{16,158,104}, atmospheric disturbances¹⁰³ and cultural noise^{158,71,200} generate the ambient field, ambient noise cross-correlation recovers a band-limited approximation to the Green's function between a pair of seismometers^{267,279}.

Thanks to ambient noise seismology, seismometers are now considered as continuous “virtual sources” and receivers, obviating the use of repeated earthquakes, explosive sources or piezoelectric transducers. Ambient noise cross-correlations capture both direct waves and multiply scattered waves. Similar to the coda of earthquakes, coda waves from ambient noise cross-correlation are greatly sensitive to small perturbations in the medium of propagation²⁴². *Pacheco and Snieder*¹⁸⁸ suggested that local perturbations in seismic velocity could be monitored through tracking of temporal changes in the coda of ambient noise cross-correlations. *Sens-Schönfelder and Wegler*²³⁰ were the first to apply “Coda Wave Interferometry” to ambient noise cross-correlations to continuously monitor the change in seismic velocity, dv/v , at Merapi volcano, Indonesia. They found a seasonal change in dv/v due to precipitation.

The workflow to extract dv/v from ambient noise is described in Figure 1.1. First, the ambient seismic field is recorded at a pair of seismic stations, A and B . Seismic waveforms are then pre-processed, cross-correlated, and stacked temporally. This process is repeated over a time period that can span minutes to years, depending on the application. Then, phase delays, dt/t , are measured between each individual correlation function and a reference correlation, which allows one to construct a time series of dv/v . This workflow is performed over all combinations of seismometers pairs and station components (north-north, north-east, north-vertical, etc.). The average dv/v time-series over all stations and components is often reported for a synoptic view of the velocity changes.

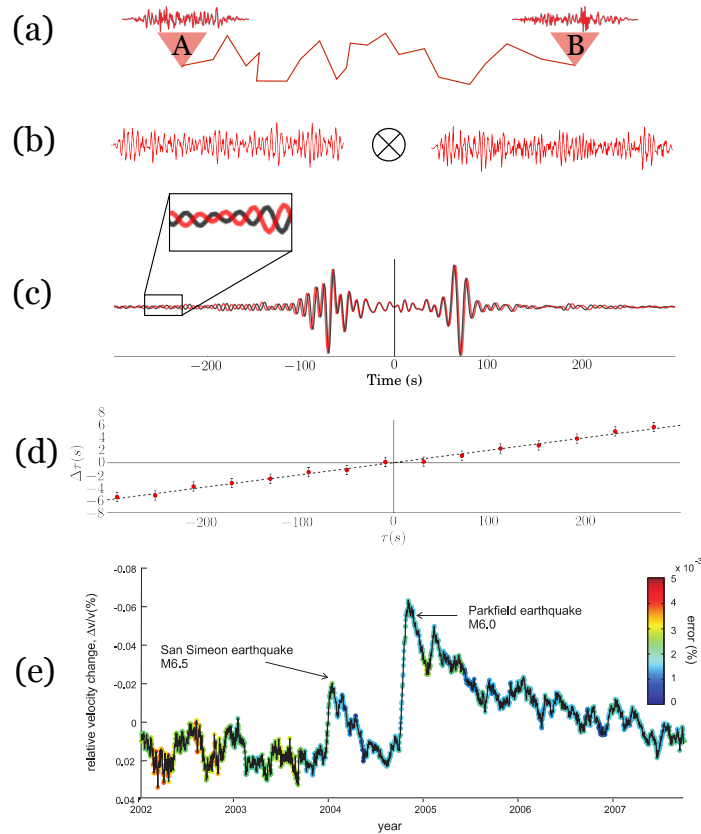


Figure 1.1: dv/v workflow. (a) Record ambient seismic noise at pair of seismometers. (b) Pre-process and cross-correlate ambient noise. (c) Phase delay between two (red and black) noise cross-correlation functions. (d) Time delay, dt/t , measurement. (e) dv/v time series near Parkfield, CA from Brenguier et al. ³³.

dv/v has been widely used in recent years to study the dynamics of Earth's crust in response to earthquakes^{33,35,287,254}, volcanic eruptions^{34,165,183,75} and ice sheet melt¹⁶⁶. dv/v is known to vary seasonally due to precipitation^{230,160,268,276}, air temperature^{160,268,106}, freeze-thaw of permafrost¹¹⁶, and decennially due to climatic forcing¹⁴⁰. Changes in seismic velocities are thus a signature of the mechanical response of the Earth material to stresses. I will now show that dv/v is closely related to volumetric strains.

3 USING NONLINEAR ELASTICITY TO TURN dv/v INTO A STRAINMETER

Lab measurements have shown that uniaxial stress¹⁷⁸ and saturation¹⁷⁹ can change seismic velocity. In low porosity rocks with high aspect ratio, penny-shaped cracks, such as granites, an increase in saturation (a decrease in effective stress) increases V_p ¹⁷⁹. In this case, when a fluid fills the cracks, the effective compressibility of the rock increases more than the density, increasing V_p . The opposite is true for sedimentary rocks with rounded pores, where the change in seismic velocity scales linearly with effective stress⁸¹. One proposed model for the velocity change in fluid-saturated sedimentary rocks is that increasing pore pressure decreases the area of grain contacts and thus lowers effective moduli, which in turns decreases seismic velocity⁵³. This relationship between effective stress and seismic velocity is especially useful when considering groundwater aquifers, which are composed of sedimentary strata and highly fractured rock.

Laboratory observations of variations in seismic velocities can be modeled using nonlinear elastic theory. *Hughes and Kelly*¹¹¹ derived the third-order equations for the velocities of elastic compression and shear waves under hydrostatic pressure p , given by,

$$\rho V_p^2 = \lambda + 2\mu - \frac{p}{3K}(6l + 4m + 7\lambda + 10\mu) \quad (1.1a)$$

$$\rho V_s^2 = \mu - \frac{p}{3K}(3m - \frac{1}{2}n + 3\lambda + 6\mu) \quad (1.1b)$$

where l , m , and n are Murnaghan's third-order elastic constants¹⁷¹, K is the bulk modulus, ρ is the density of rock, and λ and μ are the Lamé parameters or second-order elastic constants. The third-order elastic constants for Earth materials have negative values and are 1-2 orders of magnitude larger than the second-elastic constants, and do not correlate with physical properties, such as density, porosity, or strength²⁹⁰. *Toupin and Bernstein*²⁶⁵ derived the relative change in compressional and shear velocity with respect to a change hydrostatic pressure as,

$$\rho \frac{\partial V_p^2}{\partial p} = \frac{8n + 10m + 7\lambda + 10\mu}{\lambda + 2\mu} \quad (1.2a)$$

$$\rho \frac{\partial V_s^2}{\partial p} = \frac{4n + 3m + 3\lambda + 6\mu}{\lambda + 2\mu} \quad (1.2b)$$

From equation 1.2, I note that the slope $\partial V^2 / \partial p$ is related to a ratio of the third-order elastic constants to the second-order elastic coefficients. This ratio is referred to as the acoustoelastic parameter β_{ij} in the literature^{82,131,190,226}.

For a uniaxial stress, there are five unique acoustoelastic constants, β_{ij} , depending on the direction of wave propagation (parallel or perpendicular to the applied stress) and direction of particle motion⁸². E.g. β_{11} is the acoustoelastic coefficient for a compressional wave propagating parallel to the applied stress, β_{12} is the acoustoelastic coefficient for a shear wave propagating parallel to the applied stress, etc.. When measuring β_{ij} with coda waves, it is difficult to separate the individual components of β_{ij} , as compressional and shear waves equipartition under multiple scattering²⁸³.

Experimentalists thus often report the ensemble averaged acoustoelastic coefficient, $\beta = \langle a_{ij}\beta_{ij} \rangle$, where $\sum a_{ij} = 1$ and $\langle \rangle$ denotes an ensemble average^{138,190,290}.

The one-dimensional stress-strain relationship containing nonlinear effects can be reformatted with β as,

$$\sigma = M(\varepsilon + \beta\varepsilon^2 + \dots) \quad (1.3)$$

where σ is stress, M is the second- and third-order elastic modulus, given by 2 and 3 independent components, respectively, for an isotropic material¹⁸⁷, and ε is strain. In this case, β can be expressed in terms of the Murnaghan moduli as,

$$\beta = \frac{3}{2} + \frac{l + 2m}{\lambda + 2\mu} \quad (1.4)$$

Experimental values for β vary widely based on the materials but is in general a large negative values. Reported values for steel are around -10^0 ¹¹¹, concrete in the range of -10^1 to -10^2 ^{226,138,236,190,301}, Barre granite in the range of -10^2 to -10^3 ¹⁷⁹, marble around -10^3 ¹²⁵, and Fountainbleu sandstone around -10^4 ¹²⁵. Under a nonlinear elastic rheology, the local sound velocity is given by *Ostrowsky and Johnson*¹⁸⁷ as,

$$v = \sqrt{\rho^{-1}d\sigma/d\varepsilon} \approx v_0(1 + \beta\varepsilon + \dots), \quad (1.5)$$

where v and v_0 are the perturbed and unperturbed velocities, respectively. The change in velocity $\frac{\Delta v}{v} = \frac{v-v_0}{v_0}$ due to a hydrostatic stress, σ_{kk} , as a function of the volumetric strain, ε_{kk} , then becomes,

$$\frac{\Delta v}{v} = \beta\varepsilon_{kk}. \quad (1.6)$$

Recently, a number of authors have measured β values due to tidal strains using ambient noise

cross-correlation. *Mao et al.*¹⁵³ found a value of $\beta = -2 \times 10^3 - 2 \times 10^4$ for tidal strains with differing tidal periods at Piton de la Fournaise (PdF) volcano, La Réunion. *Sens-Schönfelder and Eulenfeld*²²⁹ found $\beta = -1.6 \times 10^4$ for tidal strains in the Atacama desert. *Takano et al.*²⁵⁵ found $\beta = -6.9 \times 10^4$ for tidal strains at the foot of Mount Iwate, Japan.

4 dv/v TURNED INTO A MEASURE OF HYDRAULIC HEAD

Here, I attempt to determine the effect of an increase in groundwater level on seismic velocity using poroelastic theory and nonlinear elastic theory. The constitutive relations for an ordinary isotropic, linearly elastic solid are,

$$2G\varepsilon_{ij} = \sigma_{ij} - \frac{\nu}{1 + \nu} \sigma_{kk} \delta_{ij}, \quad (1.7)$$

where ε_{ij} is the strain tensor, σ_{ij} is the stress tensor, δ_{ij} is the Kronecker delta, G is the shear modulus, and ν is Poisson's ratio. Poroelastic theory augments the linear elastic constitutive relation by adding the contribution of pore pressure, p , and the change in fluid mass content per unit volume, m . Following the results of *Rice and Cleary*²⁰², the poroelastic constitutive relations are,

$$2G\varepsilon_{ij} = \sigma_{ij} - \frac{\nu}{1 + \nu} \sigma_{kk} \delta_{ij} + \frac{3(\nu_u - \nu)}{B(1 + \nu)(1 + \nu_u)} p \delta_{ij} \quad (1.8a)$$

$$m - m_0 = \frac{3\rho_0(\nu_u - \nu)}{2GB(1 + \nu)(1 + \nu_u)} \left(\sigma_{kk} + \frac{3}{B} p \right) \quad (1.8b)$$

where ν_u is the “undrained” Poisson's ratio, B is Skempton's coefficient, $m - m_0$ is the change in fluid mass content per unit volume, ρ_0 is the density of the pore fluid and p is the pore fluid pressure. I now consider the effect of precipitation over a wide area on groundwater levels under undrained conditions, where there is no fluid flow in response to stress. I argue later in Chapter 4 that considering undrained conditions, where $m - m_0 = 0$, and neglecting the drained response

is reasonable. I follow *Roeloffs*²¹¹ to derive a relation between hydraulic head Δh , strains, and dv/v . I start with the definition of the Skempton's coefficient, which relates a change pore pressure, p , to isotropic or volumetric stress σ_{kk} ²³⁸,

$$p = \frac{-B\sigma_{kk}}{3}. \quad (1.9)$$

Using equation 1.7, we can recast equation 1.9 in terms of the pore pressure due to volumetric strain, ε_{kk} , as,

$$p = -\frac{2GB}{3} \frac{1 + \nu_u}{1 - 2\nu_u} \varepsilon_{kk}, \quad (1.10)$$

where I note that a change of pore pressure, Δp , for a given change in groundwater level Δh , is given by

$$\Delta p = \rho_0 g \Delta h \quad (1.11)$$

where g is the gravitational acceleration at the surface. Substituting equation 1.11 into equation 1.10 shows that a change in groundwater level is linearly related to the change in volumetric strain, ε_{kk} , as,

$$\Delta h = -\frac{2GB}{3\rho_0 g} \frac{1 + \nu_u}{1 - 2\nu_u} \varepsilon_{kk}. \quad (1.12)$$

Equation 1.12 is similar to the one found by *Riley*²⁰⁶ for relating the compaction of an aquifer due to the instantaneous lowering of hydraulic head. The coefficient of proportionality between Δh and ε_{kk} in the case of compaction is given by the skeletal specific storage³⁹ S_{sk} , where,

$$S_{sk} = \frac{3\rho_0 g(1 - 2\nu)}{2G(1 + \nu)}. \quad (1.13)$$

Substituting equation 1.6 into equation 1.12 then gives a relation for the change in seismic wave speed as a function of change in groundwater level,

$$\frac{\Delta v}{v} = -\frac{3\rho_0 g}{2GB} \frac{1-2\nu_u}{1+\nu_u} \beta \Delta h \quad (1.14)$$

and in its reduced form,

$$\frac{\Delta v}{v} = -\frac{S_{sk}\beta}{B} \Delta h \quad (1.15)$$

5 MONITORING ATTENUATION

Thus far, I have addressed the *phase*-sensitivity of seismic waves to subsurface perturbations. The *amplitudes* of seismic waves are also sensitive to the anelastic regime. Seismic wave amplitudes decay at a faster rate than what is predicted by geometric spreading - a process we call attenuation. Attenuation takes two forms: scattering and intrinsic absorption. Scattering is an elastic process in which small heterogeneities generate incoherent phases⁸. Scattering redistributes, rather than removes, energy. Intrinsic absorption is an anelastic process in which kinetic energy is converted into heat²²⁰. Attenuation can be measured on the amplitude of distinct seismic wave amplitudes or on the decay of coda waves, the later of which is called codaQ or Q_c^{-1} ^{7,10}.

CodaQ measurements were popular in the 1980's and 1990's, when it was thought that temporal changes in Q_c^{-1} could be used to predict earthquakes⁹. *Chouet*⁵⁰ made the first measurements of codaQ in central California, where he found an increase in Q_c^{-1} over the period of a year but could not find a satisfactory theoretical explanation for the changes. *Febler et al.*⁸⁵ observed an increase in codaQ prior to the 1981 eruption of Mount St. Helens, suggesting the role of fluids in codaQ variations. Overall, it is expected that codaQ may vary by as much as 50%^{121,220}.

Ambient noise cross-correlations also carry the signature of codaQ. *Sens-Schönfelder and We-*

gler²³⁰ and Wegler and Sens-Schönfelder²⁸⁶ measured reasonable values of coda Q in ambient-noise cross correlation and single-station correlations, respectively. Hirose *et al.*¹⁰⁷ calculated the mean free path length using the envelope decay of ambient noise correlation coda windows. Soergel *et al.*²⁴⁴ measured coda Q on inter-station cross correlations at the regional scale and found spatial variations that were consistent with tectonic structure. Soergel *et al.*²⁴⁴ needed at least 200 days of stacking for sufficient stability in the measurements. van Dinther *et al.*²⁶⁹ observed fivefold variations in Q_c^{-1} across the North Anatolian Fault using ambient noise correlations.

6 SYNOPSIS

Now that we have the foundations to understand the seismic measurements and their relation to hydrological conditions, we turn to practical application in the California. In the second chapter, I present results from monitoring groundwater fluctuations in the San Gabriel Valley, CA due to drought from 2000 - 2018⁵⁸. In the third chapter, I develop a framework for accelerating ambient noise cross-correlations using the computing language Julia and high-performance computing resources⁵⁹. In the fourth chapter, I develop a framework for using a single seismometer to monitor groundwater level changes. I compare my results to groundwater well measurements, precipitation measurements, and gravity measurements from the Gravity Recovery and Climate Experiment (GRACE)²⁰⁹. In the fifth chapter, I apply the methods developed in the third and fourth chapters to monitor groundwater levels across the state of California over the last two decades. In the final chapter, I show the potential of using internet-of-things (IoT) devices for seismology at the edge.

2

Tracking Groundwater Levels Using the Ambient Seismic Field

Too many sins, I'm runnin' out

Somebody send me a well for the drought

Kendrick Lamar Duckworth

This study⁵⁸ presents the perturbations in seismic velocity (dv/v) in the San Gabriel Valley (SGV), Eastern Los Angeles County, California. The SGV contains three unconfined, urban aquifers: the San Gabriel, the Puente, and the Raymond Basins. The east-northeast-striking Raymond Fault acts as a barrier to flow between the Raymond and San Gabriel Basins, while the San Gabriel and Puente Basins are hydraulically connected^{41,295,151}. Water-bearing sediments reach a maximum thickness of 1,200 m in the central part of the SGV⁴¹. The SGV Basin is recharged by a combination of infiltration from rainfall, runoff from the San Gabriel Mountains, stormwater capture, and imported water from the State Water Project.

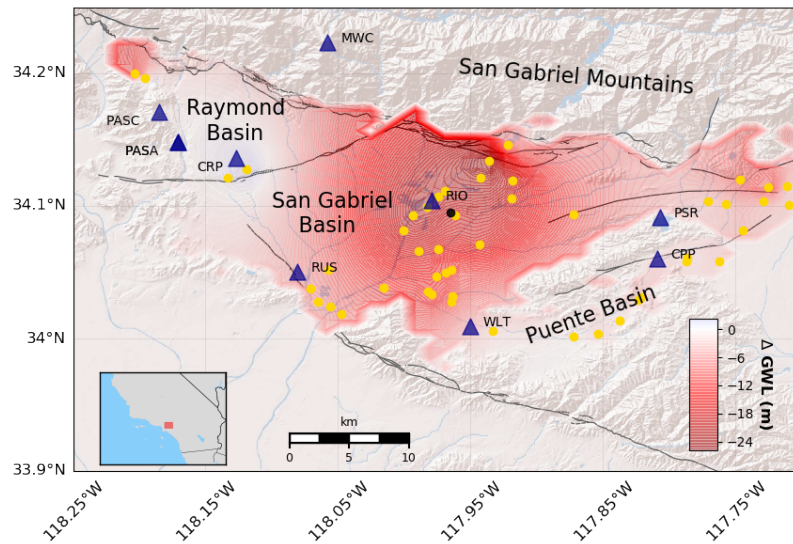


Figure 2.1: Groundwater level change in San Gabriel Valley during most recent drought (Fall 2012 - Fall 2016). Seismic stations are shown as blue triangles, and groundwater wells are shown as yellow circles. Black circle indicates the position of the Baldwin Park Key well.

We consider changes in SGV groundwater in the period Jan 2000 - Jul 2017. This period is notable for having three major droughts in southern California (2002-2004, 2007-2009, and 2012-2016)⁴². During droughts, groundwater supplies over 40% of water demand in the SGV¹⁵¹. By the

end of the 2012-2016 drought, groundwater levels dropped 18 *m* in the SGV in the Baldwin Park Key Well (Figure 2.1), reaching all-time low levels in Oct 2016. Even with above average precipitation in the winter of 2016-2017, groundwater levels only recovered 1.7 *m* in the SGV basin due to uptake by drought-parched soil¹⁵¹.

1 DATA AND METHODS

AMBIENT SEISMIC CROSS-CORRELATION

We use continuous data from broadband vertical component seismometers in the California Integrated Seismic Network (CI) from Jan 2000 - Jul 2017 (Fig 2.1). All raw waveforms are down-sampled to 20 *Hz*, demeaned, and detrended. Hourly windows of raw data with maximum amplitude greater than ten times the standard deviation of the daily trace are discarded. We apply one-bit normalization and whiten in the frequency domain from 0.05 to 4 *Hz*^{21,140}. Daily time series are segmented into 1-hour windows with 30 minutes of overlap between successive windows and cross-correlated using the MSNoise package¹³⁹. Noise cross-correlation functions (NCFs) are computed for all station pairs in all available date ranges. Instrument corrections are applied after cross-correlating. A daily NCF is formed by stacking all hourly NCFs from each day.

Daily changes in seismic velocity are computed using the Moving Window Cross-Spectrum (MWCS) technique^{193,56}. We compute time shifts, dt , in the coda of daily NCFs relative to a reference NCF, the stack of all NCFs for each station pair, in the 0.5 - 2 *Hz* frequency band. We apply a 30 day running mean to the NCFs to improve the stability of the MWCS analysis. Time shifts dt and coherency c between the reference and daily NCF are calculated beginning after the 0.5 *km/s* arrival in the coda in 10 *s* windows, shifted by 20% of the window length. dt measurements with time shift $dt \leq 0.2$ *s* in each window and coherency $c \geq 0.5$ are included. A daily time shift dt/t is measured by regressing time shifts dt from each window in the causal and acausal part of the coda.

Assuming that there is linear relation between relative time lags and that the velocity change is homogeneous throughout the sampling medium, the daily velocity variation is just $-dt/t = dv/v$.

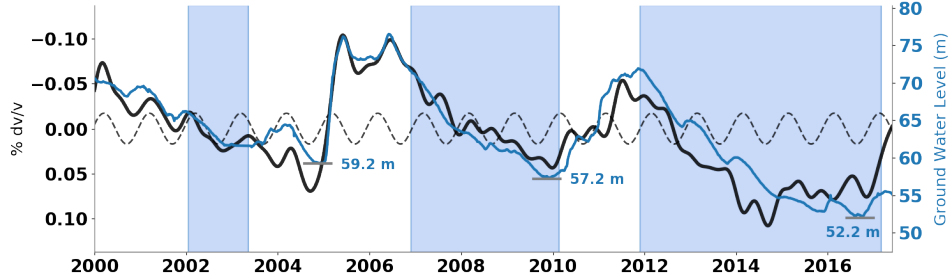


Figure 2.2: Observed dv/v stacked over all station pairs (black) with modeled dv/v due to thermo-elastic strain (dashed) removed compared with groundwater change (blue) in the Baldwin Park Key Well. Grey bars indicate lowest historical water levels of the Baldwin Park Key Well. Blue patches indicate times of drought.

DV/V REGIONALIZATION

We map dv/v spatially in $1\text{ km} \times 1\text{ km}$ grid cells using the regionalization method of *Brenguier et al.*³⁴. This simple inversion scheme approximates the scattering sensitivity of each station pair as an ellipse. We set dv/v in all grid cells within 3 km of the straight line path between each station pair as the difference in dv/v between the starting and end date of the period of interest. We then average all grid cells over all the station ellipses. A gaussian smoothing function has been applied to the dv/v maps in Fig 2.3 and 2.4. We did not use the sensitivity kernels of *Obermann et al.*¹⁸² that assume homogeneous diffuse properties, which are unlikely to be satisfied in resonating sedimentary basins.

1.1 WATER STORAGE FROM DV/V .

We calculate the change in groundwater storage dV_w in each grid cell in the SGV basin from dv/v using a modified version of the recharge equation,

$$dV_w = S_y A \Delta_{dv/v} \beta^* \quad (2.1)$$

where S_y is the specific yield, A is the area of a grid cell in the regionalization of dv/v , $\Delta_{dv/v}$ is the change in seismic velocity between two dates, and β^* is the ratio of a unit change in hydraulic head, Δh , to a unit change in $\Delta_{dv/v}$ ⁸⁸. The product $\Delta_{dv/v} \beta^* = \Delta h$ gives the average change in hydraulic head in a grid cell. S_y varies from 0.03 to 0.24 across the SGV, with averages of 0.14, 0.08 and 0.09 in the central, eastern, and western parts of the SGV, respectively⁴¹. We take $S_y = 0.12$ as a representative, average value for the entire SGV basin. Assuming that the inflation of the aquifer was purely elastic¹³⁰, we use the 2005 rain event (Jan 1 - Jun 1 2005) to calibrate β^* for the SGV. A 16.8 m increase in groundwater level in the Key Well and a -0.00125 (-0.125%) change in dv/v in the SGV basin gives a value of $\beta^* = -13280 \text{ m}/(\frac{\text{m/s}}{\text{m/s}})$. We find a similar negative value of $\beta^* = -10900 \text{ m}/(\frac{\text{m/s}}{\text{m/s}})$ using the dv/v and groundwater level changes found by *Lecocq et al.*¹⁴⁰. To estimate a volume change within the SGV basin over the 2012-2016 drought, we integrate dV_w over all grid cells.

2 RESULTS AND DISCUSSION

The dv/v variations we measure at 0.5-2.0 Hz, which is sensitive to the upper 1 km of the basin, are the most promising for groundwater monitoring at basin scale¹⁸². The change in groundwater level in the Baldwin Park Key Well explains most of the variance in the evolution of dv/v in the SGV. We observe three distinct functional forms in our dv/v measurements : 1) seasonal periodicity, 2) impulsive events, and 3) multi-year linear trends (Figure 2.2).

We use a thermo-elastic model²⁶⁸ to remove seasonal dv/v due to surface temperature variations (Fig. 2.2). We find that seasonal thermo-elastic strains induce perturbations in wavespeed of about 0.03%, much lower than the hydrological effects that perturb elastic wavespeeds that are about

0.15%. The seasonal residual in dv/v we measure is thus a component of the seasonal recharge in groundwater within the SGV basin¹¹⁷.

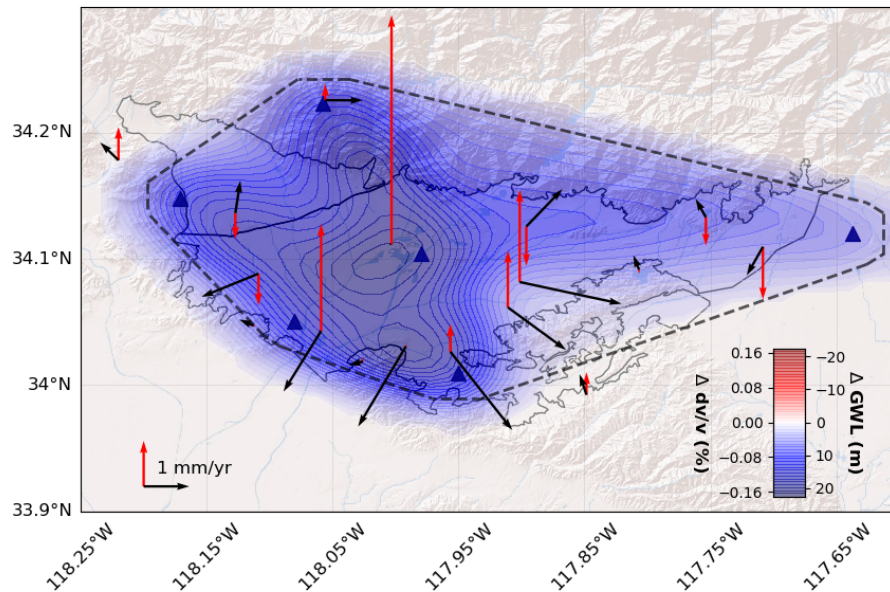


Figure 2.3: Regionalization of dv/v changes Jan 2005 - Jun 2005 following large precipitation event in the SGV. GPS stations (red = vertical, black = horizontal) uplift and move away from center of aquifer. The dashed black lines indicate extent of ray coverage. Scaling of dv/v and groundwater level is from 2005 rain event.

At the end of 2004, groundwater levels in the SGV were at an all-time low in the Baldwin Park Key Well since measurements began in 1932. In contrast, the winter of 2004-2005 recorded the largest rainfall in a 100-year period in Los Angeles with 1 m of total precipitation. Water levels in the Baldwin Park Key Well increased by over 16 m in a span of five months. GPS stations recorded more than 40 mm of uplift in the central part of the SGV^{130,118}. We find that dv/v decreased by 0.15% in the same time frame. This impulsive drop in dv/v is similar in amplitude to that seen after nearby earthquakes^{33,287}. The largest decrease in dv/v was in the center of the SGV (Fig. 2.3), where the basin is deepest²⁹⁵, as were the largest deformations recorded with InSAR¹³⁰. There is no statistically significant phase lag between the groundwater levels and dv/v response, suggesting a purely elastic response of the aquifer.

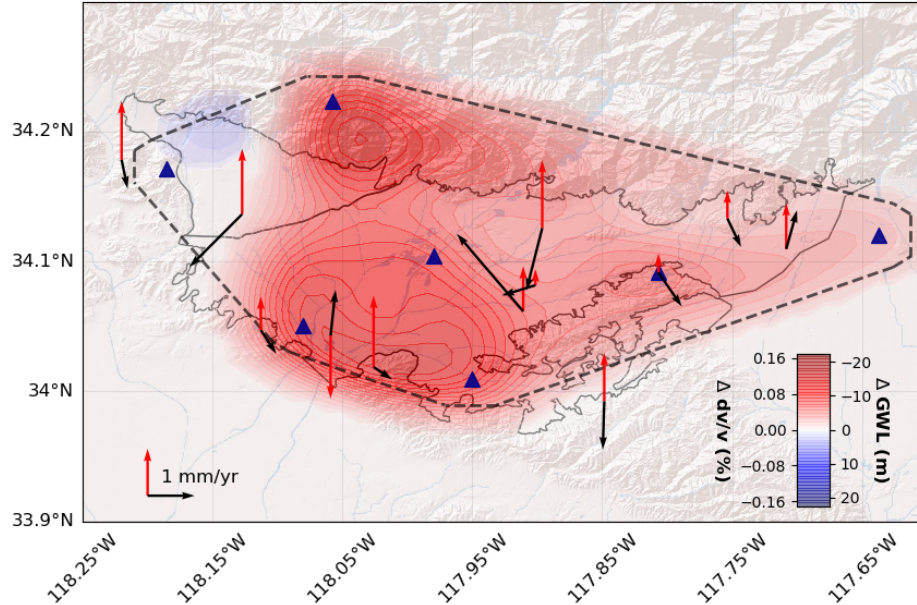


Figure 2.4: dv/v and GPS measurements after the 2012-2016 drought. Regionalization of dv/v changes (Jan 2012- Jan 2017) during California's worst drought. GPS stations move toward center of aquifer. Symbols are same as in Fig. 2.3.

During the drought of 2012-2016, groundwater levels declined in the SGV at a rate of 450 mm/yr , which is one of the highest rates seen globally²⁷². This multi-year drawdown, during a period of low precipitation²⁵⁹, was the dominant process in the increases in dv/v we measured in the SGV. The largest increase in dv/v during the drought occurred at two stations (RIO,RUS) located within the basin and atop the thickest part of the aquifer⁴¹. Over the period Jan 2012- Jan 2017, when additional well data is available throughout the SGV, we find spatial correlation between the change in dv/v and spatial and temporal patterns of groundwater change. The strongest increase in dv/v occurs in the south of the SGV (Fig. 2.4). A small decrease in dv/v in the Raymond Basin over the same time frame suggests that the SGV and Raymond basins are hydraulically separated^{41,118}. GPS stations during the same time period measured a contraction of the ground surface that may result from an elastic response of the basin.

The strong temporal correlation between groundwater levels in the Baldwin Park Key Well and

dv/v (Fig. 2.2) and spatial agreement between GPS displacements and well levels at key periods of time provide us confidence to map the change in groundwater level. We use the instantaneous elastic response of the 2005 rainfall event to calibrate the conversion between dv/v and groundwater level. Applying this calibration factor to the regionalization of dv/v from Jan 2012 - Jan 2017 yields a water storage loss of 0.48 km^3 . This matches well with the additional $0.45 - 0.5 \text{ km}^3$ of water that was pumped by from the main SGV Basin during the drought to meet water demand¹⁵¹.

Our results imply that the change in seismic velocity, dv/v , has tremendous potential to monitor groundwater fluctuations in basins of moderate-size aquifers. We find an inverse and linear scaling between dv/v and groundwater in the San Gabriel Valley. Our analysis is able to provide the water volume change, at much higher spatial resolution than GRACE data. It also the capabilities to provide direct and continuous monitoring of the spatial variations in ground water levels, complementing monitoring efforts from groundwater wells and GPS inversions of subsidence.

3

Ambient Seismic Noise Cross-Correlation on the CPU and GPU in Julia

People who like this sort of thing will find this the sort of thing they like.

Charles Farrar Browne

1 INTRODUCTION

Noise correlation functions (NCFs), generated from the cross-correlation of the ambient seismic field recorded simultaneously at a pair of seismic stations, recover information about Earth structure at spatial scales ranging from the near surface (meters) to the lowermost mantle (thousands of kilometers)^{233,145,294,230,181,199}. When averaged over long periods of time, NCFs converge to band-limited approximations of the elastodynamic Green's function between the pair of receiver stations^{284,68,241,232,218}. With the wide range of length and time scales involved in such geophysical studies, computing NCFs can be a data-intensive technique, utilizing dense, high sample-rate deployments (High/Large N) for near surface study¹⁴⁵ or multi-year to permanent deployments lasting up to decades (High/Large T) for long term monitoring¹⁴⁰.

In the last two decades, increases in computing power have allowed ambient noise processing to become practical. While the total global seismic data stored by Incorporated Research Institutions for Seismology (IRIS), the leading global manager and distributor of seismic data, is 0.66 PB total (as of July 1, 2020), Distributed Acoustic Sensing (DAS), which re-purposes fiber optic cables as seismometers, will soon be generating PetaBytes of seismic data a year⁵. Due to the increasing availability of complex and large (> 1 TB) seismic data sets, research in ambient noise seismology requires high performance software to compute NCFs.

Seismology software has been written in a number of languages, including C/Fortran (SAC - Goldstein *et al.*⁹⁴), Python (ObsPy - Beyreuther *et al.*²⁶), MATLAB (GISMO - Celso *et al.*⁴⁶), R (RSeis - Lees¹⁴²), and Unix (Computer Programs in Seismology - Herrmann¹⁰²). Software suites for computing NCFs have been written in Python (Lecocq *et al.*¹³⁹, Goutorbe *et al.*⁹⁶, Jiang and Denolle¹²⁰), C++/CUDA (Fichtner *et al.*⁸⁶) and UNIX (Herrmann¹⁰²), among unpublished codes written in other languages. The ideal computing language for ambient noise seismology would allow researchers to write high-level, performant code that scales from laptop to cluster. Currently,

the choice of language dictates needs; compiled languages (C, Fortran, Shell) allow for fast execution time at the cost of slow development, while vectorized languages (Python, Matlab, R) allow for rapid development and ease of use at the cost of slow execution time. The benefits/trade-offs of static vs dynamic languages for use in ambient noise seismology are emblematic of the two language problem: prototyping is best done in a high-level language, while implementation should be done in a low-level language⁸⁶.

To solve the two-language problem in ambient noise seismology, we developed SeisNoise.jl, a package for ambient noise cross-correlation on the CPU and GPU written entirely in the computing language Julia. Julia is a dynamically typed, high-level language that compiles at run time into low-level machine code²⁷. In Julia for loops execute as fast as C or Fortran, yet code is succinct and readable like in vectorized languages such as Python or MATLAB. Julia is the only high-level dynamic language to achieve Petaflop performance²⁶². SeisNoise.jl was designed to follow these principles: it should be open source, use only a single computing language, be easy to use in a command-line REPL (read-eval-print loop), scripting, and high-performance computing environments, use intuitive syntax, and be both fast and memory efficient. Here, we describe the structure and example use cases of SeisNoise.jl. To install SeisNoise.jl from the Julia REPL, type:

```
julia > ]  
(v1.5) pkg> add SeisNoise  
julia > using SeisNoise
```

Listing 3.1: Installation of SeisNoise.jl using the Julia package manager.

SeisNoise.jl's source code, documentation and testing are available on Github at <https://github.com/tclements/SeisNoise.jl>.

2 SEISNOISE.JL STRUCTURE AND WORKFLOW

While there are a wide range of strategies to recover NCFs^{21,243,227,87}, typical ambient noise cross-correlation workflows usually involve three steps: pre-processing, correlation, and post-processing. The goal of pre-processing is to make raw seismic data amenable for cross-correlation, which usually entails dividing day-long segments of ambient noise into shorter (sometimes overlapping) time windows²²⁷ and down-weighting large amplitude signals, such as earthquakes or instrument irregularities²¹. The correlation step is largely computational. Cross-correlation is often performed in the frequency domain due to the speed and efficiency of the Fast Fourier Transform (fft). The cross-correlation for a single time window is calculated in the frequency domain (ω) as the cross spectrum,

$$C_{AB}(\omega) = U_A^*(\omega)U_B(\omega), \quad (3.1)$$

where $U_A(\omega)$ and $U_B(\omega)$ are the Fourier transforms of the seismograms $u_a(t)$ and $u_b(t)$ at seismometers A and B , respectively, and $*$ denotes the complex conjugate. The resulting time-domain cross-correlation, $c_{AB}(\tau)$, is the inverse Fourier transform (\mathfrak{F}^{-1}) of the frequency-domain cross-spectrum,

$$c_{AB}(\tau) = \mathfrak{F}^{-1}(C_{AB}(\omega)), \quad (3.2)$$

where τ denotes the cross-correlation lag time that is a real number (positive and negative).

The goal of post-processing is to generate one or more NCFs, $n_{AB}(\tau)$, for each station pair AB , where

$$n_{AB}(\tau) = \langle c_{AB_1}(\tau), c_{AB_2}(\tau), \dots, c_{AB_n}(\tau) \rangle \quad (3.3)$$

and $\langle \rangle$ represents stacking across distinct time windows $t = 1, 2, \dots, n$.

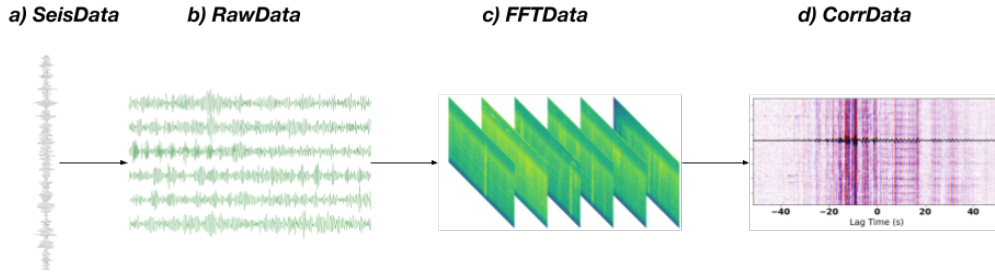


Figure 3.1: Data flow through objects in SeisNoise.jl. a) Ambient noise data read into SeisIO SeisData object for preprocessing and instrument response removal b) Windowed ambient noise data processed in a RawData object c) Fourier transforms (shown as spectrograms) stored in a FFTData object d) Cross-correlations stored in a CorrData object.

SeisNoise.jl provides three custom data structures for computing NCFs: RawData stores ambient seismic noise data in short, overlapping time windows, FFTData stores Fourier transforms of these time windows, and CorrData stores the corresponding NCFs, as shown in Figure 3.1. Because Julia uses column-major ordering, each ambient noise time window in SeisNoise.jl is stored as a column in a 2D array. The start time of each ambient noise window is stored as the number of seconds since 1970 (UNIX time) and is accessible via the `.t` field for all three types of structure in SeisNoise.jl. Ambient noise data is accessible via the `.x`, `.fft`, and `.corr` fields for RawData, FFTData, and CorrData structures, respectively. In addition to data, SeisNoise.jl structures hold numerous types of metadata, including the start time of each time window, station name and geographic location, sampling rate, instrument response, and processing notes (e.g. minimum and maximum frequencies of the data after filtering, time and spectral normalization information, maximum lag time in correlation, etc.). We introduce a general workflow for NCF processing using SeisNoise.jl structures and functions in the following sections.

2.1 CROSS-CORRELATION PRE-PROCESSING

SeisNoise.jl relies upon SeisIO.jl for reading of ambient noise data and initial data handling¹²⁷.

Seismic data (SEED, SAC, SEG Y, Nodal, etc...) can be read locally using the SeisIO.jl `read_data` function or downloaded from IRIS or FDSN data centers using the `get_data` function. SeisIO.jl stores raw seismic data and station meta-data in either `SeisChannel` (single-channel data) or `SeisData` (multichannel data) structures. SeisIO.jl provides methods for merging data, filling or removing time gaps, downsampling, detrending, tapering, removing instrument responses, and syncing channel start/end times²¹.

Cross-correlating short time windows, or overlapping time windows, has been shown to improve signal-to-noise (SNR) ratio in NCFs^{194,227}. The `RawData` structure in SeisNoise.jl allows one to operate on many short, overlapping windows of ambient noise at once. `RawData` structure take either `SeisData` or `SeisChannel` as input. For example, a day-long `SeisData` segmented into 30-minute windows with 75% overlap will yield a `RawData` structure `R` with 187 30-minute time windows. The `RawData` structure facilitates the second phase of pre-processing: preparation of time windows for cross-correlation. SeisNoise.jl provides functions for common time-window pre-processing steps, including demeaning, detrending, tapering, filtering (lowpass, highpass, bandpass and bandstop), time-domain normalizing (e.g. one-bit, root mean square clipping, running-mean normalization), and spectral-whitening²¹. SeisNoise.jl processing functions are applied to SeisNoise.jl structures, which then call low-level kernels on data²⁸. For example, calling `detrend!(R)` applies a detrending kernel to each time window in `R.x`. The advantage of this paradigm is concise production code: detrending requires a single line of code, rather than a for loop over each window. In our opinion, this concise syntax allows users to focus on workflow rather than on computational bookkeeping.

2.2 CROSS-CORRELATION PROCESSING

We proceed with cross-correlation in the frequency domain. The first step to cross-correlation is taking the discrete Fourier transform of windowed ambient noise data. Since ambient noise data is real valued, SeisNoise.jl computes discrete Fourier transforms using the real fast Fourier transform (rfft), which offer a 2-3x speed over a regular fft. The rfft function computes the real Fourier transform of the data in a RawData object and returns a FFTData object. The FFTData structure in SeisNoise.jl stores ambient noise spectra ($U(\omega)$) and allows users to apply spectral operations, such as whitening, in-place before cross-correlating.

Computing a cross-correlation in the frequency domain necessitates an element-wise multiplication of two noise spectra and an inverse Fourier transform. The SeisNoise.jl correlate function accepts two FFTData structures, computes the Fourier-domain cross-correlation between common time-windows, and returns a two-sided time-domain cross-correlation, stored in a CorrData structure. The CorrData structure provides additional metadata, such as the maximum positive and negative cross-correlation lag time and distance, azimuth, and back azimuth between the two stations.

2.3 POST-PROCESSING

The goal of post-processing is to convert raw NCFs into a more amenable format for scientific input. This is usually achieved through "stacking", or adding, NCFs over time, which increases the signal-to-noise ratio²¹. The number of NCFs generated in post-processing varies with the geophysical application. Time-independent studies, such as tomography²³³ or virtual earthquakes⁶⁶, require a single NCF, while time-dependent applications, such as noise-based monitoring using coda-wave interferometry^{181,36}, require many NCFs. Stacking in SeisNoise.jl is possible for arbitrary time periods from minutes to years, as specified by the user. SeisNoise.jl has multiple routines for stacking

correlations including the linear, phase-weighted²²⁴, robust¹⁸⁹, and autocovariance filter stacks¹⁷⁴. Additionally, SeisNoise.jl has routines for rotating NCFs in arbitrary reference frames to a vertical-radial-transverse reference frame¹⁴⁴.

After stacking or rotating, CorrData structures (along with RawData and FFTData) can be saved to and loaded from disk using the JLD2.jl package, a pure Julia HDF5-compatible file format. When saved to disk, CorrData are saved by component, e.g. "ZZ" or "EN", then by start date in "YYYY-MM-DD" form.

3 SEISNOISE.JL PARALLELIZATION

Ambient noise cross-correlation lends itself to parallel processing. For a dataset with N stations, pre-processing and computation of Fourier transforms scale with N in time, while the computation of cross-correlation and post-processing scales with N^2 in time. Thus, most computational time in ambient noise cross-correlation is spent cross-correlating and post-processing when working with a reasonably large number of stations. Cross-correlations across distinct time windows (e.g Day 1 and Day 2) are independent, while cross-correlations within a single time window (e.g Day 1) are dependent on the N Fourier transforms, $U_1(\omega), \dots, U_N(\omega)$. Computing cross-correlations across distinct time windows is thus "embarrassingly parallel", while cross-correlating within a single time-window requires either movement of individual $U(\omega)$'s among processing units, i.e. shared-memory parallelism, or computation-level parallelism, such as multi-threaded matrix multiplication¹⁰¹.

Numerous options exist for parallelizing ambient noise cross-correlation. Previous authors parallelized cross-correlation using the MapReduce framework via Hadoop³, distributed memory parallelism via the Message Passage Interface (MPI)^{48,120}, job-based parallelism via a Structured Query Language (SQL) database¹³⁹, shared memory parallelism via Open Multi-Processing (OpenMP)⁸⁶, and Graphical Processing Unit (GPU) parallelism via Compute Unified Device Ar-

chitecture (CUDA)⁸⁶. We use two types of parallelism in SeisNoise.jl: distributed-memory parallelism via the Distributed.jl Julia module and GPU-parallelism via the JuliaGPU suite.

Our CPU-based parallelism is simple. We apply distributed memory parallelism across cross-correlation time windows using the parallel map function, `pmap`, from Distributed.jl. `pmap` takes a functional input and “maps” it in parallel on a set of variable inputs. Unlike MPI-style parallelization, which is also available in Julia, `pmap` handles all parallelization, data distribution, and load balancing for the user⁶⁴. Cross-correlating an entire ambient noise data set using a `pmap`-based workflow requires only defining a map function that computes NCFs between all station-pairs for a single day, a list of day-long seismic trace file paths as inputs, and a list of parameters for cross-correlations, such as the ambient noise window length, overlap between windows, and maximum lag-time in the cross-correlation to save. We believe the parallel map framework is a particularly simple to implement cross-correlation parallelism.

3.1 GPU PARALLELIZATION

GPUs are highly parallelized processors designed for high throughput and computationally intensive applications. GPUs are widely used to accelerate data-intensive seismic workloads, such as stacking³⁰⁰, earthquake detection^{161,19}, reverse time migration², or wave propagation simulations¹³². Order of magnitude decreases in processing time for ambient noise cross-correlation are possible using Nvidia’s CUDA platform for writing GPU code^{86,270,19}. The trade-offs associated with writing GPU code in CUDA C or C++ are the difficulty and time to write CUDA code, along with a lack of code reuseability, i.e. code for the CPU cannot be reused on the GPU. This is not the case when using Julia.

Julia provides a high-level interface to the CUDA GPU programming toolkit through the `CUDA.jl` package²⁵. `CUDA.jl` provides an GPU-based array type for transferring and storing data on the GPU and access to CUDA kernels, such as matrix multiplication and Fourier transforms. Data in

SeisNoise structures (`R.x`, `F.fft`, and `C.corr` fields, for `RawData`, `FFTData`, and `CorrData`, respectively) can move between an `Array` on the CPU to a `CuArray` on the GPU using the `gpu` and `cpu` functions and the Julia pipe function, `|>`, as shown below.

```
# create raw data and send to GPU
R = RawData(S, cc_len, cc_step) |> gpu

# send data back to the CPU
R = R |> cpu
```

Listing 3.2: Sending data to and from the GPU using Julia pipes.

In most cases, `SeisNoise.jl` uses the same code to process data on the CPU and GPU. For example, processing functions in `SeisNoise.jl`, such as `detrend`, `taper` and `correlate`, are written at a high level of abstraction, such that they can accept a CPU `Array` or a GPU `CuArray` with no modification. This leads to, in our opinion, GPU code that is comprehensible for users with little GPU-programming experience.

There is a variable time cost, on the order of 1-100 ms, which scales linearly with data size and depends on GPU model, to transfer data from the CPU to GPU, or vice-versa. The GPU-based processing strategy for `SeisNoise.jl` is thus to keep data on the GPU for as long as possible. This entails transferring `RawData` to the GPU, then doing all pre-processing with `RawData` on the GPU, computing Fourier Transforms and spectral whitening on the GPU, and finally computing cross-correlations and stacking on the GPU. Stacked cross-correlations must be transferred back to the CPU before saving to disk. A GPU-based cross-correlation code in `SeisNoise.jl`, thus looks almost identical to a CPU-based single-core code, with the exception of added `|> gpu` and `|> cpu` syntax for memory transfer to and from the GPU. See section 4.2 for an example of `SeisNoise.jl`

GPU cross-correlation. We detail the speedup of GPU vs. CPU processing in the next section.

3.2 PERFORMANCE BENCHMARK

We benchmark SeisNoise.jl's single CPU core, many CPU cores, and GPU performance using a dataset similar to *Fichtner et al.*⁸⁶. We selected 188 vertical component LHZ channels with 1 Hz sampling rate, operating from January 1, 2019 to January 1, 2020, with up-time greater than 98% (more than 360 days) for the entire year. The entire dataset of LHZ waveforms is 23 GBs in size.

For the first benchmark, we test SeisNoise.jl's processing performance on the CPU and GPU against our previous ambient noise cross-correlation code written in Python using Numpy and Scipy⁵⁸. We cross-correlate one year of data for a single LHZ-LHZ station pair using a single core on an Intel i7-8700K CPU with 32 GB RAM and a Nvidia GeForce GTX 1070 Ti GPU with 8 GB VRAM (video RAM). The code for each test is the same - we read the data using SeisIO.jl, then convert immediately to RawData with a window length of 32,768 (2^{15} to maximize fft efficiency) seconds and an overlap of 6,000 seconds between windows. We apply a simple pre-processing scheme of detrending, tapering, highpass filtering above 0.001 Hz, and time-domain normalizing the RawData with one-bit²¹. The data is then cross-correlated in the frequency domain, inverse Fourier transformed, cut to a maximum lag time of $\pm 12,000$ seconds and stacked (instrument response was not removed in this benchmark).

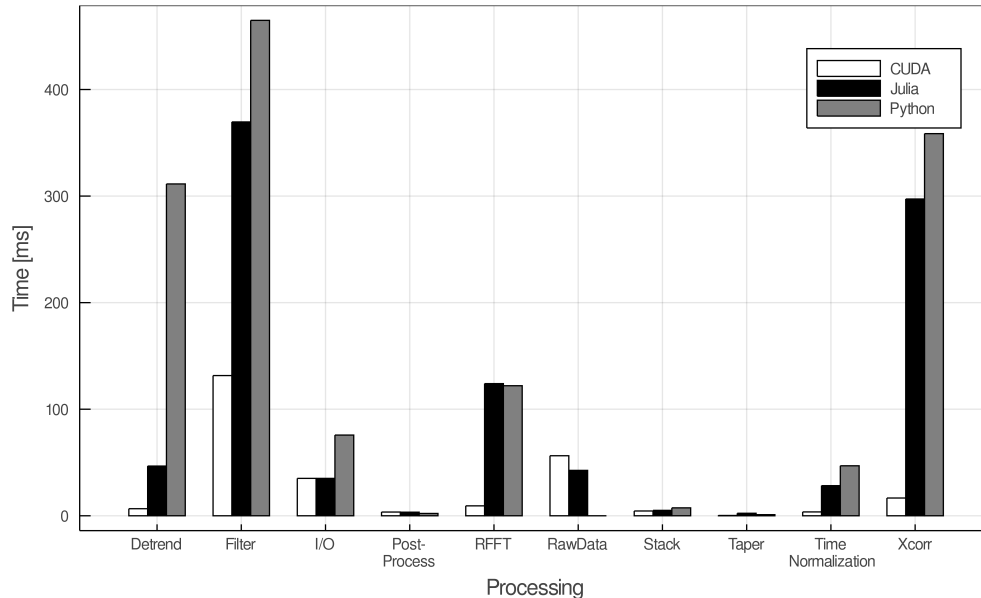


Figure 3.2: Median run times for common ambient-noise processing steps using 1 year of LHZ cross-correlations between stations BK.CMB.00.LHZ and US.WVOR.00.LHZ. Times for Julia/GPU, Julia/CPU and Python/GPU are colored white, grey and black, respectively. Median run time for conversion from SeisData to RawData in Python is not included.

The total time for cross-correlation with Julia/GPU is 0.503 s, Julia/CPU 1.7 s, and Python/CPU 2.7 s, giving a 3.5x speedup for Julia GPU vs CPU, and 1.5x speedup for Julia vs Python. Benchmarking results for each step in the cross-correlation workflow is given in Figure 3.2. Interestingly, the most time-intensive processing for a single station-pair is detrending and filtering. Julia/GPU performs better than Julia/CPU and Python/CPU on detrending and cross-correlation, as these operations involve matrix multiplications. Even though our test GPU has 18x theoretical floating point performance than the test CPU, we only achieve a 3.5x speedup using the GPU. We ascribe the CPU's relatively higher performance to 1) non-negligible time taken to transfer data to and from the GPU, 2) all SeisIO.jl pre-processing is on the CPU and 3) our non-optimal implementation of Butterworth filtering on the GPU.

To test many-core performance, we cross-correlate all station-pairs in the year-long LHZ dataset

(17,578 pairs in total). We apply the same processing steps as above, with the addition of instrument response removal, following *Fichtner et al.*⁸⁶. As a baseline, a stacked NCF for each station pair in the dataset on a single core of the Intel i7-8700K took just under 1 hour and 50 minutes, while our GPU implementation took 28 minutes on the Nvidia GTX 1070 Ti. The cross-correlations generated during this test are shown in Figure 3.4. This compares to 163 days and 16 hours reported to compute a similar dataset on the CPU and GPU, respectively, by *Fichtner et al.*⁸⁶. We tested the SeisNoise.jl's many core CPU performance using an Intel Xeon Platinum 8000 with 48 cores and 192 GB RAM, which took 12 minutes. This shows that SeisNoise.jl can take on relatively heavy computational workloads using only desktop resources.

4 EXAMPLES

4.1 PROTOTYPING IN THE REPL

Following the easy to use SAC⁹⁴, Obspy²⁶ and SeisIO.jl¹²⁷ suites for seismic data processing, SeisNoise.jl is designed for rapid prototyping in a command-line REPL (read-eval-print loop) environment. Below is a *minimal* working example of SeisNoise.jl meant to be used in a REPL. In this example, SeisIO.jl is used to download one day of data from channels CI.SDD.BHZ and CI.PER.BHZ in the Southern California Seismic Network. The complete cross-correlation process is then implemented in about 10 lines of code using SeisNoise.jl functions. The Julia dot notation (e.g. `detrend!.(R)`) for broadcasting function calls to each element of an array is used to eliminate duplicated processing steps for both stations or the use of a for loop.

```
using SeisNoise, SeisIO

# parameters
fs = 40. # sampling frequency in Hz
```

```

freqmin, freqmax = 0.1, 0.3 # min and max freq in Hz
cc_step, cc_len = 450, 1800 # corr step/length in S
maxlag = 60. # maximum lag time in correlation
smoothing_half_win = 12

# download data
S = get_data("FDSN", "CI.SDD..BHZ, CI.PER..BHZ", src="SCEDC",
            s="2019-02-03", t="2019-02-04")

# pre-process data
process_raw!(S, fs)
R = RawData.( [S[1], S[2]], cc_len, cc_step)
detrend!(R)
taper!(R)
bandpass!(R, freqmin, freqmax, zerophase=true)

# compute correlation
FFT = rfft.(R) # Fourier transform
whiten!(FFT, freqmin, freqmax) # whiten [fmin, fmax] Hz
C = correlate(FFT[1], FFT[2], maxlag) # cross-correlate
bandpass!(C, freqmin, freqmax) # filter data
abs_max!(C)

# plot
corrplot(C) # plot

```

Listing 3.3: Cross-correlation workflow testing in the REPL

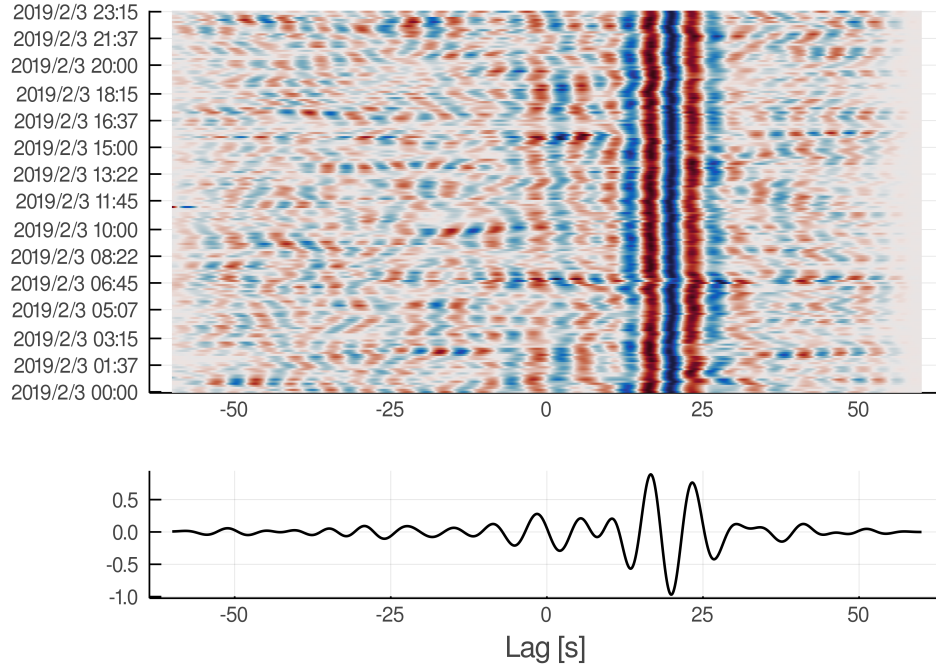


Figure 3.3: Example output of SeisNoise.jl code example. (Top) 30-minute cross-correlations from 2019/2/3 between stations CI.SDD and CI.PER, filtered between 0.1-0.2 Hz. (Below) Stacked daily correlation for CI.SDD - CI.PER.

This examples produces Figure 3.3. Functions in Julia are just-in-time (JIT) compiled the first time they are run. Consequently, the example above will be slower the first time it is run than on subsequent runs.

4.2 INTERACTIVE GPU PROTOTYPING

While most large-scale GPU applications are run on large clusters, researchers often test and benchmark GPU codes using local resources. One benefit to GPU computing in Julia is the ability to do interactive programming through the REPL. This avoids the test-recompile-test procedure required by CUDA C. Users can inspect data stored in memory on the GPU using functions from Julia's Base library such as `min`, `abs`, or `sqrt`. This allows for rapid data iteration and testing. For example,

here is a complete cross-correlation routine on the GPU that can be tested in the REPL, where S1 and S2 are SeisData objects.

```
# send data to GPU
R1 = RawData(S1, cc_len, cc_step) |> gpu
R2 = RawData(S2, cc_len, cc_step) |> gpu
R = [R1, R2]

# preprocess on the GPU
detrend!(R)
taper!(R)
bandpass!(R, freqmin, freqmax, zerophase=true)

# FFT on GPU
FFT = rfft.(R)
whiten!(FFT, freqmin, freqmax)

# compute correlation and send to cpu
C = correlate(FFT[1], FFT[2], maxlag) |> cpu
```

Listing 3.4: GPU-based cross-correlation workflow testing in the REPL

4.3 LARGE-N PROCESSING

Seismic arrays with many sensors, so called “large-N” arrays, attempt to record the complete seismic wavefield¹²⁸. Ambient noise cross-correlation analyses can benefit greatly from large-N datasets, as close station spacing and high sampling rates increase spatial and temporal resolution for tasks such as tomography^{145,196} or near surface monitoring using the change in seismic velocity³⁷. One challenge of using large-N datasets is handling the massive amounts of waveform data, on the order

of hundreds of gigabytes to terabytes per day, that sensors such as seismic nodes^{145,78,253}, distributed acoustic sensing^{5,77} or mixes of sensors²⁷⁵, can produce. Processing these data is challenging for all but the largest supercomputers. Strategies for processing large-N data products include locally sparse tomography²⁹, local similarity¹⁴³, graph clustering²⁰¹, the subarray method with double beam-forming⁴⁵, and data compression¹⁵⁴, among others.

Here, we present a strategy for computing ambient noise cross-correlations from large-N datasets with SeisNoise.jl. We use data from a single day of the LARge-n Seismic Survey in Oklahoma (LASSO) Experiment as a test dataset. The LASSO experiment used $\approx 1,830$ single-component nodal seismometers in North Central Oklahoma, shown in Figure 3.4⁷⁸, from April - May 2016 to capture induced seismicity due to wastewater injection. We use data from the $N = 1,825$ stations operating on May 1, 2016. The data are sampled at 500 Hz and the total size of raw data stored in SAC files is 294 GBs. We consider all combinations of station-pairs, giving 1,664,400 cross-correlation pairs total. We follow a similar pre-processing strategy as in the benchmark in Section 3.2: we down-sample the data to 250 Hz, remove the instrument response, bandpass filter between 0.1 and 20 Hz, and finally segment the data into 15 minute windows with 50% overlap (191 windows per day). We compute a Fourier transform of the pre-processed data once for each station, then write the frequency-domain data to disk. This takes 69 minutes using 6 CPU cores and would scale linearly with increased CPU cores.

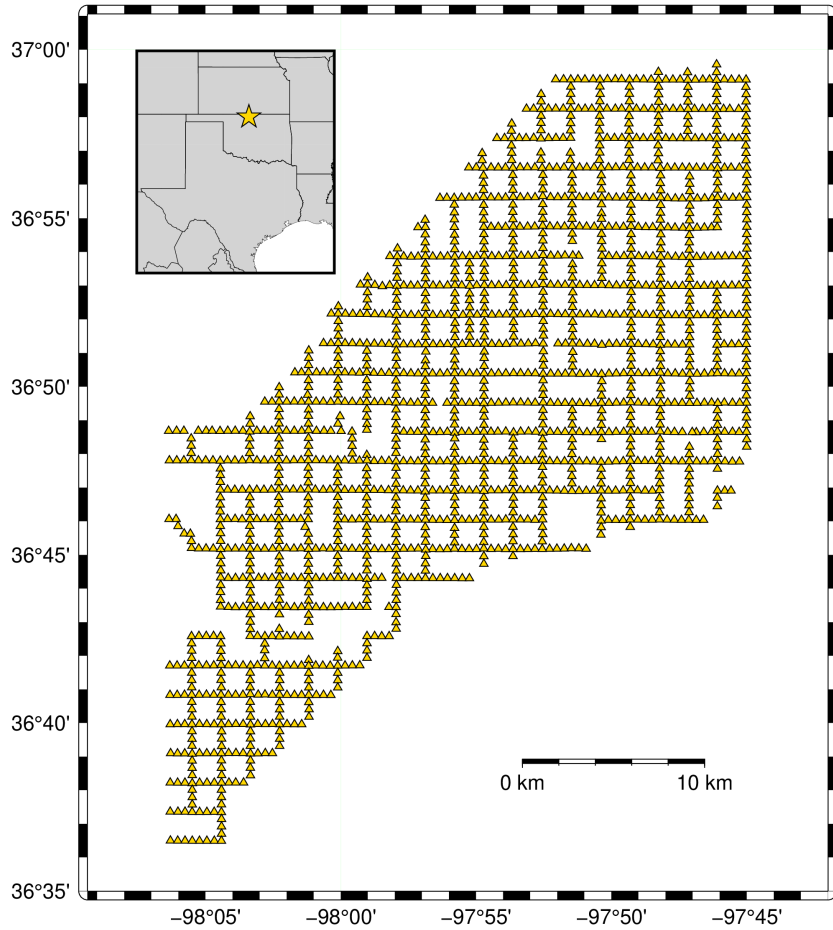


Figure 3.4: Location of 1,825 stations used from the LARge-n Seismic Survey in Oklahoma deployment operating on May 1, 2016. Average station spacing is 400m. Location of array in Northern Oklahoma is shown on inset map.

We use GPUs to cross-correlate the LASSO dataset. We use 2 different GPUs in this example: desktop (NVIDIA 1070 Ti) and server (NVIDIA V100), with 8GB and 32GB of VRAM, and 8 TFLOP (1 TFLOP = 10^{12} floating point operations per second) and 14 TFLOP of single-precision performance, respectively. Each of these GPUs does not have enough VRAM to store all FFTs in the dataset (293 GB) all at once. We thus use a processing strategy similar to the MPI-based strategy of *Chen et al.*⁴⁸ to minimize I/O. We load stations into memory on the GPU in groups such that the number of stations per group, N_{pg} , times the size of each FFT in memory, 163 MB in

our case, is about 30-40% of the total VRAM for each GPU ($N_{pg} = 18$ and 73, for NVIDIA 1070 Ti and V100 GPUs, respectively). This allows us to hold two groups of FFTs in memory on the GPU at a time, while also reserving memory for output cross-correlations. We first load one group on the GPU and cross-correlate all stations in the group. Intra-group cross-correlation give $(N_{pg} * (N_{pg} - 1))/2$ station pairs per group. While the first group is cross-correlating on the GPU, we use asynchronous I/O to load the second group of FFTs into RAM on the CPU. We then cross-correlate each FFT in the first group with every FFT in the second group, for a total of N_{pg}^2 station pairs. We then loop through every combination of groups, using asynchronous I/O to maximize the use of our GPUs. Daily cross-correlations are stacked before being transfer back to the CPU and written to disk. A moveout plot for all 1,664,400 pairs stacked for a day in 50m inter-channel distance is shown in Figure 3.5.

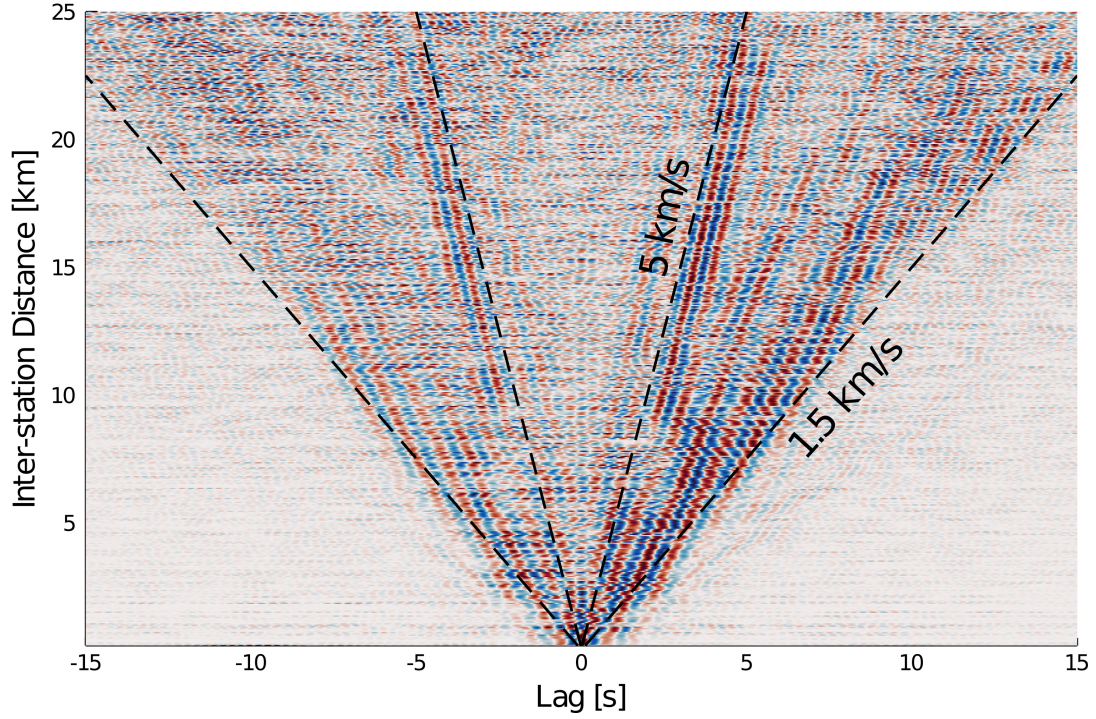


Figure 3.5: Stack of 1,664,400 cross-correlations from the LASSO array filtered in the frequency band 1-2 Hz and stacked in 50 m bins.

GPU are efficient for computationally demanding tasks. By our calculations, a single core of the CPU used in Section 3.2 would take 6.75 days (or 162 hours) to compute the cross-correlations alone of the single-day LASSO dataset. Moving the computation to the GPU takes 31 hours on a 1070 Ti GPU (a speed up of more than 5 compared to the CPU) and 16 hours on a V100 GPU. Processing time scales linearly with GPU performance. As seen with the performance of neural network computations¹³⁶, we expect Large-N cross-correlation results will improve with faster GPUs and even larger datasets.

5 FUTURE DEVELOPMENT

We have taken a functional, minimal, and modular approach to the development of SeisNoise.jl. As the breadth of research topics in ambient seismic noise continues to grow¹⁷⁵, we do not expect the core of SeisNoise.jl, its data structures and core functions, to change. This will allow future researchers to employ and extend SeisNoise.jl for their particular needs. For example, we have implemented a simple package for ambient seismic velocity monitoring (<https://github.com/tclements/SeisDvv.jl>) that is compatible with SeisNoise.jl. We anticipate that computation-heavy ambient noise research, such as global full-waveform inversion²¹⁹ or structural imaging using Distributed Acoustic Sensing⁵, could be ported to the GPU using SeisNoise.jl and Julia. Contributions to SeisNoise.jl and new packages based on SeisNoise.jl are welcome on Github.

Plotting in Julia is still in its infancy. The time-to-first plot in Julia is long (think seconds, instead of milliseconds) because Julia code is just-in-time compiled, rather than pre-compiled. Subsequent plot calls are relatively fast. All plots (except Figure 3.1) in this manuscript were made using the native Julia plots library, Plots.jl, which is based upon GR framework (gr-framework.org). There are other options for plotting in Julia, including, but not limited to, the web-based Plotly framework (<https://github.com/plotly/Plotly.jl>), the popular Python-based PyPlot (<https://github.com/JuliaPy/PyPlot.jl>)¹¹², and Makie (<https://github.com/JuliaPlots/Makie.jl>), a library high level plotting on the GPU in Julia. Maps, such as Figure 3.4, can be made with the Julia wrapper for the Generic Mapping Tools²⁸⁹. Due to the current flux of plotting in Julia, we may reevaluate the use of Plots.jl for plotting with SeisNoise in the future.

While GPU programming in SeisNoise.jl is promising, it is in no way yet optimized. Our performance benchmarks indicate a 3-4x speedup going from cross-correlation on the CPU to GPU, whereas the theoretical maximum speedup is somewhere between 10-20x. We suggest three future improvements to close that gap. The first is porting all SeisIO.jl functionality to the GPU. Seis-

Noise.jl depends heavily on SeisIO.jl for initial preprocessing, such as downsampling and instrument response removal. Modifying SeisIO.jl, such that it can process SeisChannel on the GPU, will allow for an end-to-end GPU-based cross-correlation workflow. At the time of writing, this project is under development. The second improvement is to write optimized GPU kernels using CUDA.jl. Currently, the GPU code in SeisNoise.jl is written in pure Julia. We anticipate writing custom CUDA kernels using CUDA.jl, will result in large speedups for our filtering operations. Our final suggested improvement is to use NVIDIA Tensor Cores for cross-correlation. Tensor Cores are specialty GPU hardware that offer 10x the matrix multiplication performance of CUDA cores. We anticipate implementing time-domain cross-correlations in SeisNoise.jl to use Tensor Cores could result in a significant speedup in the case where no pre-whitening of the time series is required¹⁵⁴. GPU-compute in SeisNoise.jl is only available with NVIDIA devices at the time of writing. GPU-compute using devices manufactured by Intel and Advanced Micro Devices are not yet supported in Julia, though packages for both the AMD Radeon Open Compute platform (ROCm) ([AMDGPUnative.jl](#)) and the Intel oneAPI unified programming model ([oneAPI.jl](#)) are currently under development.

In the coming years, we expect that cloud computing, which is optimal for embarrassing parallel jobs, will become the dominant platform for assembling massive ambient seismic noise datasets, performing ambient noise cross-correlation and sharing results¹⁴⁹. Cloud-based workflows will be essential for decreasing time-to-science. While we have only tested SeisNoise.jl on Amazon Web Services using the Julia AWS API ([AWSCore.jl](#) - <https://github.com/JuliaCloud/AWSCore.jl>), integration of SeisNoise.jl with other cloud-based compute systems, such as Google Cloud, Microsoft Azure and the NSF-funded Jetstream cloud²⁴⁵, will give researchers options for their cloud computing needs.

4

Single Seismic Station Groundwater Monitoring

You can observe a lot just by watching.

Yogi Berra

I INTRODUCTION

The autocorrelation of ambient seismic data is a suitable technique for monitoring changes in groundwater level. Heterogeneities in Earth's upper crust readily scatter seismic energy¹⁰. Scattered waves are more sensitive to perturbations in the subsurface than direct waves, as scattered waves sample the same location multiple times. This allows scattered waves to accumulate phase delays in response to perturbations. *Grêt et al.*⁹⁷ showed that the seismic velocity in a laboratory sandstone as measured with scattered coda waves²⁴² decreased linearly with increasing saturation, whereas there was no noticeable change in velocity for direct waves. At frequencies above 1 Hz, anthropogenic sources, such as vehicles, wind farms, oil and gas production, and other industrial activities generate high-frequency surface waves that are observable on seismometers kilometers away^{158,71,225}.

*Claerbout*⁵⁵ was the first to theoretically derive the relation between the autocorrelation of earthquake seismograms and the surface response to an impulse force, or the zero-offset Green's function. It was only twenty years after this result that *Claerbout et al.*⁵⁴ speculated, though did not prove, that the autocorrelation of ambient seismic noise *should* yield a reflection seismogram, based on the fact that the autocorrelation of white noise yields an impulse function and that the autocorrelation of ambient seismic noise yields an impulse function. Physically, the autocorrelation of ambient noise, which is the cross-correlation of waves ascending from depth with down-going waves reflected from the surface⁵⁴, yields singly and, at later lag times, multiply-scattered waves from reflectors at depth. *Rickett and Claerbout*²⁰⁵ further speculated that continuous autocorrelation of ambient noise could yield real-time monitoring of the subsurface.

There are two options for single-station ambient noise cross-correlation: auto-correlation (AC), which is cross-correlation of a ground velocity component (east, north, vertical) with itself (e.g. east-east, north-north, or vertical-vertical), or single-station cross-correlation (SC), which is the cross-correlation of differing channels (e.g. the causal and anti-causal sides of the east-north, east-

vertical, etc.. components)¹⁰⁸. I focus the following analysis on SC functions, as they provide more stable results and more independent measurements (6 vs 3) than AC functions^{108,271}. SC functions have differing sensitivities: the east-north component is sensitive to Rayleigh and Love waves, while the east-vertical and north-vertical components are sensitive to Rayleigh waves only¹⁰⁸. However, empirically, there are numerous examples that show that body-wave information may be recovered from ambient-seismic noise AC and SC functions, in a similar fashion to receiver functions^{263,65,57,222}.

Changes in groundwater level (pore pressure) affect both the phase and amplitude of seismic waves. Amplitude changes can be recovered through characterizing seismic attenuation, the rate of decay of seismic waves. In most rocks, attenuation increases with increasing saturation¹⁵⁵. Usually it is easier to measure the attenuation of coda waves (Q_c^{-1}), which sample more area for longer time, than direct waves¹⁰. Determining the cause of temporal changes Q_c^{-1} is often inconclusive - many authors have not found a link between Q_c^{-1} and earthquakes^{50,24}, though Q_c^{-1} in California varies on the order of $\sim \pm 30\%$ through time^{247,152}. A Q_c^{-1} decrease after eruption at Mt. St. Helens suggest that Q_c^{-1} is sensitive to fluids⁸⁵. In Chapter 1, we introduced a method to monitor seismic attenuation continuously over decades using SC functions from a single seismometer.

Phase delays of seismic waves are the result of changes in seismic velocity. Calculating seismic velocity change measurements, dv/v , from single-station ACs and SCs has gained popularity in the last decade due to their localized spatial sensitivity and small computational footprint - a network of n stations scales with n in time for AC vs n^2 for CC. dv/v from ACs has been employed to monitor volcanic eruptions⁶³, earthquakes^{204,271,108} and more recently, for groundwater¹²⁹ and soil moisture monitoring¹¹⁴ in local settings. In aquifers, there is an inverse relationship between saturation (groundwater level) and seismic velocity¹⁸⁵. There are several scales at which the perturbation in seismic velocities have been associated with hydrological processes - the near surface, upper 10 meters, soil moisture and shallow water table^{114,180}, changes in permafrost at depth¹¹⁵, groundwa-

ter level changes from 50 - 500m depth^{230,268,58} and deep changes in pore pressure at up to 8 km depth^{207,276}.

Subsurface hydrology operates at a range of scales: from the surface to kilometers in depth, from meters to hundreds of kilometers spatially, and from seconds to decades temporally. The spatial sensitivity of a groundwater well can be on the order of meters to kilometers, depending on the rate at which an aquifer transmits water and the amount of water released from storage when the head in the aquifer falls^{260,223}. To understand their spatial sensitivity, I take the example of a pumping well. *Theis*²⁶¹ expressed the drawdown (difference in hydraulic head) $\Delta b(r, t)$ in the vicinity of a well with the pumping rate Q ,

$$\Delta b(r, t) = \frac{Q}{4\pi T} \int_u^\infty \frac{e^{-z}}{z} dz, \quad u = \frac{r^2 S}{4Tt}, \quad (4.1)$$

based on the exact analogy to the cooling of a metal plate and where r is the radial distance from the well, T is the coefficient of transmissibility, u is a dimensionless quantity, S is the coefficient of storage, and the definite integral $\int_u^\infty \frac{e^{-u}}{u} du$ is the exponential integral^{261,260}. The radius of the cone of depression due to groundwater pumping is dependent only on S , T , time of pumping t , whereas the amplitude of $\Delta b(r, t)$ is linearly proportional to the pumping rate Q ²⁶⁰. Due to their decreasing spatial sensitivity with increasing r , groundwater level measurements from groundwater wells can be thought of as point measurements.

Groundwater well measurements and single-station cross-correlation functions share a diffusive spatial sensitivity. Using a diffusion approximation, *Pacheco and Snieder*¹⁸⁸ found the 3D sensitivity kernel for the propagation of scattered waves, K_{3D} , for a co-located source and receiver to be,

$$K_{3D}(\mathbf{r}, t) = \frac{1}{2\pi D r} \exp\left(\frac{-r^2}{Dt}\right) \quad (4.2)$$

where \mathbf{r} is the location of the receiver, r is the distance from the source, t is the propagation time,

and D is the diffusion constant. From equation 4.2, I note that for a fixed r , sensitivity increases with t . In a diffusive regime, the distance travelled from the source is related to the square root of time, $r \approx \sqrt{4Dt}$ ¹⁸⁸. Ambient seismic autocorrelations (AC), where the source and receiver are co-located, are thus most sensitive to the area directly below the seismometer, which agrees with Claerbout's conjecture that an AC is a reflectivity response. This is in contrast to ambient noise *cross-correlations* (CC), which are sensitive to the regions beneath a pair of seismometers *and* the path between them^{188,182}. Noting the similarity of equations 4.1 and 4.2 with respect to the scaling of r and t , I believe ACs are a more natural analog to groundwater wells in the case of groundwater monitoring than CCs. The new contribution of this analysis is a comprehensive comparison observations derived from hydrological, seismic, and geodetic measurements, and of empirical and theoretical hydrological models. In particular, I develop a transfer function between dv/v and groundwater level changes, which I derive using poro-elastic and nonlinear elastic theory and validate using hydrological observations and geodetic water proxies.

2 DATA

For this chapter, I use 18 years of three-component, continuous seismic velocity data from Southern California Seismic Network station CILLJR to calculate the change in seismic velocity, dv/v , through time. CILLJR is located in the Tejon Pass between the San Emigdio and Tehachapi Mountains⁴⁰, about 1.5 km east of both the San Andreas Fault and Interstate I-5. Additionally, I gather daily precipitation data from the Parameter-elevation Regressions on Independent Slopes Model (PRISM) dataset, weekly Liquid Water Equivalent (LWE) estimates from the Gravity Recovery and Climate Experiment (GRACE) and Gravity Recovery and Climate Experiment Follow-On (GRACE-FO) missions²²¹, annual groundwater measurements from USGS well 344614118454101, and weekly groundwater measurements from wells in the Castac Lake Valley Basin⁴⁴. Groundwater

proxy data have differing time spans and sampling rates - GRACE LWE data spans from 2002 until now, with a slight gap in 2017-2018 between the end of GRACE and the launch of GRACE-FO. PRISM precipitation data is daily from 1985 until now. Groundwater level measurements from well 344614118454101 start in 2005. The location of instruments and dataset grid cells used in this chapter are shown in Figure 4.1. GRACE data from CSR GRACE/GRACE-FO are sampled in $0.25^\circ \times 0.25^\circ$ grid cells, but their true resolution is near 250–300 km due to the band-limited nature of GRACE and smoothing applied during processing²²¹.

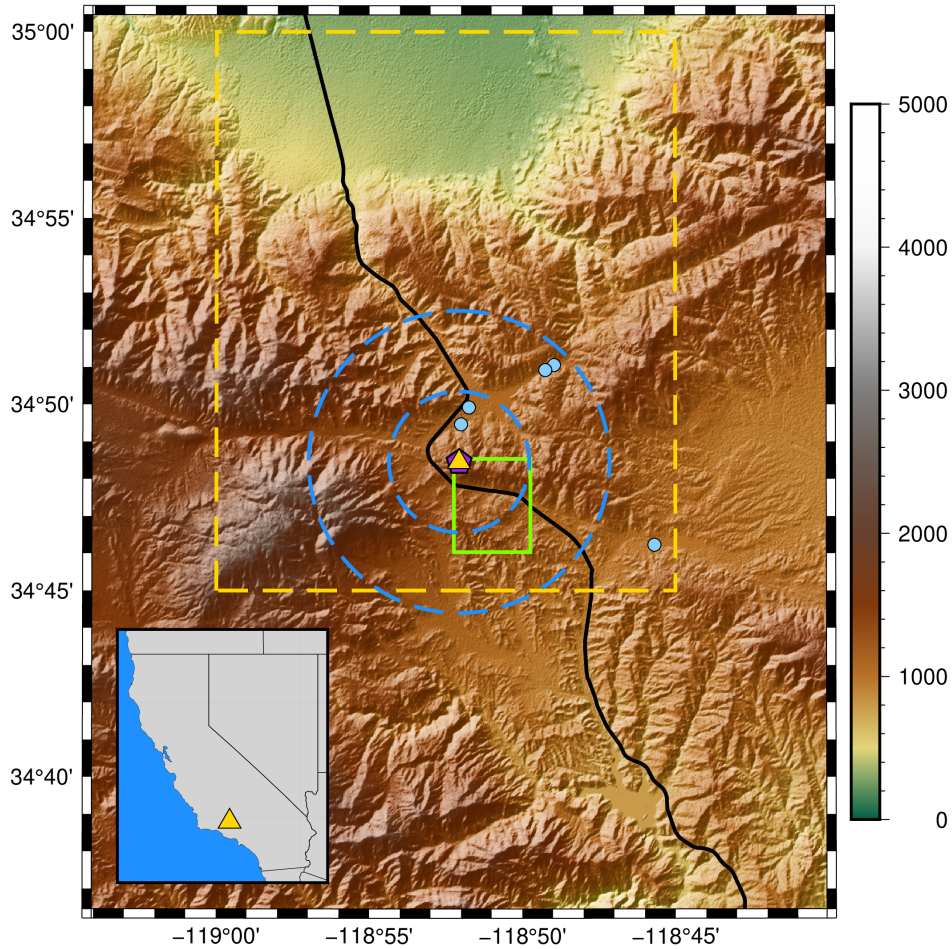


Figure 4.1: Location of seismometer CI.LJR (gold triangle) and groundwater proxies. The chartreuse rectangle denotes the 4km x 4km precipitation grid cell from PRISM dataset. The black line indicates the path of Interstate 5 through the Tejon pass. The dashed blue circles approximate limit of spatial sensitivity of CI.LJR autocorrelation at lag times of $\tau = 2$ and 8 seconds, respectively. The filled blue dots indicate position of nearest groundwater wells to CI.LJR. The orange rectangle denotes $0.25^\circ \times 0.25^\circ$ grid cell from CSR GRACE/GRACE-FO RL06 version 2 Liquid Water Equivalent (LWE) dataset.

2.1 PRECIPITATION AT CI.LJR

CI.LJR is located in a Mediterranean climate, typified by mild, wet winters and hot, dry summers⁷⁶ - nearly all rainfall occurs from October to May. Daily precipitation levels at CI.LJR are well approximated by an exponential distribution ($P < 1e - 4$)²⁰³. Annual precipitation totals are heavily

dependent on large storms - the wettest 10% of days account for 49% of the annual rainfall⁶⁹, as shown in Figure 4.2.

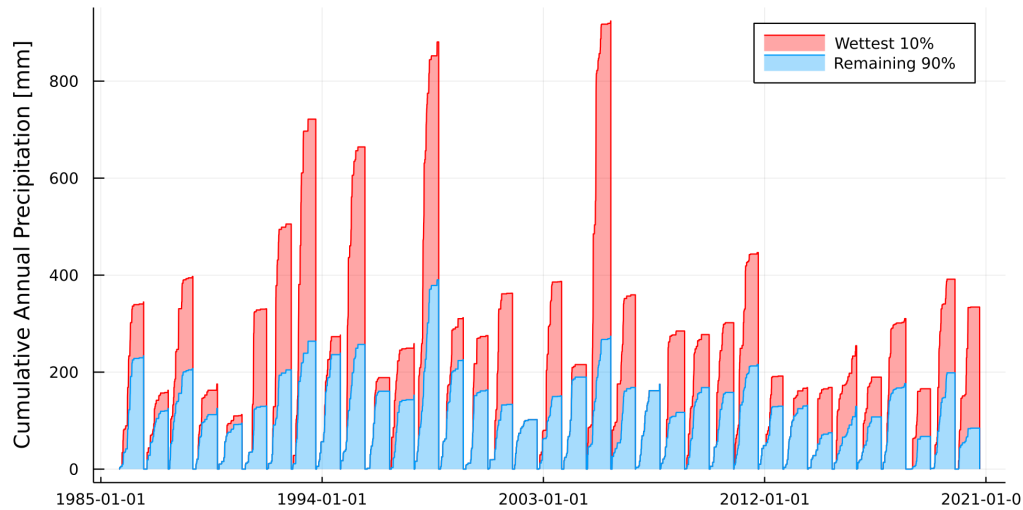


Figure 4.2: Cumulative annual (Oct 1 - Sep 30) precipitation levels near seismic station C.I.LJR for wettest 10% of days (red) and remaining 90% of days (blue) with rainfall.

In California, annual precipitation totals are heavily dependent on large storms - the wettest 10% of days account for 49% of the annual rainfall⁶⁹, as shown in Figure 4.2. The variance in annual precipitation at C.I.LJR is strongly linked to the the number and intensity of large storms in a given year. Since 1992, there has been a 54% reduction in cumulative annual rain contribution from the remaining 90% of wet days ($P = 0.002$). There is also a clear absence in extremely wet years, the last one being in 2005. Two time periods stand out from the precipitation record. First, in the winter of 2004-2005, the annual precipitation was over 3 times the median annual value and there were 18 days with large storms. Second, in the drought years of 2012-2016, annual precipitation was below the median annual value for five consecutive years and there were on average only 3 large storms a year. The years 2012-2015 were without precedence in paleoclimatic history, representing a more than 20,000 year event²⁰⁸. These swings from deluge to drought are due to the presence/absence

of a high-pressure ridge off the west coast²⁷⁶, dubbed the "Ridiculously Resilient Ridge"²⁴⁹, which prevent large storms from reaching inland California⁶⁹.

2.2 GROUNDWATER NEAR C.I.LJR

C.I.LJR sits on a local topographic high, the Tejon Lookout, in the southwestern edge of the Tehachapi Mountains, where pre-Cretaceous metasedimentary rocks outcrop at the surface⁴⁰. Groundwater in the Tejon Lookout flows into the Cuddy Canyon Basin to the West, Peace Valley to the South, and Castac Lake Valley Basin (CLVB) to the north. Flow is likely constrained by the San Andreas Fault to the South and the southern branch of the Garlock Fault to the north. The CLVB is a small ($\sim 14\text{km}^2$) groundwater basin that provides drinking water for the town of Lebec, CA and irrigation for nearby agriculture. Groundwater is thought to be unconfined in the entire CLVB. Groundwater wells in the CLVB have declined by 25 m since 2008, due to the combined effects of drought and groundwater extraction for residential use, irrigation and maintaining the level of Castac Lake⁴⁴, as shown in Figure 4.3. Groundwater levels in the Peace Valley to the southeast of C.I.LJR show modest $\sim 2\text{m}$ declines over the last decade. C.I.LJR is located 2 km away from and 300 m above the nearest pumping well. I thus assume changes in groundwater levels at C.I.LJR are due to deep percolation from precipitation⁷⁹ and not due to pumping.

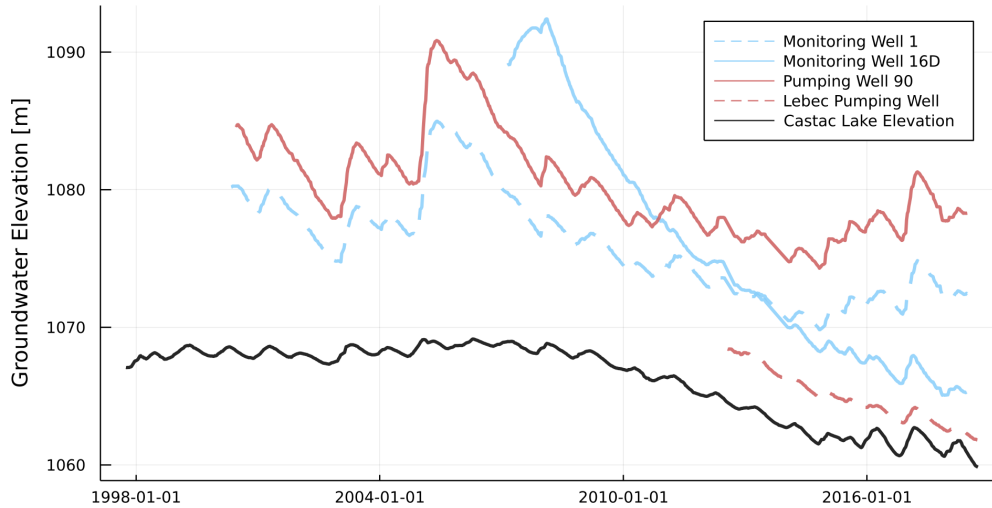


Figure 4.3: Groundwater levels in the Castac Lake Valley Basin. The monitoring and pumping wells are shown in light blue and red, respectively. The elevation of Castac Lake level shown in black.

3 METHODS

3.1 AMBIENT SEISMIC SINGLE-STATION CROSS-CORRELATIONS

I calculate daily SC functions for station CILJR using all available data from 2002 - 2021 in the Southern California Earthquake Data Center (SCEDC) public dataset. I process the east, north, and vertical components in daily chunks using the SeisIO.jl module for seismic data processing in the Julia language¹²⁷. To minimize the impact of sensor or data transmission issues, I taper data gaps with a 100 second cosine window. Daily waveforms are demeaned, detrended, and highpass filtered above 0.4 Hz before removing the instrument response. I then resample data to 40 Hz before extract windows of 30 minutes with an overlap of 75% between windows from the daily trace²²⁷. I use the SeisNoise.jl module detailed in Chapter 3 to further process the windowed ambient noise data⁵⁹. Windows are demeaned, detrended, and tapered with a 20 second cosine window. I then whiten data between 0.5 and 19 Hz and apply one-bit amplitude normalization²¹. I cross-correlate

the east-north (EN), east-vertical (EZ), and north-vertical (NZ) components in the frequency domain before transforming them back to the time domain. A daily SC is then created for each component by stacking all 30 minute windows for each day using the robust stack algorithm (Yang et al., 2021 to be submitted to Geophysical Journal International). Daily NZ correlations for station CI.LJR for lag times $\tau \in [2, 10]$ seconds are shown in Figure 4.4.

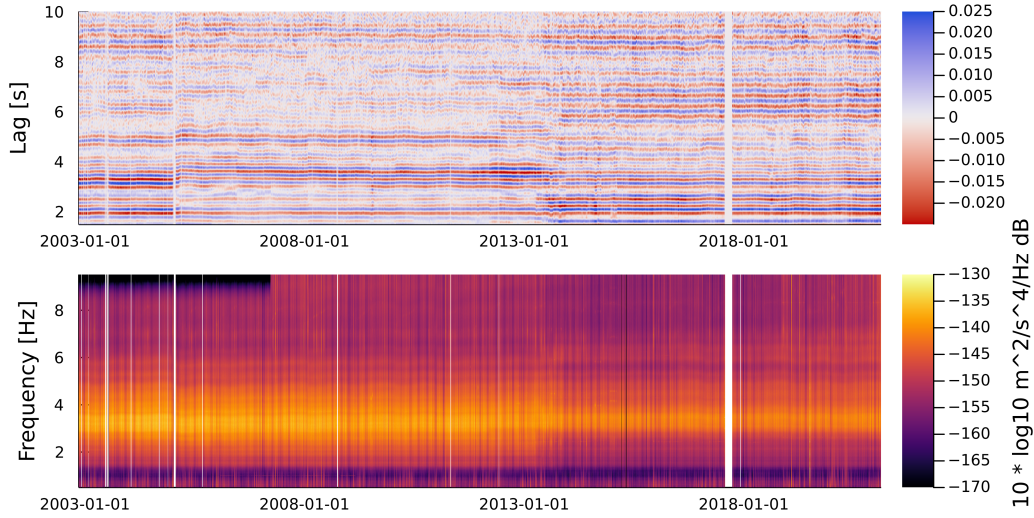


Figure 4.4: North - vertical single-station cross-correlation for station CI.LJR. Top: Daily NZ cross-correlations from 2003-2021 for lag times $\tau \in [2, 10]$ seconds in the 2-4 Hz frequency band from CI.LJR with amplitude scaled by τ . Bottom: Daily power spectral density for station CI.LJR. White regions indicate data gaps or instrument failures.

3.2 CHANGE IN SEISMIC VELOCITY, dv/v

I compute changes in seismic velocity, dv/v , using the stretching technique²³⁰, which assumes that relative time delays, $d\tau/\tau$, in the arrival of coda waves are linearly related to changes in velocity of the medium, $dv/v = -d\tau/\tau$. To recover dv/v , daily SC functions are “stretched” at times $\tau(1 - \varepsilon)$ by various stretching factors ε and compared to a reference SC waveform. The recovered velocity change is then given by the stretching factor $\varepsilon = dv/v$ that maximizes the correlation coefficient,

$$CC_i(\varepsilon) = \frac{\int_{t_1}^{t_2} SC_i[t(1-\varepsilon)]SC_0[t]dt}{\sqrt{\int_{t_1}^{t_2} SC_i^2[t(1-\varepsilon)] \cdot \int_{t_1}^{t_2} SC_0^2[t]}} \quad (4.3)$$

where SC_i is a daily SC, SC_0 is the reference SC, and t_1 and t_2 are the start and end of the coda window used to calculate dv/v , respectively⁹⁸. Before stretching, I filter the SC functions in 4 frequency octaves, 1-2 Hz, 2-4 Hz, 4-8 Hz and 8-16 Hz. Rayleigh wave sensitivity kernels for CILJR for each octave are shown in Figure 4.7. I calculate dv/v in each frequency band using 10-day stacks of SCs against the reference stack of all ~ 20 years of SCs. I apply stretching to both the causal and acausal sides of SCs in a window between $4 \cdot T_{min}$ and $16 \cdot T_{min}$ seconds lag time, where $T_{min} = 1/f_{min}$, and f_{min} is the minimum frequency in a frequency band, with trace stretching between $\varepsilon = -5\%$ and 5% , in increments of 0.001% . I then take a weighted average of dv/v values using the correlation coefficient of the stretched window for the causal and acausal sides of the EN, EZ, and NZ components,

$$dv/v_i = \frac{\sum_{k=1}^N c_{ik}^2 \cdot dv/v_{ik}}{\sum_{k=1}^N c_{ik}^2} \quad (4.4)$$

to give a single dv/v_i measurement per day, where $N = 6$ and c_{ik} is the maximum correlation coefficient between the k^{th} component of the daily and reference SCs after stretching^{108,271}, as shown in Figure 4.5.

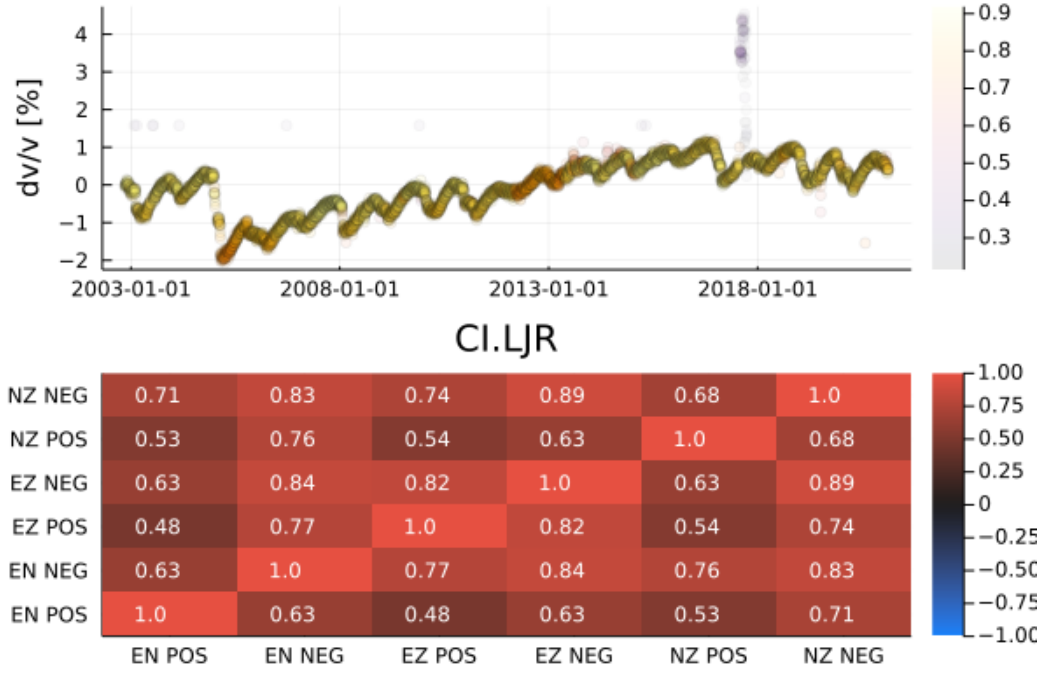


Figure 4.5: dv/v in the 2-4 Hz frequency band for CI.LJR. Top: Average dv/v for station CI.LJR using equation 4.4, where the colorbar indicates c_{ik} value. Bottom: Correlation matrix for all SC component dv/v times series.

3.3 CHANGE IN CODA ATTENUATION

Here, I propose to monitor the relative change in coda wave attenuation, Q_c^{-1} , from ambient noise SC functions. Given the source-receiver configuration, the single-station noise-derived Q_c^{-1} is similar to a local site effect. Unlike measurements of Q_c^{-1} for earthquakes, here the source is an impulse delta force at the receiver site. I follow the derivation of *Got et al.*⁹⁵, who introduced temporal monitoring of Q_c^{-1} using pairs of nearly identical earthquakes (doublets), which we call ΔQ_c^{-1} . I assume that the early coda of autocorrelations are made of surface waves^{182,184,298}. The spectral amplitude of coda waves at lapse time t and angular frequency ω is given by *Aki and Chouet*¹⁰,

$$A_i(\omega, t) = C_i t^{-\alpha} e^{-\omega t Q_c^{-1}/2} \quad (4.5)$$

where C_i is the amplitude of the single-station correlation on day i and $\alpha = 1$ for surface waves. The reference SC function is a simple linear stack over all daily SC functions. *I.e.* $A_{ref}(t) = \frac{1}{N} \sum_{i=1}^N A_i(t)$, where A_i is a daily SC function. The logarithm of the spectral ratio of an SC function on day i and the reference is,

$$\ln \left(\frac{A_i(\omega, t)}{A_{ref}(\omega, t)} \right) = \ln \left(\frac{C_i}{C_{ref}} \right) + \left(\frac{\omega}{2} \Delta Q_i^{-1} \right) t \quad (4.6)$$

where $\Delta Q_i^{-1} = Q_i^{-1} - Q_{ref}^{-1}$. I note that equation 4.6 is linear in terms of t and does not depend on the geometric spreading factor α . The workflow to measure ΔQ_i^{-1} is similar to the Moving-Window Cross-Spectrum (MWCS) technique used to measure phase-delays in coda waves⁵⁶. I first smooth $A_i(t)$ over the previous 90 days to improve signal to noise ratio. I then select 3.2 second Hanning windows in the SC functions starting at $t = 5$ second lag time. I then apply a 0.5 second spectral smoothing before computing the log spectral ratio in each window. I then slide through the lag times in 0.25 second steps until a lag time of $t = 10$ seconds. For each frequency ω , I use a weighted linear regression to solve equation 4.6 for ΔQ_i^{-1} , where the weights are the cross-coherence between $A_i(\omega, t)$ and $A_{ref}(\omega, t)$ in each window. I then take the average of each $\Delta Q_i^{-1}(\omega)$ measurement in the frequency band 2-4 Hz as the representative ΔQ_i^{-1} .

3.4 MODELING GROUNDWATER CHANGES FROM PRECIPITATION

Precipitation is one of the sources of groundwater recharge and thus can be used to estimate changes in groundwater levels. In Earth system science, inference methodologies fall into two categories: empirical and theoretical. Among the theoretical approaches to estimate groundwater level, physics-based models numerically simulate three-dimensional groundwater flow, but rely on a detailed knowledge of the aquifer and host rock properties¹⁵⁶. For areas where detailed hydrological parameters are unknown or undetermined, simpler theoretical methods, such as the pore pressure diffusion

model have been adopted by numerous authors to estimate changes in groundwater levels from precipitation^{207,276,147,276}. In this contribution, I discuss 3 main methods to predicting ground water: two based on poro-elasticity and one based on empirical measurements.

The simplest models of groundwater level changes couple Darcy's or Boussinesq's equations with precipitation input. *Sens-Schönfelder and Wegler*²³⁰ developed a model for groundwater levels after precipitation based on the assumption that under a linearized Dupuit-Boussinesq flow, drainage occurs exponentially as,

$$\Delta h_i(t) = \sum_{n=0}^i \frac{p(t_n)}{\phi} e^{-a(t_i-t_n)} \quad (4.7)$$

where ϕ is the porosity and $p(t_i)$ is the amount of precipitation on day t_i . This model approximates the classic baseflow recession curve $Q = Q_0 e^{-at}$, where Q is the rate of flow, t is time, Q_0 is the flow when $t = 0$, and a is a constant that depends on the time scale of recession²⁵⁶.

Poro-elasticity couples pore pressure, stress fields, and input source terms²²⁸. *Roeloffs*²¹² calculated the coupled poro-elastic response at depth r due to a load of amplitude p_0 at the surface as,

$$P(r, t) = \frac{B(1 + \nu_u)}{3(1 - \nu_u)} p_0 \operatorname{erf} \left[\frac{r}{(4ct)^{1/2}} \right] + p_0 \operatorname{erfc} \left[\frac{r}{(4ct)^{1/2}} \right], \quad (4.8)$$

where erf and erfc are the error and complementary error functions, respectively, c is the diffusivity of porous material, t is the time since the load was applied, ν_u is the "undrained" Poisson's ratio, and B is the Skempton's coefficient. B is close to 1 at the surface and decreases with depth²¹¹. The first term on the right side of equation 4.8 is the undrained poro-elastic response due to elastic loading, whereas the second term on the right side of equation 4.8 is the drained poro-elastic response due to diffusion. The medium response is "undrained" when there is no fluid flow in response to a change in stress $\Delta\sigma_{ij}$ ²⁰². At zero lag time, the response is undrained, while at infinite lag time, the response is fully drained. Earthquakes and elastic loading from precipitation are natural examples of

stresses that stimulate undrained responses.

*Talwani et al.*²⁵⁷ modified *Roeloffs*²¹²'s model to accommodate the change in pore pressure at depth due to a series of precipitation loads, given by,

$$P_i(r, t) = \frac{B(1 + \nu_u)}{3(1 - \nu_u)} \sum_{i=1}^n \delta p_i \operatorname{erf} \left[\frac{r}{(4c(n-i)\delta t)^{1/2}} \right] + \sum_{i=1}^n \delta p_i \operatorname{erfc} \left[\frac{r}{(4c(n-i)\delta t)^{1/2}} \right] \quad (4.9)$$

where $n \cdot \delta t$ is the number of days since the start of the rainfall time series and $\delta p_i = \rho g \delta h$ is the groundwater load change variation due to precipitation δh on day i . *Rivet et al.*²⁰⁷ and *Wang et al.*²⁷⁶, among others, have used the drained part of equation 4.9 to model pore pressure changes due to precipitation at depths down to 8 km, which is reasonable due to the low values of B at these depths. On the contrary, near-surface groundwater level changes, such as I expect at CLJR, may rather be the result of an undrained response.

Recently, *Smail et al.*²³⁹ introduced the empirical approach of cumulative deviation from the moving mean (CDMk) of precipitation to estimate deviations in groundwater level from precipitation measurements alone. The CDMk method assumes that groundwater levels respond to deficits or surpluses of precipitation in the last k days, where $k \gg 365$, which is a rough approximation to Darcy's law. Given a daily precipitation time series, P_i , the CDMk for each day i is simple to compute,

$$CDM_{ik} = \sum_{j=1}^i P_j - \bar{P}_{ik} \quad (4.10)$$

where $P_i - \bar{P}_{ik}$ is the daily deviation from the moving or rolling mean $\bar{P}_{ik} = \frac{1}{k} \sum_{j=i-k+1}^i P_j$ of k days. CDMk time series for moving mean values of 1, 6, 24 and 92 months are shown in Figure 4.6. Increasing k increases the memory of groundwater to longer-term trends in precipitation. *Smail*

*et al.*²³⁹ found that CDM_k of 60 months correlated well to groundwater levels in both bedrock and unconfined aquifers but had no correlation to levels in highly confined aquifers. The CDM_k and *Talwani et al.*²⁵⁷ models are similar. In fact, the *Talwani et al.*²⁵⁷ model evaluated at $r = 0$ m converges to the CDM_k with $k = \infty$, or just the cumulative deviation from the mean of precipitation.

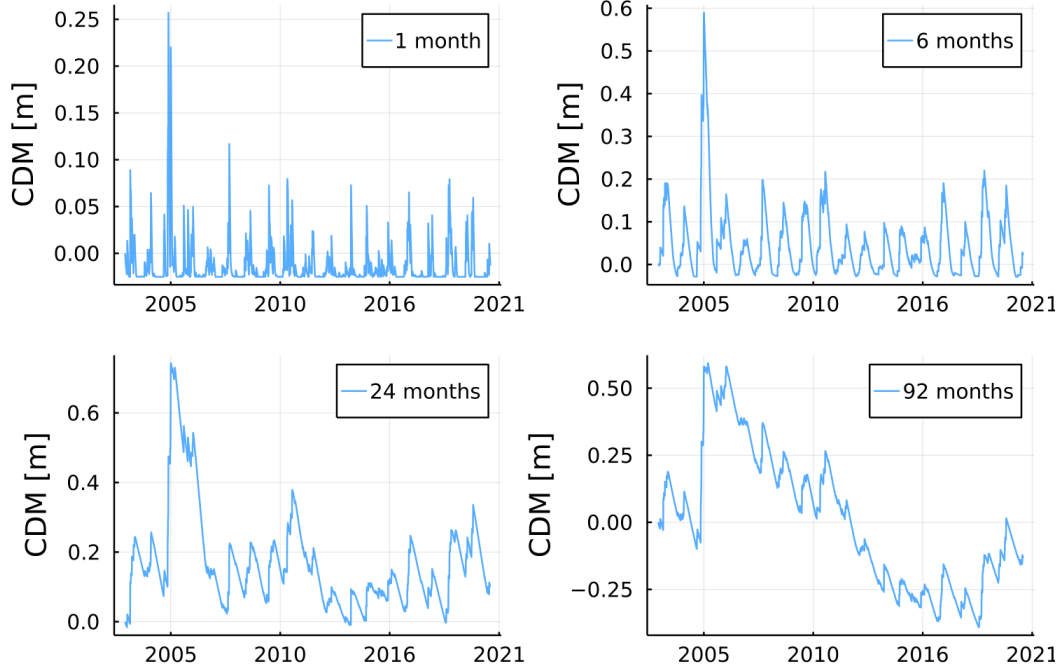


Figure 4.6: Cumulative deviation from the moving mean (CDM_k) of precipitation using 1, 6, 24, and 92 month moving means.

I test the baseflow model, the undrained poro-elastic model, and the CDM_k method to assess the influence of precipitation on the perturbation in seismic velocities at C.I.L.J.R. I calculate the best moving mean value k by correlating CDM_k against dv/v ,

$$CDM_{best} = \max_k \frac{cov(dv/v, CDM(t_k))}{\sigma_{dv/v} \sigma_{CDM(t_k)}}, \quad (4.11)$$

using k values from 1 to 14 years. I similarly calculate a best fitting baseflow and poro-elastic

models by varying a (in equation 4.7) and c (in equation 4.9).

3.5 MODELING GROUNDWATER TIME SERIES FROM dv/v

Here, I model the change in groundwater levels at station CI.LJR, given our measured dv/v and the transfer coefficients developed in Chapter 1 Sections 3 and 4. I recall the relation between dv/v and the changes in hydraulic head or groundwater levels Δb , $dv/v = \frac{\beta S_{sk} \beta}{B} \Delta b$. My model depends on the undrained Poisson ratio ν_u , the Skempton's B coefficient, the effective rigidity G , and the acoustoelastic constant β , of which B and β are the least constrained. CI.LJR sits atop the Tejon Lookout granite, a salmon-colored biotite granite that is interspersed with metasedimentary rocks, such as marbles, quartzites, slates, phyllites, schists and limestones¹⁸⁶. *Hart and Wang*⁹⁹ measured values in the range of 0.22 - 0.26 and 0.28-0.31 for the drained and undrained Poisson's ratios ν and ν_u , respectively, of Indiana Limestone, while granite has a Poisson's ratio of ~ 0.2 at 10 MPa¹¹⁹. Given my uncertainty of the poro-elastic properties of the rock at CI.LJR, I use the formula given by *Christensen*⁵² for the undrained Poisson's ratio as a function of V_p/V_s to get a $\nu_u = 0.27$. Given that Skempton's B coefficient has been measured to be 1 at low effective stress²¹¹, I assume $B = 1$ for groundwater recharge near the surface. I derive a shear modulus $G = 20 \text{ GPa}$ using $G = V_s^2 \rho$ from the Southern California Earthquake Center velocity model. Using these values, $g = 9.81 \text{ m/s}^2$, and $\rho_0 = 1000 \text{ kg/m}^3$ for the pore fluid, gives a value of $S_{sk} = 2.7 \times 10^{-7} \text{ m}^{-1}$ for the skeletal specific storage, which is $\sim 35\%$ less than the value $S_{sk} = 9.3 \times 10^{-6} \text{ m}^{-1}$ *Riley*²⁰⁶ reported in Central California.

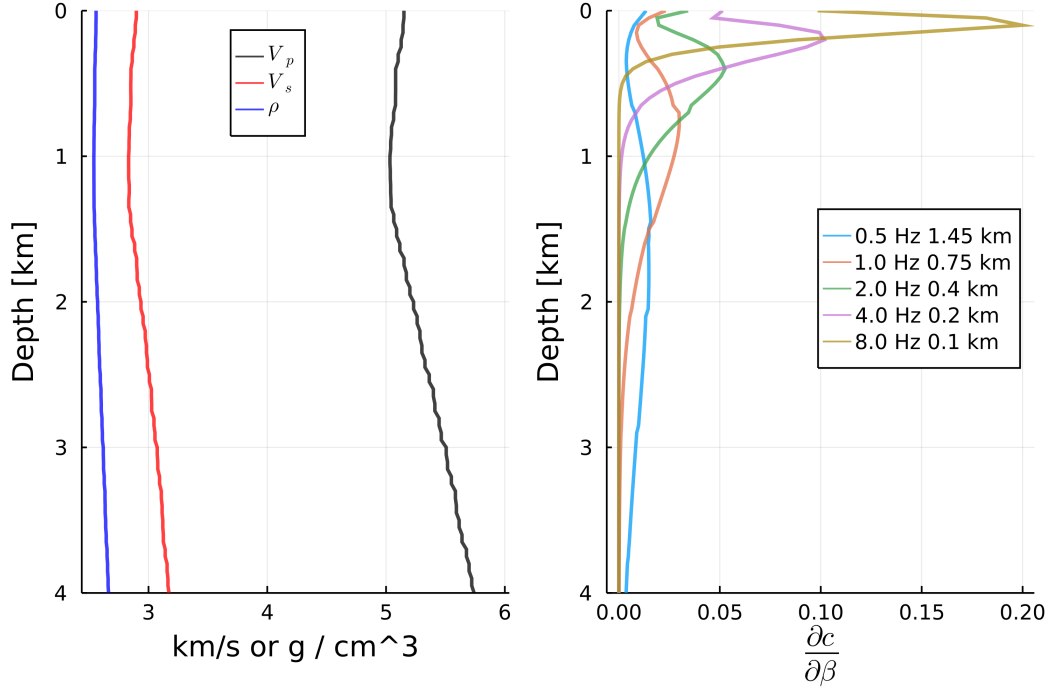


Figure 4.7: Surface wave sensitivity for CI.LJR. Left: Velocity model for CI.LJR from SCEC Community Velocity Model. Right: Surface wave sensitivity kernels ($\frac{\partial c}{\partial V_s}$ = partials of the phase velocity with shear velocity for fixed frequency and density) for CI.LJR computed with Computer Programs in Seismology¹⁰².

Choosing an accurate value of β is difficult without *a priori* information on the local volumetric strain, which in the case of CI.LJR is likely dominated by groundwater changes. Given the range in experimental values for $|\beta| = 1 \times 10^3 - 6.9 \times 10^4$ found by previous authors (*Takano et al.*²⁵⁵, *Sens-Schönfelder and Eulenfeld*²²⁹, *Mao et al.*¹⁵³) I test values of $\beta = [5 \times 10^3, 7.5 \times 10^3, 1 \times 10^4]$ with our model. I find that a $\beta = 3.7 \times 10^3$ explains the relation between our measured dv/v and the change in groundwater level at a well in the CLVB 3 km from CI.LJR, as shown in Figure 4.8. This β is over an order of magnitude higher than the $\beta = 2.2 \times 10^2$ value reported by *Nur and Simmons*¹⁷⁸ for Barre granite in a laboratory, which suggests that the groundwater level change at CI.LJR is a factor of 10 or so less than in the CLVB. Further measurements of Murnaghan's constants in a wide variety of rocks will lead to better constraints on β .

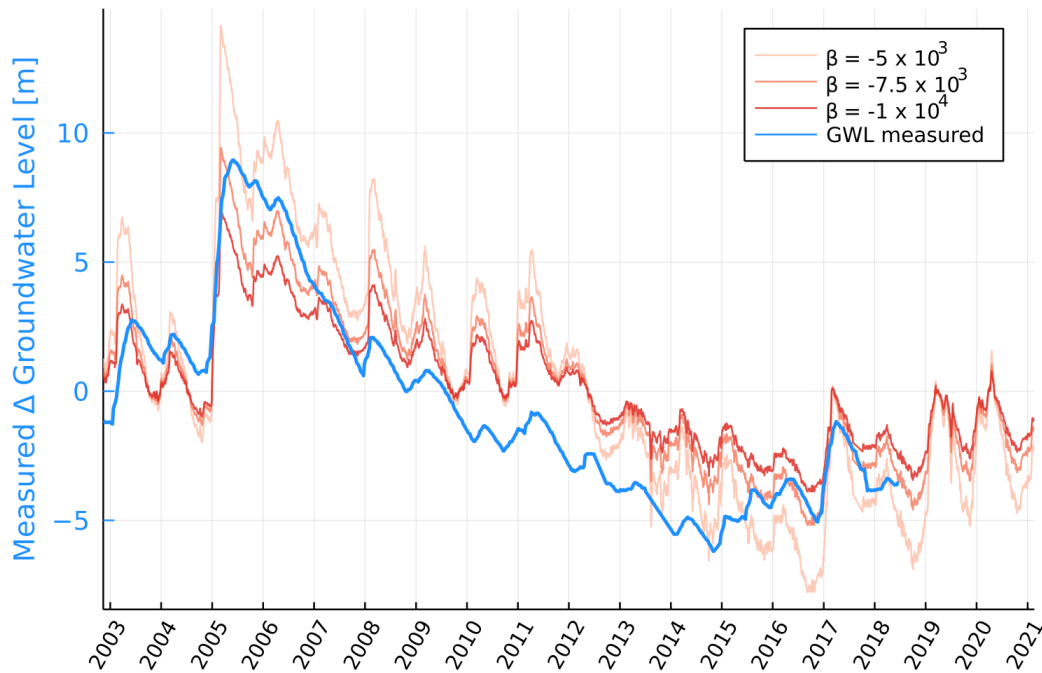


Figure 4.8: Groundwater level changes measured at monitoring well MW1 compared to modeled groundwater levels using equation 1.15.

4 RESULTS AND DISCUSSION

I see three distinct contributions to dv/v changes measured at CI.LJR: (1) rapid velocity decreases due to sudden groundwater recharge, (2) long-term velocity increases due to drought, and (3) seasonally varying velocities that are explained by groundwater level fluctuations. These trends are similar to our previous work⁵⁸ in the San Gabriel valley, where groundwater levels declined in response to the record setting drought of 2012-2016. Here, I focus our discussion of dv/v to the 2-4 Hz frequency band, which has the high and most stable power of any frequency band. Short and long-term trends in dv/v associated with groundwater at CI.LJR are discussed in Chapter 5.

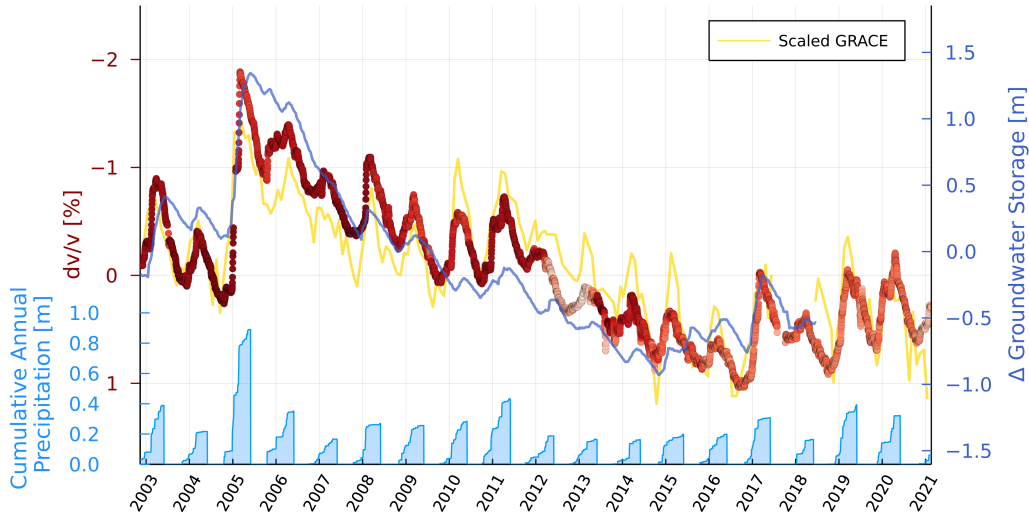


Figure 4.9: dv/v compared to CDMk of precipitation, GRACE LWE and groundwater level near station CI.LJR. Left axis: dv/v (red dots colored by correlation coefficient to stack), scaled CDMk of precipitation (chartreuse) with correlation coefficient = -0.97, scaling factor = -42.1, and scaled GRACE LWE (orange) with correlation coefficient = -0.88, scaling factor = -20.7. Right axis: Groundwater level w.r.t surface (blue with cyan dots) for a well 10 km away from CI.LJR with correlation coefficient = -0.90, scaling factor = -79.3. Bottom left axis: Cumulative annual water year precipitation (Oct 1 - June 1) for PRISM grid cell containing station CI.LJR.

4.1 SEASONALITY

I observe moderate levels of seasonality in dv/v at CI.LJR in the 2-4 Hz frequency range. Previously, *Meier et al.*¹⁶⁰ measured seasonal changes in dv/v across Southern California in the 0.1-0.2 Hz band and suggested that seasonality was more likely to come from thermo-elastically induced strains than poro-elastically induced strains, but did not investigate correlations with measured temperature or groundwater variations. *Hillers and Ben-Zion*¹⁰³ found that seasonal variations in the amplitude of ambient noise in Southern California in the 2-18 Hz band were dominated by annual temperature and wind speed changes and not seasonal groundwater recharge, and ruled out groundwater variations without investigating. Seasonal variations in dv/v at CI.LJR should not be due to seasonal changes in noise source or amplitude, as the station is located within 1 km of traffic on interstate I-5, where ~ 1 vehicle per second entered the Tejon Pass from the north and south in 2019 (data

accessed from California Department of Transit). Tsai²⁶⁸ compared Meier et al.¹⁶⁰'s dv/v measurements to thermo-elastic, poro-elastic, and hydro-elastic models of dv/v , and found that either the poro-elastic or direct elastic response from water table variations were more likely to induce seasonal variations in dv/v than thermo-elastic strains at the particular site of investigation of the Los Angeles Basin¹⁶⁰.

To assess the contributions of temperature or precipitation-driven seasonality, I cross-correlate daily mean temperatures from PRISM, dv/v , and CDMk filtered between 0.5 and 1.5 year period. Filtered daily temperature has a maximum Pearson's correlation coefficient of 0.93 with dv/v at a lag of 63 days while dv/v and filtered CDMk have correlation coefficient -0.99 at zero lag. This compares well with the 55-day and 0-day lags Tsai²⁶⁸ found between temperature and hydrological changes, respectively, and dv/v in the Los Angeles basin. From 2012-2016, the seasonal amplitudes of dv/v are reduced, as compared to the non-drought years 2008-2011 and 2017-2019. Due to the higher correlation between seasonal dv/v and CDMk than dv/v and temperature, and the correlation of low seasonal dv/v amplitudes during the drought of 2012-2016, I suggest that seasonality in dv/v is controlled low-level and low-variance annual precipitation levels at C.I.L.J.R.

4.2 COMPARISON OF dv/v TO GROUNDWATER PROXIES

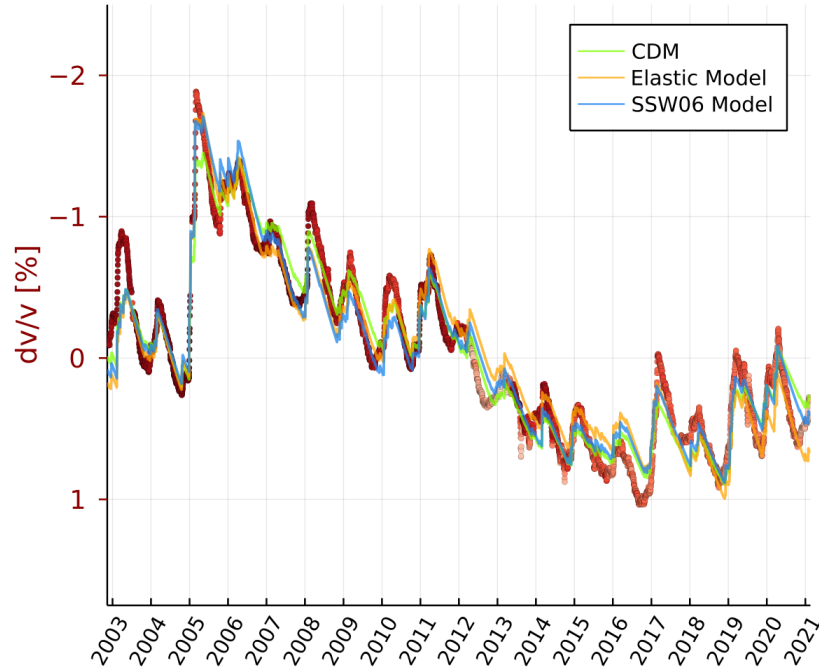


Figure 4.10: dv/v compared to CDMk of precipitation (chartreuse), elastic model of *Talwani et al.*²⁵⁷ (orange) and model of *Sens-Schönfelder and Wegler*²³⁰ (SSW06 - blue) at station CI.LJR. All models have been scaled by fitting coefficients.

The tight relationship between dv/v and GRACE liquid water equivalent changes suggest that hydrologic changes are the main driver of dv/v at CI.LJR. I confirm this with the three groundwater - precipitation models mentioned in Section 3.4. The baseflow model of *Sens-Schönfelder and Wegler*²³⁰ and (CDMk) of *Smail et al.*²³⁹ are correlated with dv/v with Pearson correlation coefficient ≈ -0.97 at zero lag. These models suggest long-term memories for past precipitation - the best fitting k for the CDMk model was 2,819 days or 7.7 years, while the best-fitting a constant for the baseflow model was 0.0008 days^{-1} , or a half-flow period of ~ 900 days. The fully-coupled poro-elastic model of *Talwani et al.*²⁵⁷ did not fit the observed dv/v , though a purely undrained model, obtained by disregarding the drained response, fit dv/v well at zero lag (Pearson correlation

coefficient with $dv/v = -0.96$). This strongly suggests that dv/v at 2-4 Hz at C.I.L.J.R responds to the elastic load due to precipitation and not the diffusion response. I retrieve a diffusivity constant $c = 0.0038 \text{ m}^2 \text{ s}^{-1}$ from the undrained model, which indicates very slow flow. This value of c falls in between the ranges found for unfractured and fractured igneous rocks²¹¹. The stress-strain relationship at C.I.L.J.R is likely recoverable, as the hysteresis curve between dv/v and CDMk is linear, with the exception of the transition into the drought in 2012-2013, as shown in Figure 4.11.

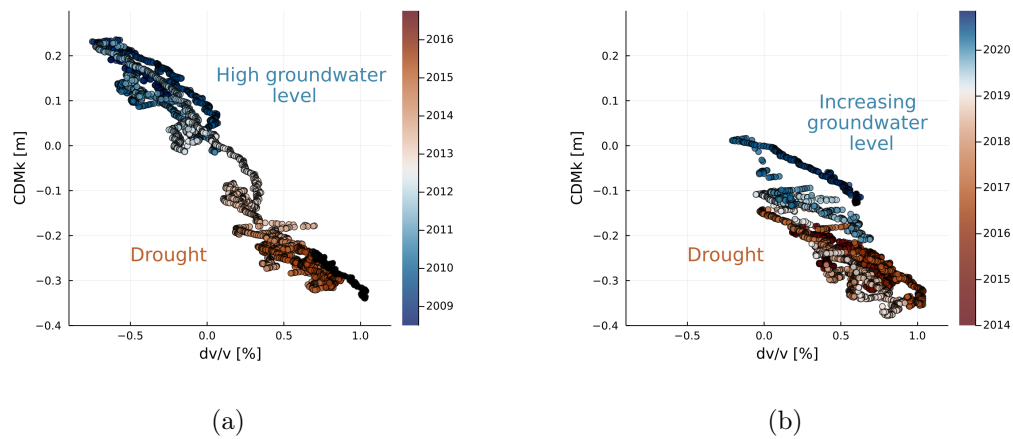


Figure 4.11: Hysteresis curve for dv/v and CDMk. (a) Hysteresis between 2008-2016. Transition period starts around 2012-2013 when dv/v increases above previous high. (b) Hysteresis between 2014-2021. Slight recovery in groundwater levels suggests recoverable deformation.

4.3 RELATIVE CHANGE IN CODA ATTENUATION, ΔQ_c^{-1}

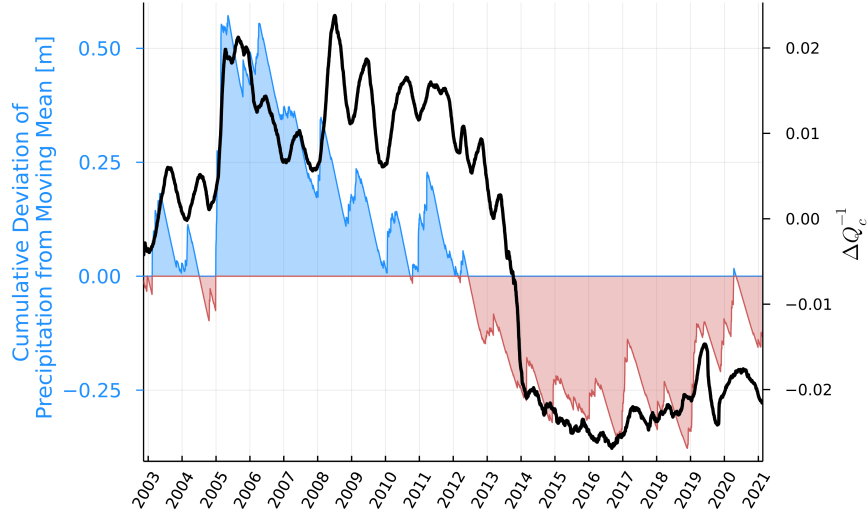


Figure 4.12: Cumulative deviation from the moving mean of precipitation and average ΔQ_c^{-1} measured on EN and NZ components of C.I.LJR for lag times $\tau \in [5, 10]$ seconds in the 2-4 Hz frequency band.

The temporal change in coda attenuation (ΔQ_c^{-1}) derived from SC functions with the method introduced in Section 3.3 suggests that attenuation variations are related to groundwater. Previous authors have documented 20–30% temporal changes in attenuation in California^{247,152}. At station C.I.LJR, ΔQ_c^{-1} increased with influx of groundwater in the winter of 2004-2005. An increase in Q_c^{-1} is expected, and visually apparent from $\tau \in [5, 10]$ in Figure 4.4B, as laboratory measurements have shown attenuation increases with increasing saturation^{126,291}. ΔQ_c^{-1} values remained high through 2013, as precipitation was higher than average during this period, as shown in Figure 4.12.

We see a large decrease in attenuation in early 2013 when dv/v goes above its highest previous level. We expect that groundwater levels achieved all-time lows at this point. When groundwater level falls below its previous low, the aquifer skeleton may undergo permanent rearrangement⁸⁹, which results in some anelastic compaction. We suggest some form of compaction in less-permeable layers underlying C.I.LJR occurred in 2013, resulting in a drying of the pores, which led to a rapid

decrease in attenuation. This effect is most likely to occur in the near surface, where lithostatic pressure is low - attenuation decreases with increasing pressure and decreasing saturation¹²⁶. Our results agree with recent findings from *Malagnini and Parsons*¹⁵² that Q_c^{-1} changes at the San Andreas Fault at Parkfield are modulated by groundwater level changes.

These observations could have a large impact on localized seismic hazard estimates, where site effects come from a combination of amplification and attenuation³⁰. Seismic amplification is related to the ratio of seismic impedance (density times velocity) at depth to the seismic impedance averaged near the surface. For a unit decline in groundwater level, the change in velocity for rocks with rounded pores is quite small, on the order of a few percent²⁹¹. The resulting change in density due to drying depends on the specific yield. The decrease in attenuation due to a unit decline in head should be an order of magnitude greater than the change in velocity, on the order of 10 – 30%. The spectral acceleration of earthquake ground motions is given by

$$a(f) = A_0 e^{-\pi f t^*} \quad (4.12)$$

where A_0 depends on source properties, and t^* is an attenuation time for seismic waves that travel as rays,

$$t^* = \int \frac{Q_S^{-1}(r)}{V_S(r)} dr \quad (4.13)$$

where Q_S^{-1} is the shear wave attenuation factor and V_S is the shear wave velocity with depth. As V_S decreases rapidly near the surface, t^* is most sensitive to near surface Q_S^{-1} . This suggests that decreases in groundwater levels will result in decreased attenuation in future ground motion. Further research is required to assess if a groundwater-attenuation mechanism could cause anomalously high ground motions¹⁶⁴.

5 CONCLUSION

I showed that monitoring the change in seismic velocity, dv/v , at a single seismometer is similar to measurements made by groundwater wells, GRACE, and precipitation modeling. Unlike satellite methods, our method directly samples the subsurface and its depth sensitivity can be tuned based on frequency band. I suggest that the cumulative deviation from the moving mean of precipitation is a simple and effective technique to infer the effect of precipitation on dv/v . The combination of dv/v , CDMk and GRACE can be used to predict/calibrate groundwater level changes where groundwater level measurements are sparse, whereas precipitation records are readily available in the United States⁶⁰ and available at varying resolution globally²⁴⁸.

I find that groundwater levels in the vicinity of C.I.L.J.R. are strongly linked to the variance in precipitation from large storms coming to California. I explored three separate models for predicting groundwater levels from a precipitation time series. I find that models based on Boussinesq flow, Darcy's law and poro-elasticity readily explain observed groundwater levels and dv/v at C.I.L.J.R. Importantly, I found that the undrained response dominates over the drained response for near-surface recharge, which I do not believe has yet been observed.

Initial measurements of the temporal change in coda attenuation, Q_c^{-1} from single-station cross-correlations suggest that near surface attenuation is also modulated by changes in pore pressure. With the drying of California expected to continue in the future, we expect reduced attenuation could lead to stronger ground shaking from future earthquakes.

5

The Seismic Signature of California's Extreme Events

Water, water, everywhere,
And all the boards did shrink;
Water, water, everywhere,
Nor any drop to drink.

Samuel Taylor Coleridge

I INTRODUCTION

California is subject to substantial extreme natural events. It hosts large magnitude 7 earthquakes²¹⁶, its climate is highly variable with short lived, but extreme precipitation events as atmospheric events⁷⁰, and multi-year droughts²⁷⁶. As Earth's climate is warming, volatility in precipitation is expected to increase in Mediterranean climates²⁵². In the last decade California has been emblematic of this predicted trend: the 2012-2016 drought was unprecedented in the observational record²⁵⁰, with the lowest three-year rainfall recorded in the last hundred years. Snowpack levels in the Sierra Nevada mountains were 95% below average in 2015, due in part to record temperatures and inadequate snowfall⁴². The 2012-2016 drought was followed by one of the wettest January and Februaries on record in the winter of 2017²⁷⁶.

Rapid shifts from deluge to drought will increasingly stress groundwater supplies in California²⁵⁸. Groundwater is often a last resort when imported water and precipitation are scarce during drought - Declining groundwater levels lead to a permanent loss in groundwater storage due to compaction²⁴⁰ and to dry wells in rural areas¹⁹². Satellite observations indicate groundwater levels are in decline once again in California after increased rainfall in the winter of 2017¹⁷².

The coupling between earthquakes and groundwater levels is well established²⁷⁴. At time scale of seconds, strong ground motion from earthquakes can cause groundwater level changes in distant wells³⁸ and rapidly increase discharge of groundwater into streams²⁷³. Recent evidence suggests earthquakes may temporarily increase aquifer permeability²⁷⁶, which leads to flow. On the annual scale, seasonal elastic hydrological loading may drive seismicity in California^{15,123} and Taiwan¹¹⁰.

Here, we investigate the impact of declining groundwater levels and earthquakes on seismic wave propagation in California. Changes in groundwater level (pore pressure) affect the phase of seismic waves. Phase delays of seismic waves are the result of changes in seismic velocity. In aquifers, there is an inverse relationship between saturation (groundwater level) and seismic velocity¹⁸⁵. Seismic

velocity changes due to recharge/decline in groundwater level have been measured with ambient noise cross correlation^{230,58}.

2 DATA AND METHODS

Seismic monitoring has occurred in California for nearly 100 years, with digitized, continuous measurements going back to 1999¹¹³. Recently, the Southern California Earthquake Data Center (SCEDC) uploaded their entire catalog of continuous seismic data, totaling more than 100 Terabytes (TB), from the Southern California Seismic Network (SCSN) as a Public Data Set (PDS) on Amazon Web Services²⁹⁷. Using data from the SCEDC PDS and the Northern California Seismic Network, we analyzed all available continuous seismic data from broadband seismometers located in California going back to 1999. We computed daily single-station cross-correlations (SC) for all stations in the dataset with more than 1 year of data using the same parameters and methods as detailed in Chapter 4 and my Julia package `SeisNoise.jl`⁵⁹. SCs approximate the reflection seismogram with ambient seismic data, allowing for continuous monitoring of the subsurface at high frequency⁵⁴. We compute relative velocity changes, dv/v , for all stations in our dataset in the 2-4 Hz frequency band²⁰⁴. These dv/v changes are sensitive to the upper 500 m of the subsurface (see Figure

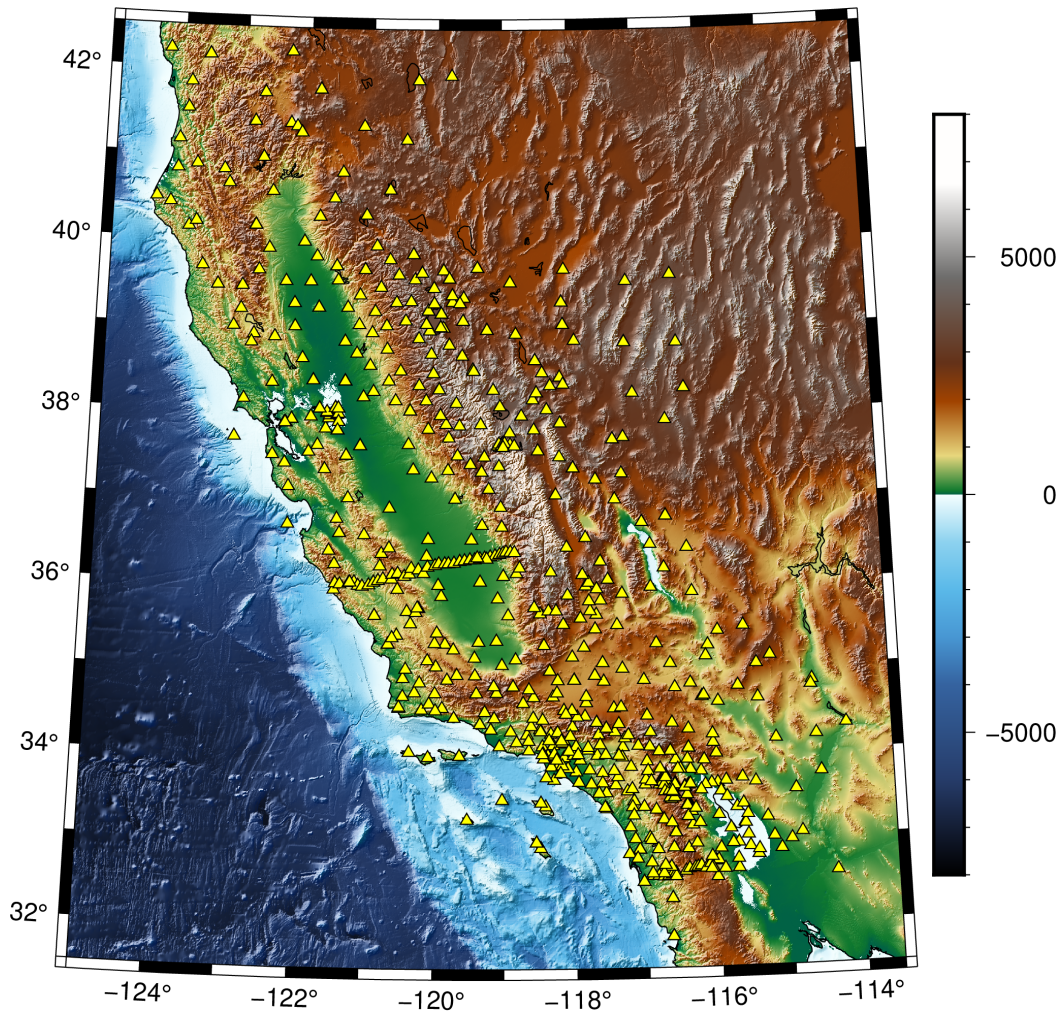


Figure 5.1: Location of all seismometers (≈ 700) used in this study. Time of observations is 1999-2021.

3 RESULTS

Overall, the relative contributions between the tectonic, thermal, and hydrological strains vary across sites. The spatial coherence between these effects is related to the location and intensity of the events. For instance, the deluge of precipitation in the winter of 2004-2005 lowered seismic velocities across most of Southern California⁷⁰, while the effects of the M7.0 2019 Ridgecrest earthquake

were limited to within less than 50 km from the epicenter of the earthquake.

Hydrologic modulation is the dominant factor in seasonal and long-term dv/v in California. Atmospheric rivers brought record-setting rainfall to Southern California in the winter of 2004-2005, with 11 separate storms sweeping across the region in 6 months^{195,176}. Cumulative rainfall for the time period was 60 cm greater than the mean, which is the highest deviation in the last 20 years. Groundwater levels in the Los Angeles Basin increased by 20m in response to the extreme precipitation^{130,58}, leading to a more than 1% decrease in dv/v at stations in coastal Southern California. At station CI.LJR, dv/v decreased more than 1% following a set of storms on December 27th-29th, January 2nd-4th, and January 7th-11th and more than 0.85% following a single storm on February 17th-23rd. These dv/v decrease are similar in amplitude to the dv/v decrease *Sens-Schönfelder and Wegler*²³⁰ measured for a 40 m increase in groundwater level at Merapi volcano, Indonesia. Similar decreases in dv/v have been recorded at close distances to large earthquakes¹⁸¹, though we rule out earthquakes as a cause, as no $M_w > 6$ earthquakes occurred within 1,000 km of CI.LJR from October 2004 to May 2005. The increase in storage in the Los Angeles Basin after winter 2004-2005 was $\Delta S \approx 2m$, assuming a basin-wide specific yield $S_y = 0.15$. This was $5 \times$ greater than the 30 cm storage change measured by GRACE in the same time-frame averaged over the Los Angeles region. Previously, *Clements and Denolle*⁵⁸ estimated that 0.5 km³ of groundwater was lost during the drought in the San Gabriel Basin from 2012-2016. The undrained poroelastic response to this deluge is the most likely mechanism for the decrease in dv/v at shallow depths during this time, as shown in Chapter 4. At greater depths though, the drained poroelastic response is expected to dominate^{207,276}.

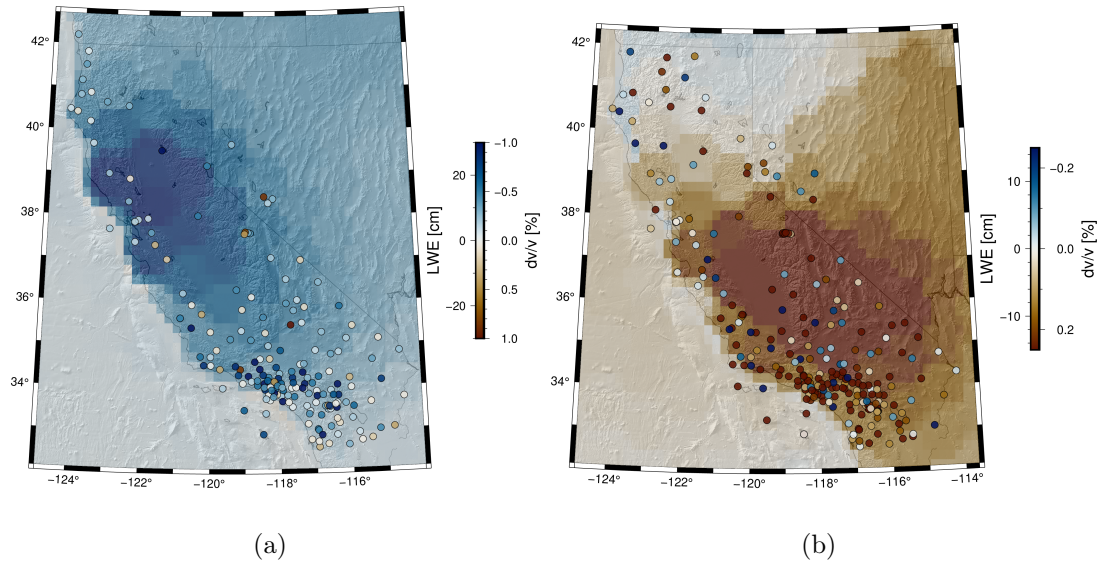


Figure 5.2: Changes in GRACE Liquid Water Equivalent and dv/v across California in 2-4 Hz frequency band. (a) Change in dv/v measured between January 2005 and May 2005 (scatter points) and GRACE LWE (shading) over the same time frame. (b) Change in dv/v and GRACE LWE between September 2011 and September 2016.

Following the precipitous drop in dv/v in the winter of 2004-2005, dv/v steadily increased 2.5% from 2005-2017. This period contained two major droughts, the first from 2007-2009 and the second 2012-2016. The long-term increase in dv/v from 2005-2017 was controlled by the lack of significant storms - years with less precipitation from large storms (wettest 10% of days mentioned in Chapter 4 Section 2.1 had smaller decreases in dv/v between October and May than years with fewer large storms (correlation coefficient = 0.93). Similarly, summertime increases in velocity were greater after years with larger amounts of precipitation than drought years, suggesting that summertime increases in velocity are modulated by discharge and not temperature. As seen in Figure 5.3A, seasonality in dv/v was greatly reduced during the drought, as compared to before and after. A short atmospheric river in 2017 brought much needed rain to Southern California²⁸⁸ but represented only a brief interlude in the long-term increase in dv/v at C.I.LJR since 2005. dv/v at C.I.LJR remains stabilized at its 2016 end-of-drought level from 2017-now.

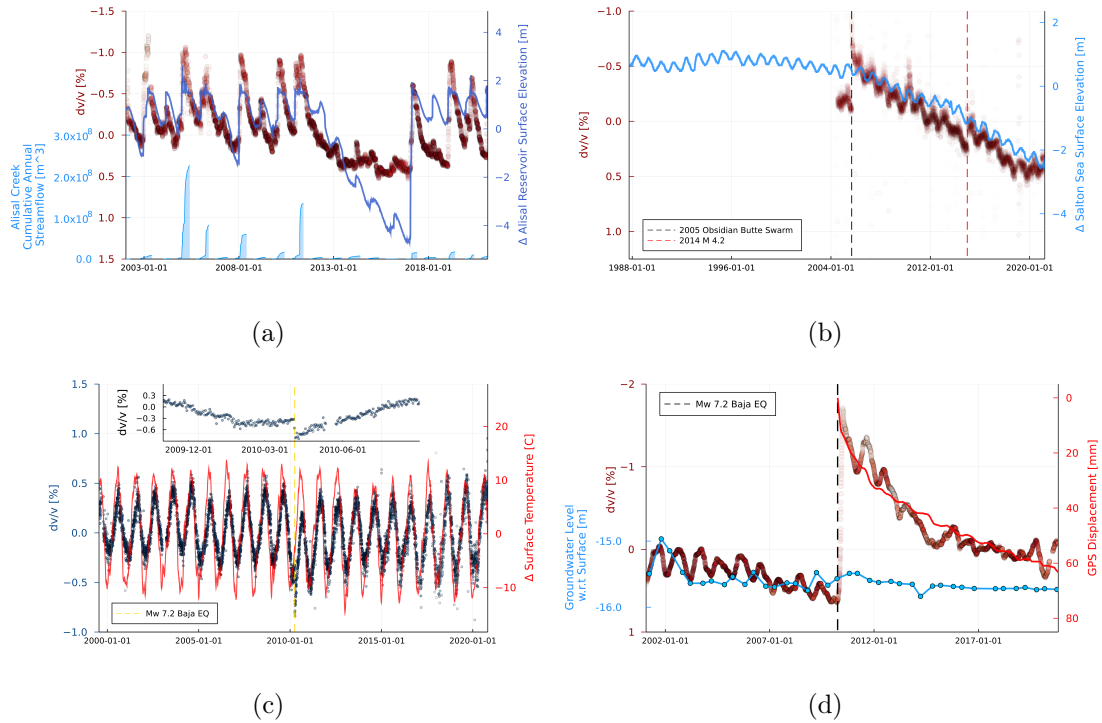


Figure 5.3: Representative dv/v time series. (a) dv/v at CI.NJQ compared with surface level of nearby Alisal Reservoir and stream flow from Alisal creek. (b) dv/v at CI.RXH and elevation of the Salton Sea. Dashed vertical lines are timing of nearby earthquakes. (c) dv/v and surface temperature change at CI.SAL. Dashed line indicates timing of 2010 Mw 7.2 Baja Earthquake. (d) dv/v at CI.WES, groundwater level at USGS well 324603115480501 and root-mean-square E-N displacement at GPS station P494 after 2010 M_w 7.2 Baja Earthquake.

dv/v is strongly correlated with seasonal temperature change in dry areas such as the Mojave desert, with the lone exception of the drying of the Salton Sea. At station CI.RXH, which is located 100 m inland from the southeast edge of the Salton Sea, dv/v has been steadily increasing since 2005 as the sea levels have dropped more than 2 m, as shown in Figure 5.3B. Station CI.SAL in Salton City, as shown in Figure 5.3C, exemplifies the near perfect sinusoidal change in dv/v modulated by thermoelastic strain induced by atmospheric temperature changes²⁰. We do not observe any long-term trends in dv/v due to thermoelastic strain, though extreme temperatures were thought to account for up 25% of the drought’s moisture deficit.

In California's inland areas, $M_w > 6$ earthquakes occur on average once every three years. Earthquakes complicate groundwater monitoring using dv/v - strong ground motion during earthquakes is thought to open cracks in the near surface and allow for increased groundwater flow^{213,38}, which decreases seismic velocities^{185,217}. Large drops in dv/v due to earthquakes look similar to rapid recharge events. Cross-comparison of dv/v drops to earthquake catalogs and groundwater proxies can rectify groundwater vs seismic causation. In the case of CI.RXH, the 2005 Obsidian Butte swarm masked an increase in water level in the Salton Sea¹⁴⁸, as shown in Figure 5.3B. In the Imperial Valley, dv/v at station CI.WES decreased 2.5%, while CI.SAL saw more modest 0.3% drop in dv/v , following the 2010 Baja earthquake. The $M_w 7.1$ 2019 Ridgecrest earthquake was the largest earthquake in the last 20 years in the state of California - though noticeable drops in velocity are only seen within 25 km of the epicenter³¹.

4 CONCLUSIONS

We measured relative seismic velocity changes, dv/v across California using single-station cross-correlations from 1999-2021. dv/v time series in the 2-4 Hz frequency have remarkable sensitivity to near-surface changes. The dominant signal in dv/v in California since 1999 is a long-term increase in velocity due to receding groundwater levels, punctuated by drops in velocity from recharge from large storms. This temporal pattern is most coherent in Southern California's coastal basins. We anticipate that future changes in seismic velocity in California will be modulated by swings between floods and droughts.

6

Earthquake Detection at the Edge

The Edge...There is no honest way to explain it
because the only people who really know where it is
are the ones who have gone over.

Dr. Hunter S. Thompson

Earthquake detection is the critical first step in Earthquake Early Warning (EEW) systems. For robust EEW systems, detection accuracy, detection latency and sensor density are critical to providing real-time earthquake alerts. Traditional EEW systems use fixed sensor networks or, more recently, networks of mobile phones equipped with micro-electromechanical systems (MEMS) accelerometers. Internet of things (IoT) edge devices, with built-in machine learning (ML) capable microcontrollers, and always-on, always internet-connected, stationary MEMS accelerometers provide the opportunity to deploy ML-based earthquake detection and warning using a single-station approach at a global scale. Here, we test and evaluate deep learning ML algorithms for earthquake detection on Arduino Cortex M4 microcontrollers. We show the trade-offs between detection accuracy and latency on resource-constrained microcontrollers for possible use as edge-connected seismometers.

1 EARTHQUAKE DETECTION

Earthquake detection is the process of picking earthquake signals from a continuous stream of recorded ground velocity or acceleration measurements. Most often this is in the form of distinguishing P (primary) waves and S (shear) waves from seismic noise and other anthropogenic signals, such as quarry or mine blasts, nuclear tests, machinery or transportation. The goal of earthquake detection is to detect seismic events with low and high signal-to-noise (SNR) in a computationally efficient manner. Due to the exponentially increasing volume of seismic data collected each year, earthquake detection in real time necessarily must be automatic.

Since the first attempt at automated earthquake detection 40 years ago with single trace detection on microprocessors¹³, numerous techniques have been developed. The short-term average over long-term average (STA/LTA) technique, where an event detection is triggered when the ratio of short-term average (few seconds) of seismic amplitude to long-term average (≈ 10 seconds) of seismic amplitude is above an arbitrary threshold, is the most widely-used automated algorithm

for earthquake detection¹¹. Template detection, which relies on the correlation of known seismic events to detect new seismic events, can detect earthquake signals at lower SNR levels than the STA/LTA method but is computationally unfeasible for real-time use⁹³. Fingerprint detection, which detects earthquakes using an unsupervised data mining approach, is faster than template detection²⁹⁶ but does not scale well for real-time use nor is it memory efficient.

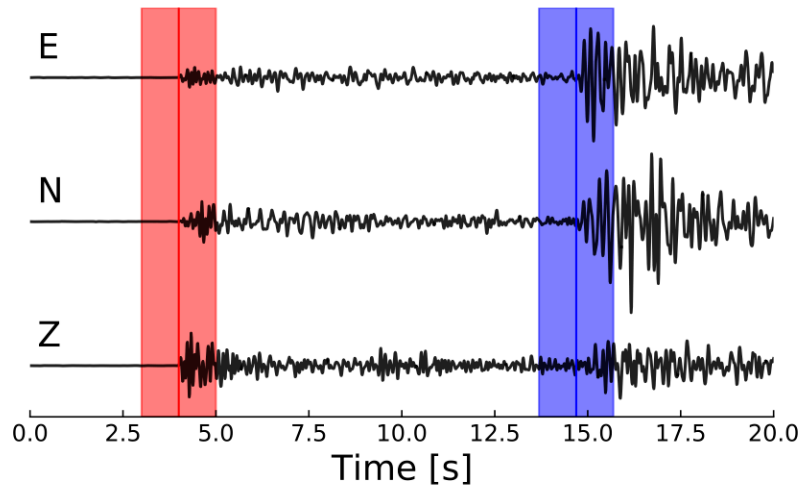


Figure 6.1: Velocity time-series of seismic event recorded on the East-West (E), North-South (N) and Vertical (Z) components of a three-component high-gain broadband seismometer. Red and blue bars and vertical lines denote the two-second window and exact arrival time of P-wave and S-wave, respectively.

1.1 EARTHQUAKE DETECTION USING MACHINE LEARNING

Machine learning is an increasingly popular approach to automate earthquake detection. Earthquake detection is one of the few applications in seismology to allow for supervised learning, where an algorithm learns a function to map known inputs to outputs²³. The first attempts at ML for earthquake detection used artificial neural networks on engineered features, such as spectrogram and STA/LTA¹²² or power spectral density¹⁵⁰ features to detect earthquake arrivals. More recently, the MyShake platform used a single-layer ANN with three features (interquartile range between the

25th and 75th percentile of the acceleration, the zero crossing rate, and the cumulative absolute velocity) to distinguish between earthquakes and human activity recorded on smart phones¹³³.

Convolutional neural networks (CNN) offer a huge leap forward in the power of ML-based earthquake detection. CNNs enable a shift away from earthquake detection based on amplitudes or similarity to previous recorded signals and toward generalized earthquake detection. Convolution-based detection is faster and more memory efficient than fingerprint or template detection and more accurate at lower SNR levels than STA/LTA methods^{191,214}. Convolutional and recurrent neural networks (RNN) are now used for P-wave polarity picking²¹⁵, earthquake magnitude estimation¹⁶⁷, and phase picking, which determines the exact arrival time of a particular seismic phase²¹⁵. To achieve higher accuracy, CNN earthquake detection networks tend to get deeper. The current state-of-the-art earthquake detection algorithms use deep residual network of convolutional and recurrent units¹⁶⁸ or transformer networks¹⁷⁰ to achieve 99% detection accuracy.

2 TINYML: MACHINE LEARNING AT THE EDGE

Tiny Machine Learning (TinyML) reverses the recent trends in ML to train and deploy larger and larger models on massive cloud servers. Instead, with TinyML, inference is performed on millions of compute and memory-constrained devices at the edge, where memory, compute and energy constraints are just as important as accuracy²⁸¹. We define the edge as any networked location between the point of data collection and cloud servers²³⁵. Micro-controllers used for TinyML have 100-1,000x less compute capability, memory availability or energy consumption than mobile phones, which leads to an emphasis on model and design efficiency⁶¹. Recent advances in deploying TinyML workflows to the edge include speech detection^{277,302,280,14,231}, arrhythmia detection⁸³, and computer vision tasks⁵¹.

Performing ML tasks on compute-constrained MCUs is very similar to previous on-site earth-

quake detection methods using neural networks³² or filter-banks¹⁵⁹, along with deep learning earthquake detection methods¹²². While earthquake detection has been deployed and run on mobile phones with success, in the form of the MyShake seismic network¹³³ and [Android phones](#), the widespread availability of always-on, always-connected IoT devices with ML capabilities could allow for a "true onsite approach, where a seismometer is installed at the user's location and provides a warning at that location, removes these telemetry delays, but with increased probability of false or missed alarms"¹². Fixed IoT devices additionally give greater guarantees for network stability over smart phones at the expense of limited RAM and compute capabilities. Here, we create a proof of concept for ML-based earthquake detection on an Arduino Nano 33 BLE board with 256 KB of SRAM and 1 MB flash storage.

3 DATA

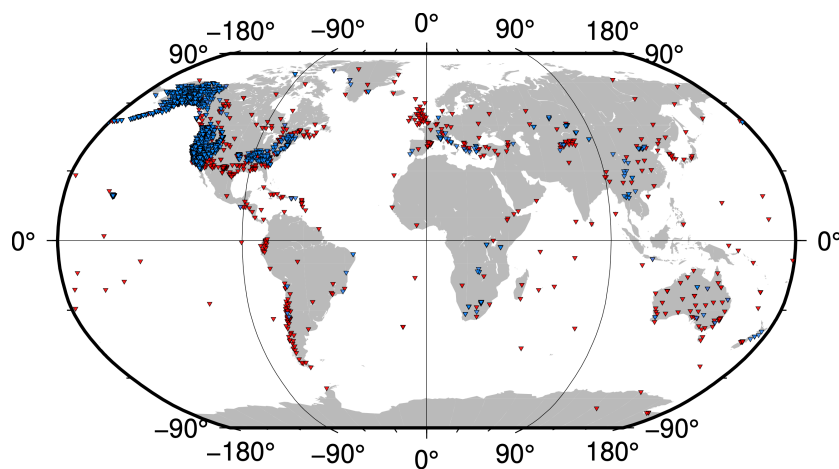


Figure 6.2: Locations of sensors used in this study. Blue triangles represent sensors providing earthquake and/or noise waveforms, whereas red triangles represent sensors providing only noise data.

To develop and test earthquake detection at the edge, we use a subset of the Stanford EArthquake Dataset (STEAD)¹⁶⁹ for training of deep neural networks (DNN). STEAD is a high-quality,

global dataset of containing labelled earthquake signals (P-wave and S-wave) and seismic noise. Data is stored as sixty-second three-channel (East-West, North-South and Vertical) traces sampled at 100 Hz recorded on a range of different seismometers. STEAD provides numerous metadata for each event and sensor, including event magnitude, event-sensor distance, P-wave and S-wave arrival times, event SNR, etc..

To construct our training dataset, we use earthquakes recorded on all station types (HH, BH, EH, etc.), with all event types (strike-slip, thrust and normal), and all event depths. We select earthquakes with source to station distances less than 100 km, which while likely larger than the maximum detection distance for MEMS accelerometer, as MyShake-enabled phones reliably detect earthquakes 10 km away¹³³, is reasonable for training. We reject all earthquakes where P and S waves arrive are within 0.2 seconds of each other. We select earthquakes with SNR > 40 dB and detection threshold above 0.5, as found by¹⁶⁹.

We split the training and test data by time using an 80/20 split. The training set is all noise, P-wave and S-wave events meeting the previous criteria before May 10, 2017, whereas the test set is similarly all noise, P-wave and S-wave events meeting the previous criteria after May, 10, 2017 in the STEAD dataset. The training and testing sets contain 250,980 and 62,748 two-second three-channel windows, respectively. We then balance the training and testing data such that they each contain 50% noise, 25% P waves and 25% S waves.

3.1 PROCESSING AND DATA AUGMENTATION

We apply very simple pre-processing workflow to mimic data as it is collected on-device with a MEMS accelerometer. For P-wave, S-wave and noise time waveforms alike, we read each three-channel 60 second window at 100Hz into a 60,000x3 tensor. Velocity waveforms are then converted to acceleration using a Fourier derivative. We then detrend, taper, and filter each channel between 2.0 and 20 Hz using a causal bandpass filter. We then select a two second window around each event

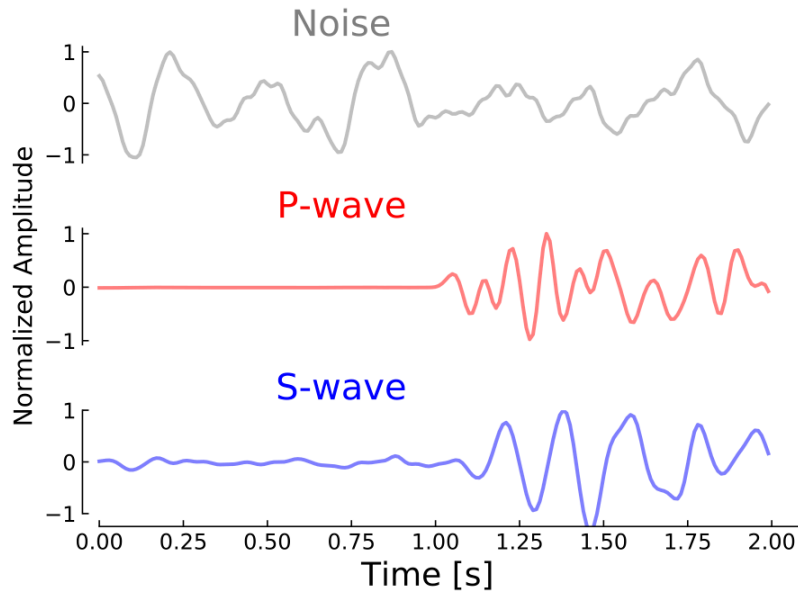


Figure 6.3: Representative two-second windows of the vertical component of noise, P-wave, and S-wave, respectively, after pre-processing.

pick (1 second before and 1 second after P-wave or S-wave arrival time), or randomly within the 60 second window for noise. Data are then resampled to 40 Hz, leaving an 80×3 tensor for each earthquake or noise.

Before training, we augment earthquake data by 1) shifting data randomly around the event pick up to 0.8 seconds 2) normalizing data to zero mean and unit standard deviation 3) randomly setting amplitude of 1 or 2 channels to zero with 0.3 probability 4) randomly scaling amplitude of channels with probability 0.3 and 5) adding Gaussian white noise. For noise data, we apply steps 2), 3) and 4) as before for P-wave and S-waves.

4 MODELS FOR TINY EARTHQUAKE DETECTION

We create a suite of simple and small model architectures, focusing on a selection of those mentioned in *Zhang et al.*³⁰², namely Deep Neural Networks (DNN), Convolutional Neural Networks

(CNN), and Depthwise-Separable Convolutional Neural Networks (DS-CNN), to test which network style is most accurate and efficient at detecting earthquakes. We did not test long short-term memory (LSTM) or gated recurrent unit (GRU)-based recurrent neural network (RNN) models as these types of networks are not available in the TensorFlow Lite Micro at the time of writing (December 2020). All models are created within the SRAM and Flash constraints of 256 KB and 1 MB, respectively, and a maximum desired model invoke time of 0.1 seconds, or 10 inferences per second. This latency threshold was chosen to minimize the time between earthquake detection and alert, which is vital for real-time EEW, while still allowing fairly complex models. The input to all three types of models is a $1 \times 80 \times 3 \times 1$ tensor, similar in shape to a tall and narrow RGB image.

All models were trained with the Keras framework with TensorFlow 2.3.0 as a backend on an Nvidia 1070Ti GPU¹. Models were trained with the Adam optimizer, categorical crossentropy as the loss function, and with the default learning rate of 0.001. Validation data for training was chosen as a random 20% split from the training set, which represents 16% of the total dataset. We train each model with a batch size of 256 for 100 epochs.

4.1 DEEP NEURAL NETWORK (DNN)

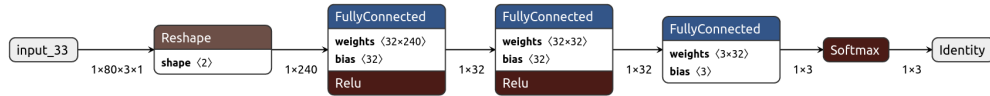


Figure 6.4: Model architecture for Deep Neural Network with two fully-connected layers of varying size and one output layer.

Deep neural networks (DNN) are a type of feedforward neural networks containing fully-connected layers. While able to generalize non-linear relationships, DNNs notably ignore the temporal relationship between inputs, which limits their effectiveness for time-series classification tasks³⁰². We flatten the input to each DNN model from a 4D tensor into a 240×1 array. We then run the 1D flattened through two fully-connected layers, each containing n neurons varying in size from 2^3 to 2^7 in

powers of two, followed by rectified linear unit (ReLU) activation functions. Each fully-connected layer is followed by a dropout layer with rate 0.1. The output layer is a fully-connected layer with three neurons with softmax activation, which gives the normalized probability, between 0 and 1, that the input is a P-wave, S-wave, or noise.

4.2 CONVOLUTIONAL NEURAL NETWORK (CNN)



Figure 6.5: Model architecture for Convolutional Neural Network with two convolutional layers of varying number of channels.

While DNN models ignore temporal relationship between inputs, Convolutional Neural Networks (CNN) use a series of small convolutional filters to extract local to high-level temporal patterns in time-series input³⁰². CNN models have far fewer trainable parameters than DNN at the expense of higher computational cost¹⁹¹. We create a suite of models with two 3x1 convolution layers and ReLU activation with number of channels varying from 2^1 to 2^8 in powers of two. Each convolutional layer is followed by a max pooling layer with pool size 3x1 and a dropout layer with rate = 0.1. The convolutional layers are followed by a fully-connected layer with 16 neurons and ReLU activation and finally with a three neuron fully-connected output layer with softmax activation.

4.3 DEPTHWISE-SEPARABLE CONVOLUTIONAL NEURAL NETWORK (DS-CNN)

Depthwise-Separable Convolution Neural Networks (DS-CNN), use spatial convolutions independently over each channel, followed by a pointwise (1x1) convolution in the depth dimension, rather than a standard 2D convolution. DS layers are often used to reduce the number of parameters and thus execution time of CNN models^{49,302}. Similar to our CCN suite of models, we create a suite of

DS-CNN models with two 3×1 convolution layers and ReLU activation with number of depth multipliers per DS layer varying from 2 to 8. All other layers and parameters are the same as the CNN models mentioned previously.



Figure 6.6: Model architecture for Depthwise-Separable Convolutional Neural Network with two DS-CNN layers of varying size.

5 RESULTS

Our most relevant criteria for evaluating models are accuracy and latency. We set a maximum latency of 10 ms as an acceptable threshold for all models, as was chosen by⁸⁴ for a speech enhancement application. For comparison, this is 20x less latency than the median alert latency of the MyShake network for $\approx M_5$ earthquakes²⁴⁶. We evaluate the accuracy of each model on the test set of earthquake and noise data of 62,748 windows recorded after May 10, 2017. The output of each model is a 3×1 tensor giving the probability that the input window is noise, P-wave or S-wave, respectively. We classify a model output as noise, P-wave or S-wave based on the maximum probability. We test for latency by measuring the invocation time (to within one millisecond) for each model on an Arduino Nano 33 BLE board with a single Cortex-M4 CPU running at 64 MHz using Arduino TensorFlowLite version 2.1.0-ALPHA.

As shown in Fig. 6.7, DNN models have the both the lowest latency and accuracy, whereas CNN models have the highest accuracy (92%) and latency (up to 2 seconds). DS-CNN models represent a mid-point between DNN and CNN, models in terms of accuracy and latency. For similar numbers of trainable parameters, CNN models have slightly higher accuracy than DS-CNN and far greater accuracy than DNN models, as shown in Fig. 6.8. The least accurate CNN and DS-CNN

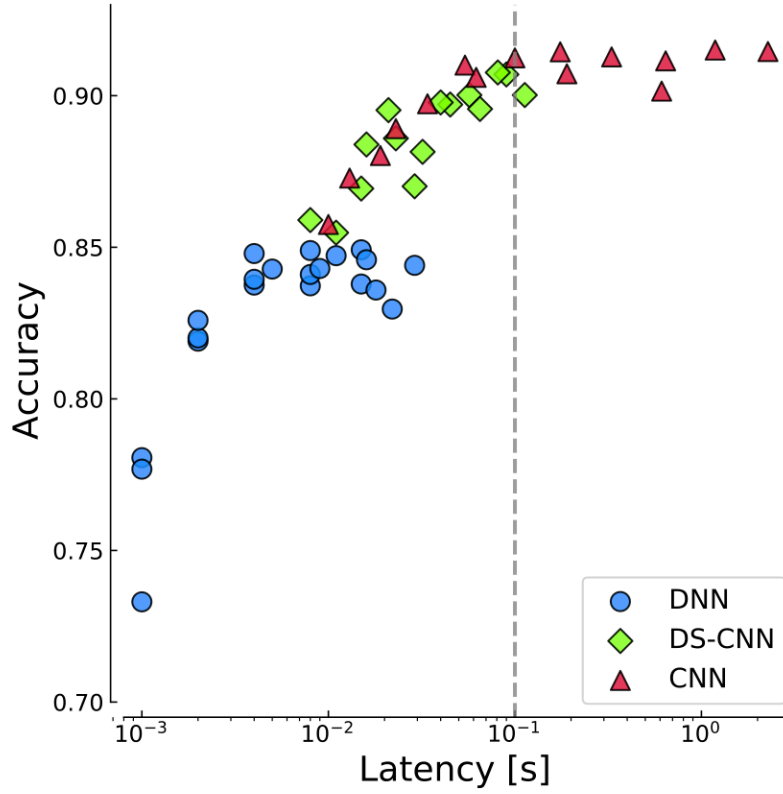


Figure 6.7: Accuracy vs. latency trade-off curve for suite of fully-connected (DNN), convolutional (CNN) and depthwise-separable convolutional (DS-CNN) models. Gray dashed line is cutoff for acceptable models, shown at latency of 0.1 seconds or 10 inferences per second.

models attain similar accuracy to the highest performing DNN model with 10-100x less trainable parameters, showing their efficient detection ability.

We test a range of minimum detection probability thresholds for P and S wave detection. Fig. 6.9 shows the precision-recall for a range of minimum probability thresholds between 0.4 and 0.99. We define precision and recall scores as

$$Precision = \frac{tp}{tp + fp}, \quad Recall = \frac{tp}{tp + fn} \quad (6.1)$$

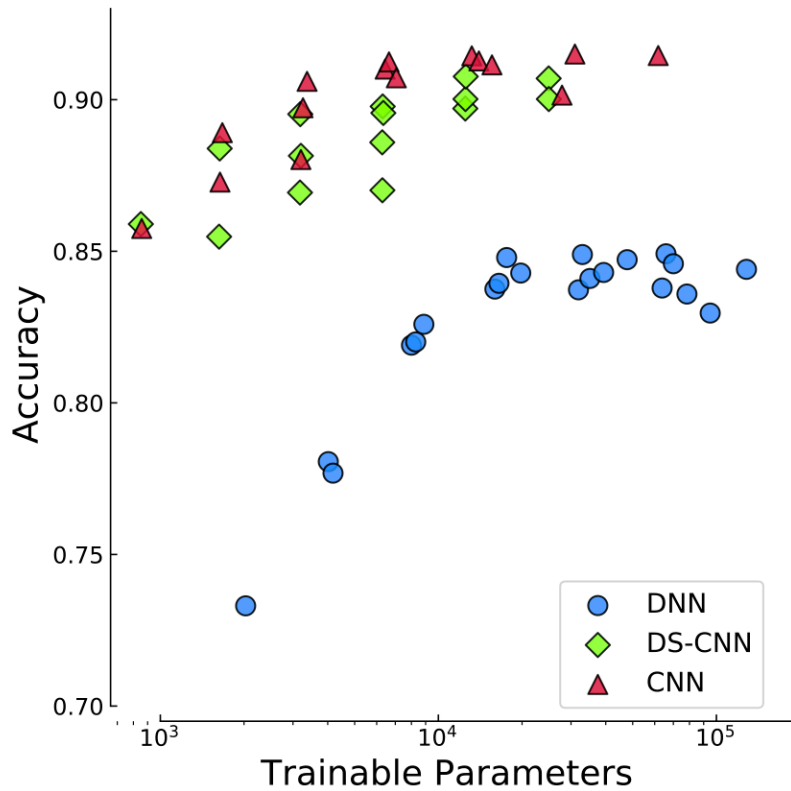


Figure 6.8: Accuracy vs. number of trainable parameters trade-off curve for suite of fully-connected (DNN), convolutional (CNN) and depthwise-separable convolutional (DS-CNN) models.

where tp is the number of true positives, fp is the number of false positives, and fn is the number of false negatives, for a given minimum probability threshold. We find that classification accuracy is maximized for a minimum P-wave and S-wave detection threshold of 0.49. Any P-wave or S-wave with model probability less than this threshold is incorrectly classified as noise.

6 DISCUSSION

Earthquake detection is the first step to most workflows in observational seismology. In that context, detection accuracy is vitally important to subsequent scientific analysis. Our best performing

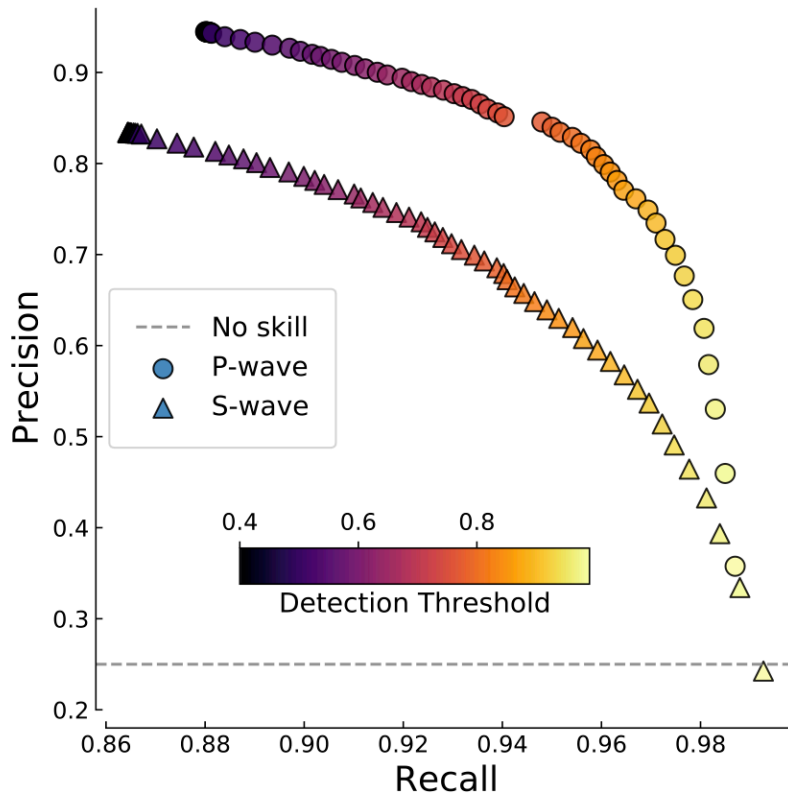


Figure 6.9: Precision vs recall trade-off curve for highest performing convolutional model on test data set. Colorbar denotes probability threshold above which test data is classified as P-wave or S-wave. Horizontal dashed line at precision = 0.25 is threshold for which network has no skill, as P-waves and S-waves both represent 25% of the test set.

model has a detection accuracy about 8% lower than vastly larger models trained to run large graphical processing units (GPU), which have accuracy around 99%^{214,170}. This is likely due to our simplified training set, as our precision-recall values are much lower than larger networks. As shown in Fig. 6.10, our most common misclassification errors are classifying P-waves as S-waves and S-waves as noise. We do a very good job of differentiating P-waves from noise, which is vital for real-time EEW¹⁶².

The difficulty of earthquake detection on the edge is the trade-off between accuracy and latency. As shown in Fig. 6.7, we can tune models to be either low-latency and high-accuracy based on archi-

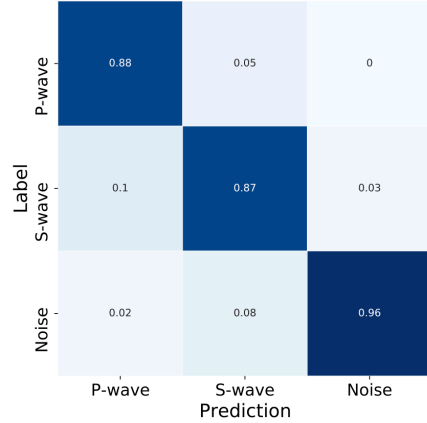


Figure 6.10: Normalized confusion matrix for highest performing convolutional model on test data set (N = 62, 748). Total number of P-wave, S-wave, and noise samples are 15,687, 15,687 and 31,374, respectively. Upper-right value has been rounded from 0.004 to 0. for display purposes.

ecture. If earthquake detection at the edge is used in an EEW context, the choice between latency and accuracy depends on the user’s tolerance for false alerts. Users who are false-alert-tolerant can choose to receive many false alerts but desire receive timely alerts for all levels of ground motion by choosing a low-latency CNN model^{162,163}. If the application is insensitive to latency, such as building an earthquake catalog or cataloging spatial variations in peak ground velocity, we can create larger and deeper models, which have longer inference times and hopefully greater detection capacity¹²⁴.

6.1 HOW TO USE SYSTEM

Details for replicating figures, preprocessing the data, training models and testing the results on device in this manuscript are available on [Github](#). Since the STanford EArthquake Dataset (STEAD) is quite large at around 90 GB, a smaller subset of test data is provided.

6.2 ETHICAL AI

While no human data was used in this dataset, seismometers can passively monitor the movement of humans at the scale of neighborhoods or cities¹⁴¹. In terms of bias, the location of earthquakes and sensors used in this study are highly skewed to those in the United States and Europe, which maintain most of the seismic networks globally but do not have the majority of earthquakes. Creating a global network of IoT seismometers could ameliorate this bias, as one of the limitations to installing seismic sensors is the significant cost of seismic sensors and telemetry equipment.

Calculating the location of an earthquake is only possible if each sensor in a seismic network meets two conditions: 1) the location of each sensor is precisely known and 2) all sensors are synced to UTC. These requirements almost always necessitate GPS for location and timing. The use of GPS combined with accelerometer data is a potential privacy concern for IoT devices installed in a home. If IoT sensor location data and accelerometer data are shared publicly, theoretically, one could know the location and movements, such as garage doors opening, doors shutting, etc., of a household. While again this is theoretically possible, the fixed nature of most IoT devices gives more privacy than current mobile phone-based seismic systems, which have the potential to track not only a person's physical location and movements throughout the day. In this way, moving earthquake detection from mobile phones to IoT devices should give end users greater autonomy and privacy.

7 CONCLUSIONS AND FUTURE DIRECTIONS

Earthquake detection is a broad starting point for many application from earthquake early warning to the study of fault zone structure. Here, we showed that earthquake detection is feasible on a memory and compute-constrained microcontroller within a latency threshold of 10 ms with 92% accuracy. For EEW applications, for which this application shows great promise, more work is required, particularly in the task of converting earthquake detection to ground motion or shaking

intensity. While Mousavi et al., 2020 have developed neural networks for estimating earthquake magnitudes from ground motion, predicting peak ground motion from small windows of ground velocity after P-wave arrival is not likely deterministic^{170,266}.

Directions for future work include creating deeper models based on the insights learned here. Inference latency could be further reduced by applying quantization to model layers and weights¹³⁵. Testing models that are either fully or partially include recurrent layers could improve accuracy, as state-of-the-art models for earthquake detection^{170,168} on large GPU-servers and speech detection on microcontrollers²³¹ include RNNs. Applying neural architecture search (NAS) could find earthquake detection architectures that are both high-accuracy and low-latency^{146,17}. There is much to do in the TinyML-earthquake detection space.

7

Conclusions

A painting is not a picture of an experience;
it is an experience.

Mark Rothko

In this dissertation, I have shown that measuring the change in seismic velocity, dv/v from the cross-correlation of ambient seismic waves recorded on a seismic network or with just a single seismic is akin to measuring groundwater levels using a single monitoring well. In addition, I have de-

veloped a framework for efficiently computing ambient noise cross-correlation functions on both the CPU and GPU using the computing language Julia and methods for processing seismic data on the cloud and at the edge. In chapter 2, I showed that dv/v tracked long-term trends in groundwater level due to drought and flood in the San Gabriel Valley, CA. Using dv/v I was able to accurately measure the loss in groundwater storage in the San Gabriel Valley during the most recent drought. In chapter 3, I developed a framework for accelerating cross-correlations using high-performance computing resources in the computing language Julia. With my Julia-language framework, I was able to handle computational workloads that previously required high-performance resources using only desktop resources. In chapter 4, I developed a methodology and theory for single-station groundwater monitoring. I showed that seismometers are sensitive to volumetric strains caused by changes in groundwater levels. I also suggest that groundwater levels could control near-surface changes in attenuation. In chapter 5, I applied the computational framework developed in chapter 3 and the methodology developed in chapter 4 to compute dv/v for every broadband seismometer in California over the last two decades. I showed that over the past two decades, seismic velocities in Southern California are increasing and are controlled by a reduction in precipitation from atmospheric rivers, which lead to long-lasting droughts. Finally, in chapter 6 I applied tiny machine learning to earthquake detection. I showed that earthquake detection is feasible using inexpensive, compute-constrained Internet-of-Things (IoT) devices, which opens the possibility for a future IoT seismic network with billions of devices.

References

- [1] Abadi, M., A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, R. Jozefowicz, Y. Jia, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, M. Schuster, R. Monga, S. Moore, D. Murray, C. Olah, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng (2015), TensorFlow: Large-scale machine learning on heterogeneous systems, *12th USENIX Symposium on Operating Systems Design and Implementation*, doi: 10.5281/zenodo.4724125.
- [2] Abdelkhalek, R., H. Calandra, O. Coulaud, J. Roman, and G. Latu (2009), Fast seismic modeling and reverse time migration on a GPU cluster, *Proceedings of the 2009 International Conference on High Performance Computing and Simulation, HPCS 2009*, pp. 36–43, doi: 10.1109/HPCSIM.2009.5192786.
- [3] Addair, T. G., D. A. Dodge, W. R. Walter, and S. D. Ruppert (2014), Large-scale seismic signal analysis with Hadoop, *Computers and Geosciences*, 66, 145–154, doi:10.1016/j.cageo.2014.01.014.

- [4] Aeschbach-Hertig, W., and T. Gleeson (2012), Regional strategies for the accelerating global problem of groundwater depletion, *Nature Geoscience*, 5(12), 853–861, doi:10.1038/ngeo1617.
- [5] Ajo-Franklin, J. B., S. Dou, N. J. Lindsey, I. Monga, C. Tracy, M. Robertson, V. Rodriguez Tribaldos, C. Ulrich, B. Freifeld, T. Daley, and X. Li (2019), Distributed Acoustic Sensing Using Dark Fiber for Near-Surface Characterization and Broadband Seismic Event Detection, *Scientific Reports*, 9(1), 1328, doi:10.1038/s41598-018-36675-8.
- [6] Aki, K. (1957), Space and time spectra of stationary stochastic waves, with special reference to microtremors, *Bulletin of the Earthquake Research Institute*, 35, 415–457, doi:http://hdl.handle.net/2261/11892.
- [7] Aki, K. (1969), Analysis of the seismic coda of local earthquakes as scattered waves, *Journal of Geophysical Research*, 74(2), 615–631, doi:10.1029/JB074i002p00615.
- [8] Aki, K. (1982), Scattering and attenuation, *Bulletin of the Seismological Society of America*, 72(6), 319–330, doi:10.1785/BSSA07206B0319.
- [9] Aki, K. (1996), Scale dependence in earthquake phenomena and its relevance to earthquake prediction., *Proceedings of the National Academy of Sciences*, 93(9), 3740–3747, doi:10.1073/pnas.93.9.3740.
- [10] Aki, K., and B. Chouet (1975), Origin of coda waves: Source, attenuation, and scattering effects, *Journal of Geophysical Research*, 80(23), 3322–3342, doi:10.1029/JB080i023p03322.
- [11] Allen, R. (1982), Automatic phase pickers: Their present use and future prospects, *Bulletin of the Seismological Society of America*, 72 (6B)(6B), 225–242.

- [12] Allen, R. M., P. Gasparini, O. Kamigaichi, and M. Bose (2009), The Status of Earthquake Early Warning around the World: An Introductory Overview, *Seismological Research Letters*, 80(5), 682–693, doi:10.1785/gssrl.80.5.682.
- [13] Allen, R. V. (1978), Automatic earthquake recognition and timing from single traces, *Bulletin of the Seismological Society of America*, 68(5), 1521–1532.
- [14] Alvarez, R., and H. J. Park (2019), End-to-end Streaming Keyword Spotting, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2019-May*, 6336–6340, doi:10.1109/ICASSP.2019.8683557.
- [15] Amos, C. B., P. Audet, W. C. Hammond, R. Bürgmann, I. A. Johanson, and G. Blewitt (2014), Uplift and seismicity driven by groundwater depletion in central California, *Nature*, 509(7501), 483–486, doi:10.1038/nature13275.
- [16] Arduin, F., L. Gualtieri, and E. Stutzmann (2015), How ocean waves rock the Earth: Two mechanisms explain microseisms with periods 3 to 300s, *Geophysical Research Letters*, 42(3), 765–772, doi:10.1002/2014GL062782.
- [17] Banbury, C., C. Zhou, I. Fedorov, R. M. Navarro, U. Thakker, D. Gope, V. J. Reddi, M. Mattina, and P. N. Whatmough (2020), MicroNets: Neural Network Architectures for Deploying TinyML Applications on Commodity Microcontrollers, *arXiv*, (Figure 1).
- [18] Bawden, G. W., W. Thatcher, R. S. Stein, K. W. Hudnut, and G. Peltzer (2001), Tectonic contraction across Los Angeles after removal of groundwater pumping effects, *Nature*, 412(6849), 812–815, doi:10.1038/35090558.
- [19] Beaucé, E., A. Romanenko, and W. B. Frank (2018), Fast matched filter (FMF): An efficient seismic matched-filter search for both CPU and GPU architectures, *Seismological Research Letters*, 89(1), 165–172, doi:10.1785/0220170181.

- [20] Ben-Zion, Y., and A. A. Allam (2013), Seasonal thermoelastic strain and postseismic effects in Parkfield borehole dilatometers, *Earth and Planetary Science Letters*, 379, 120–126, doi:10.1016/j.epsl.2013.08.024.
- [21] Bensen, G. D., M. H. Ritzwoller, M. P. Barmin, A. L. Levshin, F. Lin, M. P. Moschetti, N. M. Shapiro, and Y. Yang (2007), Processing seismic ambient noise data to obtain reliable broad-band surface wave dispersion measurements, *Geophysical Journal International*, 169(3), 1239–1260, doi:10.1111/j.1365-246X.2007.03374.x.
- [22] Benz, H. M., N. D. McMahon, R. C. Aster, D. E. McNamara, and D. B. Harris (2015), Hundreds of Earthquakes per Day : The 2014 Guthrie , Oklahoma , Earthquake Sequence, 86(5), 1–8, doi:10.1785/0220150019.
- [23] Bergen, K. J., P. A. Johnson, M. V. De Hoop, and G. C. Beroza (2019), Machine learning for data-driven discovery in solid Earth geoscience, *Science*, 363(6433), doi:10.1126/science.aau0323.
- [24] Beroza, G. C., A. T. Cole, and W. L. Ellsworth (1995), Stability of coda wave attenuation during the Loma Prieta, California, earthquake sequence, *Journal of Geophysical Research: Solid Earth*, 100(B3), 3977–3987, doi:10.1029/94JB02574.
- [25] Besard, T., C. Foket, and B. De Sutter (2019), Effective Extensible Programming: Unleashing Julia on GPUs, *IEEE Transactions on Parallel and Distributed Systems*, 30(4), 827–841, doi:10.1109/TPDS.2018.2872064.
- [26] Beyreuther, M., R. Barsch, L. Krischer, T. Megies, Y. Behr, and J. Wassermann (2010), ObsPy: A Python Toolbox for Seismology, *Seismological Research Letters*, 81(3), 530–533, doi:10.1785/gssrl.81.3.530.

- [27] Bezanson, J., A. Edelman, S. Karpinski, and V. B. Shah (2017), Julia: A Fresh Approach to Numerical Computing, *SIAM Review*, 59(1), 65–98, doi:10.1137/141000671.
- [28] Bezanson, J., J. Bolewski, and J. Chen (2018), Fast Flexible Function Dispatch in Julia.
- [29] Bianco, M. J., P. Gerstoft, K. B. Olsen, and F. C. Lin (2019), High-resolution seismic tomography of Long Beach, CA using machine learning, *Scientific Reports*, 9(1), 1–11, doi:10.1038/s41598-019-50381-z.
- [30] Boore, D. M., and W. B. Joyner (1997), Site amplifications for generic rock sites, *Bulletin of the Seismological Society of America*, 87(2), 327–341.
- [31] Boschelli, J., M. P. Moschetti, and C. Sens-Schönfelder (2021), Temporal Seismic Velocity Variations: Recovery Following From the 2019 M w 7.1 Ridgecrest, California Earthquake, *Journal of Geophysical Research: Solid Earth*, 126(4), 1–12, doi:10.1029/2020jb021465.
- [32] Böse, M., F. Wenzel, and M. Erdik (2008), PreSEIS: A neural network-based approach to earthquake early warning for finite faults, *Bulletin of the Seismological Society of America*, 98(1), 366–382, doi:10.1785/0120070002.
- [33] Brenguier, F., M. Campillo, C. Hadziioannou, N. M. Shapiro, R. M. Nadeau, and E. Larose (2008), Postseismic Relaxation Along the San Andreas Fault at Parkfield from Continuous Seismological Observations, *Science*, 321(5895), 1478–1481, doi:10.1126/science.1160943.
- [34] Brenguier, F., N. Shapiro, M. Campillo, V. Ferrazzini, Z. Duputel, O. Coutant, and A. Nercessian (2008), Towards forecasting volcanic eruptions using seismic noise, *Nature Geoscience*, 1(2), 126–130, doi:10.1038/ngeo104.

- [35] Brenguier, F., M. Campillo, T. Takeda, Y. Aoki, N. M. Shapiro, X. Briand, K. Emoto, and H. Miyake (2014), Mapping pressurized volcanic fluids from induced crustal seismic velocity drops, *Science*, 345(6192), 80–82, doi:10.1126/science.1254073.
- [36] Brenguier, F., D. Rivet, A. Obermann, N. Nakata, P. Boué, T. Lecocq, M. Campillo, and N. Shapiro (2016), 4-D noise-based seismology at volcanoes: Ongoing efforts and perspectives, *Journal of Volcanology and Geothermal Research*, 321, 182–195, doi:10.1016/j.jvolgeores.2016.04.036.
- [37] Brenguier, F., P. Kowalski, N. Ackerley, N. Nakata, P. Boué, M. Campillo, E. Larose, S. Rambaud, C. Pequegnat, T. Lecocq, P. Roux, V. Ferrazzini, N. Villeneuve, N. M. Shapiro, and J. Chaput (2016), Toward 4D Noise-Based Seismic Probing of Volcanoes: Perspectives from a Large- N Experiment on Piton de la Fournaise Volcano, *Seismological Research Letters*, 87(1), 15–25, doi:10.1785/0220150173.
- [38] Brodsky, E. E. (2003), A mechanism for sustained groundwater pressure changes induced by distant earthquakes, *Journal of Geophysical Research*, 108(B8), 1–10, doi:10.1029/2002JB002321.
- [39] Burbey, T. J. (2001), Stress-Strain Analyses for Aquifer-System Characterization, *Groundwater*, 39(1), 128–136, doi:10.1111/j.1745-6584.2001.tb00358.x.
- [40] Buwalda, J. (1954), Geology of the Tehachapi Mountains, California, *Tech. Rep. 170*.
- [41] California Department of Water Resources (1966), Planned utilization of ground water basins: San Gabriel Valley, *Tech. rep.*, Sacramento, California.
- [42] California Department of Water Resources (2015), California’s Most Significant Drought: Comparing Historical and Recent Conditions, *Tech. rep.*, Sacramento, California.

- [43] Campillo, M., and A. Paul (2003), Long-Range Correlations in the Diffuse Seismic Coda, *Science*, 299(5606), 547–549, doi:10.1126/science.1078551.
- [44] Castac Basin GSA (2020), Groundwater Sustainability Plan Castac Lake Valley, *Tech. rep.*
- [45] Castellanos, J. C., R. W. Clayton, and A. Juarez (2020), Using a Time-Based Subarray Method to Extract and Invert Noise-Derived Body Waves at Long Beach, California, *Journal of Geophysical Research: Solid Earth*, 125(5), doi:10.1029/2019JB018855.
- [46] Celso, G. Thompson, M. West, usfseismiclab, D. Ketner, and C. Tape (2018), geoscience-community-codes/GISMO: version 1.20 beta, doi:10.5281/ZENODO.1404723.
- [47] Chaussard, E., P. Milillo, R. Bürgmann, D. Perissin, E. J. Fielding, and B. Baker (2017), Remote Sensing of Ground Deformation for Monitoring Groundwater Management Practices: Application to the Santa Clara Valley During the 2012–2015 California Drought, *Journal of Geophysical Research: Solid Earth*, pp. 8566–8582, doi:10.1002/2017JB014676.
- [48] Chen, P., N. J. Taylor, K. G. Dueker, I. S. Keifer, A. K. Wilson, C. L. McGuffey, C. G. Novitsky, A. J. Spears, and W. S. Holbrook (2016), PSIN: A scalable, Parallel algorithm for Seismic INterferometry of large-N ambient-noise data, *Computers and Geosciences*, 93, 88–95, doi:10.1016/j.cageo.2016.05.003.
- [49] Chollet, F. (2016), Xception: Deep Learning with Depthwise Separable Convolutions, *arXiv*.
- [50] Chouet, B. (1979), Temporal variation in the attenuation of earthquake coda near Stone Canyon, California, *Geophysical Research Letters*, 6(3), 143–146, doi:10.1029/GL006i003p00143.

- [51] Chowdhery, A., P. Warden, J. Shlens, A. Howard, and R. Rhodes (2019), Visual Wake Words Dataset.
- [52] Christensen, N. I. (1996), Poisson's ratio and crustal seismology, *Journal of Geophysical Research: Solid Earth*, 101(B2), 3139–3156, doi:10.1029/95JB03446.
- [53] Christensen, N. I., and H. F. Wang (1985), Influence of pore pressure and confining pressure on dynamic elastic properties of Berea sandstone, doi:http://dx.doi.org/10.1016/0148-9062(86)90438-9.
- [54] Claerbout, J., S. Cole, D. Nichols, and L. Zhang (1988), Why a big 2-D array to record microseisms?, *SEP-59*.
- [55] Claerbout, J. F. (1968), Synthesis of a Layered Medium From its Acoustic Transmission Response, *GEOPHYSICS*, 33(2), 264–269, doi:10.1190/1.1439927.
- [56] Clarke, D., L. Zaccarelli, N. M. Shapiro, and F. Brenguier (2011), Assessment of resolution and accuracy of the Moving Window Cross Spectral technique for monitoring crustal temporal variations using ambient seismic noise, *Geophysical Journal International*, 186(2), 867–882, doi:10.1111/j.1365-246X.2011.05074.x.
- [57] Clayton, R. W. (2020), A detailed image of the continent-borderland transition beneath Long Beach, California, *Geophysical Journal International*, 222(3), 2102–2107, doi:10.1093/gji/ggaa286.
- [58] Clements, T., and M. A. Denolle (2018), Tracking Groundwater Levels Using the Ambient Seismic Field, *Geophysical Research Letters*, 45(13), 6459–6465, doi:10.1029/2018GL077706.

- [59] Clements, T., and M. A. Denolle (2020), SeisNoise.jl: Ambient Seismic Noise Cross Correlation on the CPU and GPU in Julia, *Seismological Research Letters*, doi:10.1785/0220200192.
- [60] Daly, C., R. P. Neilson, and D. L. Phillips (1994), A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain, *Journal of Applied Meteorology*, 33(2), 140–158, doi:10.1175/1520-0450(1994)033<0140:ASTMFM>2.0.CO;2.
- [61] David, R., J. Duke, A. Jain, V. J. Reddi, N. Jeffries, J. Li, N. Kreeger, I. Nappier, M. Natraj, S. Regev, R. Rhodes, T. Wang, and P. Warden (2020), TensorFlow Lite Micro: Embedded Machine Learning on TinyML Systems.
- [62] De Fazio, T. L., K. Aki, and J. Alba (1973), Solid Earth tide and observed change in the in situ seismic velocity, *Journal of Geophysical Research*, 78(8), 1319–1322, doi:10.1029/JB078i008p01319.
- [63] De Plaen, R. S., T. Lecocq, C. Caudron, V. Ferrazzini, and O. Francis (2016), Single-station monitoring of volcanoes using seismic ambient noise, *Geophysical Research Letters*, 43(16), 8511–8518, doi:10.1002/2016GL070078.
- [64] Dean, J., and S. Ghemawat (2008), MapReduce: Simplified data processing on large clusters, *Communications of the ACM*, 51(1), 107, doi:10.1145/1327452.1327492.
- [65] Delph, J. R., A. Levander, and F. Niu (2019), Constraining Crustal Properties Using Receiver Functions and the Autocorrelation of Earthquake-Generated Body Waves, *Journal of Geophysical Research: Solid Earth*, 124(8), 8981–8997, doi:10.1029/2019JB017929.
- [66] Denolle, M. A., E. M. Dunham, G. A. Prieto, and G. C. Beroza (2014), Strong Ground Motion Prediction Using Virtual Earthquakes, *Science*, 343(6169), 399–403, doi:10.1126/science.1245678.

- [67] Derode, A., E. Larose, M. Campillo, and M. Fink (2003), How to estimate the Green's function of a heterogeneous medium between two passive sensors? Application to acoustic waves, *Applied Physics Letters*, 83(15), 3054–3056, doi:10.1063/1.1617373.
- [68] Derode, A., E. Larose, M. Tanter, J. de Rosny, A. Tourin, M. Campillo, and M. Fink (2003), Recovering the Green's function from field-field correlations in an open scattering medium (L), *The Journal of the Acoustical Society of America*, 113(6), 2973, doi:10.1121/1.1570436.
- [69] Dettinger, M. (2016), Historical and Future Relations Between Large Storms and Droughts in California, *San Francisco Estuary and Watershed Science*, 14(2), doi:10.15447/sfews.2016v14iss2art1.
- [70] Dettinger, M. D., F. M. Ralph, T. Das, P. J. Neiman, and D. R. Cayan (2011), Atmospheric Rivers, Floods and the Water Resources of California, *Water*, 3(4), 445–478, doi:10.3390/w3020445.
- [71] Díaz, J., M. Ruiz, P. S. Sánchez-Pastor, and P. Romero (2017), Urban Seismology: On the origin of earth vibrations within a city, *Scientific Reports*, 7(1), 1–11, doi:10.1038/s41598-017-15499-y.
- [72] Diaz, J., M. Schimmel, M. Ruiz, and R. Carbonell (2020), Seismometers Within Cities: A Tool to Connect Earth Sciences and Society, *Frontiers in Earth Science*, 8(February), 1–7, doi:10.3389/feart.2020.00009.
- [73] Dodge, D. A., and G. C. Beroza (1997), Source array analysis of coda waves near the 1989 Loma Prieta, California, mainshock: Implications for the mechanism of coseismic velocity changes, *Journal of Geophysical Research-Solid Earth*, 102(B11), 24,437–24,458, doi:10.1029/97JB02024.

- [74] Döll, P., H. Hoffmann-Dobrev, F. Portmann, S. Siebert, A. Eicker, M. Rodell, G. Strassberg, and B. Scanlon (2012), Impact of water withdrawals from groundwater and surface water on continental water storage variations, *Journal of Geodynamics*, 59-60, 143-156, doi:10.1016/j.jog.2011.05.001.
- [75] Donaldson, C., C. Caudron, R. G. Green, W. A. Thelen, and R. S. White (2017), Relative seismic velocity variations correlate with deformation at Kīlauea volcano, *Science Advances*, 3(6), e1700219, doi:10.1126/sciadv.1700219.
- [76] Dong, C., G. MacDonald, G. S. Okin, and T. W. Gillespie (2019), Quantifying Drought Sensitivity of Mediterranean Climate Vegetation to Recent Warming: A Case Study in Southern California, *Remote Sensing*, 11(24), 2902, doi:10.3390/rs11242902.
- [77] Dou, S., N. Lindsey, A. M. Wagner, T. M. Daley, B. Freifeld, M. Robertson, J. Peterson, C. Ulrich, E. R. Martin, and J. B. Ajo-Franklin (2017), Distributed Acoustic Sensing for Seismic Monitoring of the Near Surface: A Traffic-Noise Interferometry Case Study, *Scientific Reports*, 7(1), 1-12, doi:10.1038/s41598-017-11986-4.
- [78] Dougherty, S. L., E. S. Cochran, and R. M. Harrington (2019), The LARge-n Seismic Survey in Oklahoma (LASSO) experiment, *Seismological Research Letters*, 90(5), 2051-2057, doi:10.1785/0220190094.
- [79] Durbin, T. J. (2010), Natural Groundwater Recharge Antelope Valley, California, *Tech. rep.*
- [80] Dziewonski, A. M., and D. L. Anderson (1981), Preliminary reference Earth model*, *Phys. Earth Planet. Inter.*, 9, 297-356.
- [81] Eberhart-Phillips, D., D.-h. Han, and M. D. Zoback (1989), Empirical relationships among seismic velocity, effective pressure, porosity, and clay content in sandstone, *GEOPHYSICS*, 54(1), 82-89, doi:10.1190/1.1442580.

- [82] Egle, D. M., and D. E. Bray (1976), Measurement of acoustoelastic and third-order elastic constants for rail steel, *The Journal of the Acoustical Society of America*, 60(3), 741–744, doi: 10.1121/1.381146.
- [83] Faraone, A., and R. Delgado-Gonzalo (2020), Convolutional-Recurrent Neural Networks on Low-Power Wearable Platforms for Cardiac Arrhythmia Detection, *Proceedings - 2020 IEEE International Conference on Artificial Intelligence Circuits and Systems, AICAS 2020*, pp. 153–157, doi:10.1109/AICAS48895.2020.9073950.
- [84] Fedorov, I., M. Stamenovic, C. Jensen, L.-C. Yang, A. Mandell, Y. Gan, M. Mattina, and P. N. Whatmough (2020), TinyLSTMs: Efficient Neural Speech Enhancement for Hearing Aids.
- [85] Fehler, M., P. Roberts, and T. Fairbanks (1988), A temporal change in coda wave attenuation observed during an eruption of Mount St. Helens, *Journal of Geophysical Research: Solid Earth*, 93(B5), 4367–4373, doi:10.1029/JB093iB05p04367.
- [86] Fichtner, A., L. Ermert, and A. Gokhberg (2017), Seismic Noise Correlation on Heterogeneous Supercomputers, *Seismological Research Letters*, 88(4), 1141–1145, doi: 10.1785/0220170043.
- [87] Fichtner, A., D. Bowden, and L. Ermert (2020), Optimal processing for seismic noise correlations, *EarthArXiv*.
- [88] Fitts, C. R. (2013), Groundwater, in *Groundwater Science*, pp. 1–22, Elsevier, doi:10.1016/B978-0-12-384705-8.00001-7.
- [89] Galloway, D., D. R. Jones, and S. Ingebritsen (1999), *Land Subsidence in the United States*.

- [90] Galloway, D. L., and T. J. Burbey (2011), Review: Regional land subsidence accompanying groundwater extraction, *Hydrogeology Journal*, 19(8), 1459–1486, doi:10.1007/s10040-011-0775-5.
- [91] Galloway, D. L., and J. Hoffmann (2007), The application of satellite differential SAR interferometry-derived ground displacements in hydrogeology, *Hydrogeology Journal*, 15(1), 133–154, doi:10.1007/s10040-006-0121-5.
- [92] Garven, G. (1995), Continental-Scale Groundwater Flow and Geologic Processes, *Annual Review of Earth and Planetary Sciences*, 23(1), 89–117, doi:10.1146/annurev.ea.23.050195.000513.
- [93] Gibbons, S. J., and F. Ringdal (2006), The detection of low magnitude seismic events using array-based waveform correlation, *Geophysical Journal International*, 165(1), 149–166, doi:10.1111/j.1365-246X.2006.02865.x.
- [94] Goldstein, P., D. Dodge, M. Firpo, and L. Minner (2003), SAC2000: Signal Processing and Analysis Tools for Seismologists and Engineers, *Invited contribution to "The IASPEI International Handbook of Earthquake and Engineering Seismology*, (Academic Press, London), 1–7.
- [95] Got, J. L., G. Poupinet, and J. Frechet (1990), Changes in source and site effects compared to codaQ-1 temporal variations using microearthquakes doublets in California, *Pure and Applied Geophysics PAGEOPH*, 134(2), 195–228, doi:10.1007/BF00876998.
- [96] Goutorbe, B., D. L. de Oliveira Coelho, and S. Drouet (2015), Rayleigh wave group velocities at periods of 6–23 s across Brazil from ambient noise tomography, *Geophysical Journal International*, 203(2), 869–882, doi:10.1093/gji/ggv343.

- [97] Grêt, A., R. Snieder, and J. Scales (2006), Time-lapse monitoring of rock properties with coda wave interferometry, *Journal of Geophysical Research: Solid Earth*, 111(3), 1–11, doi:10.1029/2004JB003354.
- [98] Hadziioannou, C., E. Larose, A. Baig, P. Roux, and M. Campillo (2011), Improving temporal resolution in ambient noise monitoring of seismic wave speed, *Journal of Geophysical Research: Solid Earth*, 116(7), 1–10, doi:10.1029/2011JB008200.
- [99] Hart, D. J., and H. F. Wang (1995), Laboratory measurements of a complete set of poroelastic moduli for Berea sandstone and Indiana limestone, *Journal of Geophysical Research: Solid Earth*, 100(B9), 17,741–17,751, doi:10.1029/95JB01242.
- [100] Healy, R. W., and P. G. Cook (2002), Using groundwater levels to estimate recharge, *Hydrogeology Journal*, 10(1), 91–109, doi:10.1007/s10040-001-0178-0.
- [101] Herlihy, M. (2006), The art of multiprocessor programming, in *Proceedings of the twenty-fifth annual ACM symposium on Principles of distributed computing - PODC '06*, vol. 57, pp. 1–2, ACM Press, New York, New York, USA, doi:10.1145/1146381.1146382.
- [102] Herrmann, R. B. (2013), Computer Programs in Seismology: An Evolving Tool for Instruction and Research, *Seismological Research Letters*, 84(6), 1081–1088, doi:10.1785/0220110096.
- [103] Hillers, G., and Y. Ben-Zion (2011), Seasonal variations of observed noise amplitudes at 2–18 Hz in southern California, *Geophysical Journal International*, 184(2), 860–868, doi:10.1111/j.1365-246X.2010.04886.x.
- [104] Hillers, G., N. Graham, M. Campillo, S. Kedar, M. Landés, and N. Shapiro (2012), Global oceanic microseism sources as seen by seismic arrays and predicted by wave action models, *Geochemistry, Geophysics, Geosystems*, 13(1), doi:10.1029/2011GC003875.

- [105] Hillers, G., L. Retailleau, M. Campillo, A. Inbal, J. P. Ampuero, and T. Nishimura (2015), In situ observations of velocity changes in response to tidal deformation from analysis of the high-frequency ambient wavefield, *Journal of Geophysical Research: Solid Earth*, 120(1), 210–225, doi:10.1002/2014JB011318.
- [106] Hillers, G., Y. Ben-Zion, M. Campillo, and D. Zigone (2015), Seasonal variations of seismic velocities in the San Jacinto fault area observed with ambient seismic noise, *Geophysical Journal International*, 202(2), 920–932, doi:10.1093/gji/ggv151.
- [107] Hirose, T., H. Nakahara, and T. Nishimura (2019), A Passive Estimation Method of Scattering and Intrinsic Absorption Parameters From Envelopes of Seismic Ambient Noise Cross-Correlation Functions, *Geophysical Research Letters*, 46(7), 3634–3642, doi: 10.1029/2018GL080553.
- [108] Hobiger, M., U. Wegler, K. Shiomi, and H. Nakahara (2014), Single-station cross-correlation analysis of ambient seismic noise: Application to stations in the surroundings of the 2008 Iwate-Miyagi Nairiku earthquake, *Geophysical Journal International*, 198(1), 90–109, doi: 10.1093/gji/ggu115.
- [109] Hotovec-Ellis, A., J. Gomberg, J. Vidale, and K. Creager (2014), A continuous record of intereruption velocity change at Mount St. Helens from coda wave interferometry, *Journal of Geophysical Research: Solid Earth*, 119, 6093–6112, doi:10.1002/2014JB010998. Received.
- [110] Hsu, Y.-J., H. Kao, R. Bürgmann, Y.-T. Lee, H.-H. Huang, Y.-F. Hsu, Y.-M. Wu, and J. Zhuang (2021), Synchronized and asynchronous modulation of seismicity by hydrological loading: A case study in Taiwan, *Science Advances*, 7(16), eabf7282, doi:10.1126/sciadv.abf7282.

- [111] Hughes, D. S., and J. L. Kelly (1953), Second-Order Elastic Deformation of Solids, *Physical Review*, 92(5), 1145–1149, doi:10.1103/PhysRev.92.1145.
- [112] Hunter, J. D. (2007), Matplotlib: A 2D Graphics Environment, *Computing in Science & Engineering*, 9(3), 90–95, doi:10.1109/MCSE.2007.55.
- [113] Hutton, K., J. Woessner, and E. Hauksson (2010), Earthquake Monitoring in Southern California for Seventy-Seven Years (1932-2008), *Bulletin of the Seismological Society of America*, 100(2), 423–446, doi:10.1785/0120090130.
- [114] Illien, L., C. Andermann, C. Sens-Schönfelder, K. L. Cook, K. P. Baidya, L. B. Adhikari, and N. Hovius (2021), Subsurface Moisture Regulates Himalayan Groundwater Storage and Discharge, *AGU Advances*, 2(2), doi:10.1029/2021av000398.
- [115] James, S. R. (2019), Insights Into Permafrost and Seasonal Active - Layer Dynamics From Ambient Seismic Noise Monitoring Journal of Geophysical Research : Earth Surface, pp. 1798–1816, doi:10.1029/2019JF005051.
- [116] James, S. R., H. A. Knox, R. E. Abbott, and E. J. Screaton (2017), Improved moving window cross-spectral analysis for resolving large temporal seismic velocity changes in permafrost, *Geophysical Research Letters*, 44(9), 4018–4026, doi:10.1002/2016GL072468.
- [117] Jasechko, S., S. J. Birks, T. Gleeson, Y. Wada, P. J. Fawcett, Z. D. Sharp, J. J. McDonnell, and J. M. Welker (2014), The pronounced seasonality of global groundwater recharge, *Water Resources Research*, 50(11), 8845–8867, doi:10.1002/2014WR015809.
- [118] Ji, K. H., and T. A. Herring (2012), Correlation between changes in groundwater levels and surface deformation from GPS measurements in the San Gabriel Valley, California, *Geophysical Research Letters*, 39(1), 1–5, doi:10.1029/2011GL050195.

- [119] Ji, S., L. Li, H. B. Motra, F. Wuttke, S. Sun, K. Michibayashi, and M. H. Salisbury (2018), Poisson's Ratio and Auxetic Properties of Natural Rocks, *Journal of Geophysical Research: Solid Earth*, 123(2), 1161–1185, doi:10.1002/2017JB014606.
- [120] Jiang, C., and M. A. Denolle (2020), NoisePy: A New High-Performance Python Tool for Ambient-Noise Seismology, *Seismological Research Letters*, doi:10.1785/0220190364.
- [121] Jin, A., and K. Aki (1986), Temporal change in coda Q before the Tangshan Earthquake of 1976 and the Haicheng Earthquake of 1975, *Journal of Geophysical Research*, 91(B1), 665, doi:10.1029/JB091iB01p00665.
- [122] Jin Wang, and Ta-Liang Teng (1995), Artificial neural network-based seismic detector, *Bulletin of the Seismological Society of America*, 85(1), 308–319.
- [123] Johnson, C. W., Y. Fu, and R. Bürgmann (2017), Seasonal water storage, stress modulation, and California seismicity, *Science*, 356(6343), 1161–1164, doi:10.1126/science.aak9547.
- [124] Johnson, C. W., D. Kilb, A. Baltay, and F. Vernon (2020), Peak Ground Velocity Spatial Variability Revealed by Dense Seismic Array in Southern California, *Journal of Geophysical Research: Solid Earth*, 125(6), 1–17, doi:10.1029/2019JB019157.
- [125] Johnson, P. A., and P. N. J. Rasolofosaon (1996), Nonlinear elasticity and stress-induced anisotropy in rock, *Journal of Geophysical Research: Solid Earth*, 101(B2), 3113–3124, doi:10.1029/95JB02880.
- [126] Johnston, H., M. N. Tokst, and A. Timur (1979), Attenuation of seismic waves in dry and saturated rocks : Mechanisms, II, 31(4).

- [127] Jones, J. P., K. Okubo, T. Clements, and M. A. Denolle (2020), SeisIO: A Fast, Efficient Geophysical Data Architecture for the Julia Language, *Seismological Research Letters*, doi:10.1785/0220190295.
- [128] Karplus, M., and B. Schmandt (2018), Preface to the Focus Section on Geophone Array Seismology, *Seismological Research Letters*, 89(5), 1597–1600, doi:10.1785/0220180212.
- [129] Kim, D., and V. Lekic (2019), Groundwater Variations From Autocorrelation and Receiver Functions Geophysical Research Letters, pp. 1–8, doi:10.1029/2019GL084719.
- [130] King, N. E., D. Argus, J. Langbein, D. C. Agnew, G. Bawden, R. S. Dollar, Z. Liu, D. Galloway, E. Reichard, A. Yong, F. H. Webb, Y. Bock, K. Stark, and D. Barseghian (2007), Space geodetic observation of expansion of the San Gabriel Valley, California, aquifer system, during heavy rainfall in winter 2004-2005, *Journal of Geophysical Research: Solid Earth*, 112(3), 1–11, doi:10.1029/2006JB004448.
- [131] Kino, G. S., J. B. Hunter, G. C. Johnson, A. R. Selfridge, D. M. Barnett, G. Hermann, and C. R. Steele (1979), Acoustoelastic imaging of stress fields, *Journal of Applied Physics*, 50(4), 2607–2613, doi:10.1063/1.326268.
- [132] Komatitsch, D., D. Göldeke, G. Erlebacher, and D. Michéa (2010), Modeling the propagation of elastic waves using spectral elements on a cluster of 192 GPUs, *Computer Science - Research and Development*, 25(1-2), 75–82, doi:10.1007/s00450-010-0109-1.
- [133] Kong, Q., R. M. Allen, L. Schreier, and Y. W. Kwon (2016), Earth Sciences: MyShake: A smartphone seismic network for earthquake early warning and beyond, *Science Advances*, 2(2), 1–9, doi:10.1126/sciadv.1501055.
- [134] Konikow, L. F., and E. Kendy (2005), Groundwater depletion: A global problem, *Hydrogeology Journal*, 13(1), 317–320, doi:10.1007/s10040-004-0411-8.

- [135] Krishnamoorthi, R. (2018), Quantizing deep convolutional networks for efficient inference: A whitepaper.
- [136] Krizhevsky, A., I. Sutskever, and G. E. Hinton (2012), ImageNet Classification with Deep Convolutional Neural Networks, in *Advances in Neural Information Processing Systems* 25, edited by F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, pp. 1097–1105, Curran Associates, Inc.
- [137] Kumm, M., H. de Moel, P. J. Ward, and O. Varis (2011), How close do we live to water? a global analysis of population distance to freshwater bodies, *PLoS ONE*, 6(6), doi:10.1371/journal.pone.0020578.
- [138] Larose, E., and S. Hall (2009), Monitoring stress related velocity variation in concrete with a 2×10^{-5} relative resolution using diffuse ultrasound, *The Journal of the Acoustical Society of America*, 125(4), 1853–1856, doi:10.1121/1.3079771.
- [139] Lecocq, T., C. Caudron, and F. Brenguier (2014), MSNoise, a Python Package for Monitoring Seismic Velocity Changes Using Ambient Seismic Noise, *Seismological Research Letters*, 85(3), 715–726, doi:10.1785/0220130073.
- [140] Lecocq, T., L. Longuevergne, H. A. Pedersen, F. Brenguier, and K. Stammer (2017), Monitoring ground water storage at mesoscale using seismic noise: 30 years of continuous observation and thermo-elastic and hydrological modeling, *Scientific Reports*, 7(1), 1–16, doi: 10.1038/s41598-017-14468-9.
- [141] Lecocq, T., S. P. Hicks, K. V. Noten, K. V. Wijk, and P. Koelemeijer (2020), Global quieting of high-frequency seismic noise due to COVID-19 pandemic lockdown measures, 1343(September), 1338–1343.

- [142] Lees, J. M. (2019), RSEIS: Seismic Time Series Analysis Tools, *available at: <https://CRAN.R-project.org/package=RSEIS>*.
- [143] Li, Z., Z. Peng, D. Hollis, L. Zhu, and J. McClellan (2018), High-resolution seismic event detection using local similarity for Large-N arrays, *Scientific Reports*, 8(1), 1–10, doi:10.1038/s41598-018-19728-w.
- [144] Lin, F.-C., M. P. Moschetti, and M. H. Ritzwoller (2008), Surface wave tomography of the western United States from ambient seismic noise: Rayleigh and Love wave phase velocity maps, *Geophysical Journal International*, 173(1), 281–298, doi:10.1111/j.1365-246X.2008.03720.x.
- [145] Lin, F. C., D. Li, R. W. Clayton, and D. Hollis (2013), High-resolution 3D shallow crustal structure in Long Beach, California: Application of ambient noise tomography on a dense seismic array, *Geophysics*, 78(4), doi:10.1190/geo2012-0453.1.
- [146] Lin, J., W.-M. Chen, Y. Lin, J. Cohn, C. Gan, and S. Han (2020), MCUNet: Tiny Deep Learning on IoT Devices, pp. 1–13.
- [147] Liu, C., K. Aslam, and E. Daub (2020), Seismic Velocity Changes Caused by Water Table Fluctuation in the New Madrid Seismic Zone and Mississippi Embayment, *Journal of Geophysical Research: Solid Earth*, 125(8), 1–13, doi:10.1029/2020JB019524.
- [148] Lohman, R. B., and J. J. McGuire (2007), Earthquake swarms driven by aseismic creep in the Salton Trough, California, *Journal of Geophysical Research: Solid Earth*, 112(4), 1–10, doi:10.1029/2006JB004596.
- [149] Maccarthy, J., O. Marcillo, and C. Trabant (2020), Seismology in the Cloud : A New Streaming Workflow, doi:10.1785/0220190357.Introduction.

- [150] Madureira, G., and A. E. Ruano (2009), A Neural Network Seismic Detector, *IFAC Proceedings Volumes*, 42(19), 304–309, doi:10.3182/20090921-3-TR-3005.00054.
- [151] Main San Gabriel Watermaster (2017), 2016 - 2017 Annual Report, *Tech. rep.*, Azusa, California.
- [152] Malagnini, L., and T. Parsons (2020), Seismic Attenuation Monitoring of a Critically Stressed San Andreas Fault, *Geophysical Research Letters*, 47(23), 1–11, doi:10.1029/2020GL089201.
- [153] Mao, S., M. Campillo, R. D. Hilst, F. Brenguier, L. Stehly, and G. Hillers (2019), High Temporal Resolution Monitoring of Small Variations in Crustal Strain by Dense Seismic Arrays, *Geophysical Research Letters*, 46(1), 128–137, doi:10.1029/2018GL079944.
- [154] Martin, E. R. (2019), A scalable algorithm for cross-correlations of compressed ambient seismic noise, in *SEG Technical Program Expanded Abstracts 2019*, pp. 3006–3010, Society of Exploration Geophysicists, doi:10.1190/segam2019-3216637.1.
- [155] Mavko, G. M. (1979), Wave attenuation in partially saturated rocks, *Geophysics*, 44(2), 161, doi:10.1190/1.1440958.
- [156] McDonald, M. G., and A. W. Harbaugh (1984), A modular three-dimensional finite-difference groundwater flow model, *U.S. Geological Survey Open-File Report 83–875*.
- [157] McGuire, V. L., M. R. Johnson, R. L. Schieffer, J. S. Stanton, S. K. Sebree, and I. M. Verstraeten (2003), Water in Storage and Approaches to Ground- Water Management , High Plains Aquifer, 2000, *U.S. Geological Survey Circular*, 1243, 1–51.

- [158] McNamara, D. E., and R. P. Buland (2004), Ambient Noise Levels in the Continental United States, *Bulletin of the Seismological Society of America*, 94(4), 1517–1527, doi:10.1785/012003001.
- [159] Meier, M., T. Heaton, and J. Clinton (2015), The Gutenberg Algorithm: Evolutionary Bayesian Magnitude Estimates for Earthquake Early Warning with a Filter Bank, *Bulletin of the Seismological Society of America*, 105(5), 2774–2786, doi:10.1785/0120150098.
- [160] Meier, U., N. M. Shapiro, and F. Brenguier (2010), Detecting seasonal variations in seismic velocities within Los Angeles basin from correlations of ambient seismic noise, *Geophysical Journal International*, 181(2), 985–996, doi:10.1111/j.1365-246X.2010.04550.x.
- [161] Meng, L., J.-P. Ampuero, J. Stock, Z. Duputel, Y. Luo, and V. C. Tsai (2012), Earthquake in a Maze: Compressional Rupture Branching During the 2012 Mw 8.6 Sumatra Earthquake, *Science*, 337(6095), 724–726, doi:10.1126/science.1224030.
- [162] Minson, S. E., M. A. Meier, A. S. Baltay, T. C. Hanks, and E. S. Cochran (2018), The limits of earthquake early warning: Timeliness of ground motion estimates, *Science Advances*, 4(3), 1–11, doi:10.1126/sciadv.aag0504.
- [163] Minson, S. E., A. S. Baltay, E. S. Cochran, T. C. Hanks, M. T. Page, S. K. McBride, K. R. Milner, and M. A. Meier (2019), The Limits of Earthquake Early Warning Accuracy and Best Alerting Strategy, *Scientific Reports*, 9(1), 1–13, doi:10.1038/s41598-019-39384-y.
- [164] Minson, S. E., A. S. Baltay, E. S. Cochran, S. K. McBride, and K. R. Milner (2020), Shaking is Almost Always a Surprise: The Earthquakes That Produce Significant Ground Motion, *Seismological Research Letters*, doi:10.1785/0220200165.
- [165] Mordret, A., A. D. Jolly, Z. Duputel, and N. Fournier (2010), Monitoring of phreatic eruptions using Interferometry on Retrieved Cross-Correlation Function from Ambient Seismic

- Noise: Results from Mt. Ruapehu, New Zealand, *Journal of Volcanology and Geothermal Research*, 191(1-2), 46–59, doi:10.1016/j.jvolgeores.2010.01.010.
- [166] Mordret, A., T. D. Mikesell, C. Harig, B. P. Lipovsky, and A. Prieto (2015), Monitoring South-West Greenland’s ice sheet melt with ambient seismic noise, *Science Advances*, (May), 1–11, doi:10.1126/sciadv.1501538.
- [167] Mousavi, S. M., and G. C. Beroza (2020), A Machine-Learning Approach for Earthquake Magnitude Estimation, *Geophysical Research Letters*, 47(1), 1–7, doi:10.1029/2019GL085976.
- [168] Mousavi, S. M., W. Zhu, Y. Sheng, and G. C. Beroza (2019), CRED: A Deep Residual Network of Convolutional and Recurrent Units for Earthquake Signal Detection, *Scientific Reports*, 9(1), 1–14, doi:10.1038/s41598-019-45748-1.
- [169] Mousavi, S. M., Y. Sheng, W. Zhu, and G. C. Beroza (2019), STanford EArthquake Dataset (STEAD): A Global Data Set of Seismic Signals for AI, *IEEE Access*, 7, 179,464–179,476, doi:10.1109/ACCESS.2019.2947848.
- [170] Mousavi, S. M., W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza (2020), Earthquake transformer—an attentive deep-learning model for simultaneous earthquake detection and phase picking, *Nature Communications*, 11(1), 1–12, doi:10.1038/s41467-020-17591-w.
- [171] Murnaghan, F. D. (1937), Finite Deformations of an Elastic Solid, *American Journal of Mathematics*, 59(2), 235, doi:10.2307/2371405.
- [172] Murray, K. D., and R. B. Lohman (2018), Short-lived pause in Central California subsidence after heavy winter precipitation of 2017, *Science Advances*, 4(8), eaar8144, doi:10.1126/sciadv.aar8144.

- [173] Nadeau, R. M., W. Foxall, and T. V. McEvilly (1995), Clustering and periodic recurrence of microearthquakes on the san andreas fault at parkfield, california., *Science*, 267(5197), 503–507, doi:10.1126/science.267.5197.503.
- [174] Nakata, N., P. Boué, F. Brenguier, P. Roux, V. Ferrazzini, and M. Campillo (2016), Body and surface wave reconstruction from seismic noise correlations between arrays at Piton de la Fournaise volcano, *Geophysical Research Letters*, 43(3), 1047–1054, doi: 10.1002/2015GL066997.
- [175] Nakata, N., L. Gualtieri, and A. Fichtner (Eds.) (2019), *Seismic Ambient Noise*, 1, 532 pp., Cambridge University Press, Cambridge, doi:10.1017/9781108264808.
- [176] National Oceanic and Atmospheric Administration (2005), The Second Wettest Rainfall Season in Los Angeles Come to an End, *Tech. rep.*, Oxnard, CA.
- [177] Nishimura, T., S. Tanaka, T. Yamawaki, H. Yamamoto, T. Sano, M. Sato, H. Nakahara, N. Uchida, S. Hori, and H. Sato (2005), Temporal changes in seismic velocity of the crust around Iwate volcano, Japan, as inferred from analyses of repeated active seismic experiment data from 1998 to 2003, *Earth, Planets and Space*, 57(6), 491–505, doi: 10.1186/BF03352583.
- [178] Nur, A., and G. Simmons (1969), Stress-induced velocity anisotropy in rock: An experimental study, *Journal of Geophysical Research*, 74(27), 6667, doi:10.1029/JB074i027p06667.
- [179] Nur, A., and G. Simmons (1969), The effect of saturation on velocity in low porosity rocks, *Earth and Planetary Science Letters*, 7(2), 99–108, doi:10.1016/0012-821X(69)90035-1.
- [180] Oakley, D. O. S., B. Forsythe, X. Gu, A. A. Nyblade, and S. L. Brantley (2021), Seismic ambient noise analyses reveal changing temperature and water signals to 10s of meters depth in the critical zone, *Journal of Geophysical Research: Earth Surface*, doi:10.1029/2020jf005823.

- [181] Obermann, A., and G. Hillers (2019), Seismic time-lapse interferometry across scales, *60*.
- [182] Obermann, A., T. Planès, E. Larose, C. Sens-Schönfelder, and M. Campillo (2013), Depth sensitivity of seismic coda waves to velocity perturbations in an elastic heterogeneous medium, *Geophysical Journal International*, *194*(1), 372–382, doi:10.1093/gji/ggt043.
- [183] Obermann, A., T. Planès, E. Larose, and M. Campillo (2013), Imaging preeruptive and coeruptive structural and mechanical changes of a volcano with ambient seismic noise, *Journal of Geophysical Research: Solid Earth*, *118*(12), 6285–6294, doi:10.1002/2013JB010399.
- [184] Obermann, A., T. Planès, C. Hadziioannou, and M. Campillo (2016), Lapse-time-dependent coda-wave depth sensitivity to local velocity perturbations in 3-D heterogeneous elastic media, *Geophysical Journal International*, *207*(1), 59–66, doi:10.1093/gji/ggw264.
- [185] O’Connell, R. J., and B. Budiansky (1974), Seismic velocities in dry and saturated cracked solids, *Journal of Geophysical Research*, *79*(35), 5412–5426, doi:10.1029/JB079i035p05412.
- [186] Olson, B. P. E., and B. J. Swanson (2017), Preliminary Geologic Map of the Lebec 7.5’ Quadrangle, Kern, Los Angeles, and Ventura Counties, California, *Tech. rep.*, doi:10.5069/G97P8W9T.
- [187] Ostrovsky, L. A., and P. A. Johnson (2001), Dynamic nonlinear elasticity in geomaterials, *La Rivista del Nuovo Cimento*, *24*(7), 1–46, doi:10.1007/BF03548898.
- [188] Pacheco, C., and R. Snieder (2005), Time-lapse travel time change of multiply scattered acoustic waves, *The Journal of the Acoustical Society of America*, *118*(3), 1300–1310, doi:10.1121/1.2000827.
- [189] Pavlis, G. L., and F. L. Vernon (2010), Array processing of teleseismic body waves with the USArray, *Computers and Geosciences*, *36*(7), 910–920, doi:10.1016/j.cageo.2009.10.008.

- [190] Payan, C., V. Garnier, J. Moysan, and P. A. Johnson (2009), Determination of third order elastic constants in a complex solid applying coda wave interferometry, *Applied Physics Letters*, *94*(1), 011,904, doi:10.1063/1.3064129.
- [191] Perol, T., M. Gharbi, and M. Denolle (2018), Convolutional neural network for earthquake detection and location, *Science Advances*, *4*(2), e1700,578, doi:10.1126/sciadv.1700578.
- [192] Perrone, D., and S. Jasechko (2017), Dry groundwater wells in the western United States, *Environmental Research Letters*, *12*(10), doi:10.1088/1748-9326/aa8ac0.
- [193] Poupinet, G., W. L. Ellsworth, and J. Frechet (1984), Monitoring velocity variations in the crust using earthquake doublets: An application to the Calaveras Fault, California, *Journal of Geophysical Research*, *89*(B7), 5719, doi:10.1029/JB089iB07p05719.
- [194] Prieto, G. A., J. F. Lawrence, and G. C. Beroza (2009), Anelastic Earth structure from the coherency of the ambient seismic field, *Journal of Geophysical Research: Solid Earth*, *114*(7), 1–15, doi:10.1029/2008JB006067.
- [195] Ralph, F. M., P. J. Neiman, G. N. Kiladis, K. Weickmann, and D. W. Reynolds (2011), A Multiscale Observational Case Study of a Pacific Atmospheric River Exhibiting Tropical–Extratropical Connections and a Mesoscale Frontal Wave, *Monthly Weather Review*, *139*(4), 1169–1189, doi:10.1175/2010MWR3596.1.
- [196] Ranasinghe, N. R., L. L. Worthington, C. Jiang, B. Schmandt, T. S. Finlay, S. L. Bilek, and R. C. Aster (2018), Upper-Crustal Shear-Wave Velocity Structure of the South-Central Rio Grande Rift above the Socorro Magma Body Imaged with Ambient Noise by the Large-N Sevilleta Seismic Array, *Seismological Research Letters*, *89*(5), 1708–1719, doi:10.1785/0220180074.

- [197] Ratdomopurbo, A., and G. Poupinet (1995), Monitoring a temporal change of seismic velocity in a volcano: Application to the 1992 eruption of Mt. Merapi (Indonesia), *Geophysical Research Letters*, 22(7), 775–778, doi:10.1029/95GL00302.
- [198] Reasenber, P. A., and K. Aki (1974), A precise, continuous measurement of seismic velocity for monitoring in situ stress, *Journal of Geophysical Research*, 79(2), 399–406, doi:10.1029/JB079i002p00399.
- [199] Retailleau, L., P. Boué, L. Li, and M. Campillo (2020), Ambient seismic noise imaging of the lowermost mantle beneath the North Atlantic Ocean, *Geophysical Journal International*, 222(2), 1339–1351, doi:10.1093/gji/ggaa210.
- [200] Riahi, N., and P. Gerstoft (2015), The seismic traffic footprint: Tracking trains, aircraft, and cars seismically, *Geophysical Research Letters*, 42(8), 2674–2681, doi:10.1002/2015GL063558.
- [201] Riahi, N., and P. Gerstoft (2017), Using graph clustering to locate sources within a dense sensor array, *Signal Processing*, 132, 110–120, doi:10.1016/j.sigpro.2016.10.001.
- [202] Rice, J. R., and M. P. Cleary (1976), Diffusion Solutions for Fluid-Saturated Elastic Porous M Constituents, *Reviews of Geophysics and Space Sciences*, 14(2), 227–241.
- [203] Richardson, C. W. (1981), Stochastic simulation of daily precipitation, temperature, and solar radiation, *Water Resources Research*, 17(1), 182–190, doi:10.1029/WR017i001p00182.
- [204] Richter, T., C. Sens-Schönfelder, R. Kind, and G. Asch (2014), Comprehensive observation and modeling of earthquake and temperature-related seismic velocity changes in northern Chile with passive image interferometry, *Journal of Geophysical Research: Solid Earth*, 119(6), 4747–4765, doi:10.1002/2013JB010695.

- [205] Rickett, J., and J. Claerbout (1999), Acoustic daylight imaging via spectral factorization: Helioseismology and reservoir monitoring, *1999 SEG Annual Meeting*, pp. 957–960, doi:10.1190/1.1820854.
- [206] Riley, F. S. (1969), Analysis of Borehole Extensometer Data from Central California, *International Association of Hydrologic Sciences*, pp. 423–431.
- [207] Rivet, D., F. Brenguier, and F. Cappa (2015), Improved detection of preruptive seismic velocity drops at the Piton de la Fournaise volcano, *Geophysical Research Letters*, 42(15), 6332–6339, doi:10.1002/2015GL064835.
- [208] Robeson, S. M. (2015), Revisiting the recent California drought as an extreme value, *Geophysical Research Letters*, 42(16), 6771–6779, doi:10.1002/2015GL064593.
- [209] Rodell, M., and J. S. Famiglietti (2002), The potential for satellite-based monitoring of groundwater storage changes using GRACE: The High Plains aquifer, Central US, *Journal of Hydrology*, 263(1-4), 245–256, doi:10.1016/S0022-1694(02)00060-4.
- [210] Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, 460(7258), 999–1002, doi:10.1038/nature08238.
- [211] Roeloffs, E. (1996), Poroelastic Techniques in the Study of Earthquake-Related Hydrologic Phenomena, *Advances in Geophysics*, 38(C), 135–195, doi:10.1016/S0065-2687(08)60270-8.
- [212] Roeloffs, E. A. (1988), Fault stability changes induced beneath a reservoir with cyclic variations in water level, *Journal of Geophysical Research*, 93(B3), 2107, doi:10.1029/JB093iB03p02107.

- [213] Rojstaczer, S., S. Wolf, and R. Michel (1995), Permeability enhancement in the shallow crust as a cause of earthquake-induced hydrological changes, *Nature*, 373(6511), 237–239, doi:10.1038/373237a0.
- [214] Ross, Z. E., M. A. Meier, E. Hauksson, and T. H. Heaton (2018), Generalized seismic phase detection with deep learning, *Bulletin of the Seismological Society of America*, 108(5), 2894–2901, doi:10.1785/0120180080.
- [215] Ross, Z. E., M. A. Meier, and E. Hauksson (2018), P Wave Arrival Picking and First-Motion Polarity Determination With Deep Learning, *Journal of Geophysical Research: Solid Earth*, 123(6), 5120–5129, doi:10.1029/2017JB015251.
- [216] Ross, Z. E., D. T. Trugman, E. Hauksson, and P. M. Shearer (2019), Searching for hidden earthquakes in Southern California, *Science*, 771(May), 767–771, doi:10.1126/science.aaw6888.
- [217] Rubinstein, J. L. (2004), Evidence for Widespread Nonlinear Strong Ground Motion in the MW 6.9 Loma Prieta Earthquake, *Bulletin of the Seismological Society of America*, 94(5), 1595–1608, doi:10.1785/012004009.
- [218] Sabra, K. G. (2005), Extracting time-domain Green's function estimates from ambient seismic noise, *Geophysical Research Letters*, 32(3), L03,310, doi:10.1029/2004GL021862.
- [219] Sager, K., C. Boehm, L. Ermert, L. Krischer, and A. Fichtner (2020), Global-Scale Full-Waveform Ambient Noise Inversion, *Journal of Geophysical Research: Solid Earth*, 125(4), doi:10.1029/2019jb018644.
- [220] Sato, H., and M. C. Fehler (2009), *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*, Springer, Berlin, Heidelberg.

- [221] Save, H., S. Bettadpur, and B. D. Tapley (2016), High-resolution CSR GRACE RLo5 mascons, *Journal of Geophysical Research: Solid Earth*, 121(10), 7547–7569, doi:10.1002/2016JB013007.
- [222] Saygin, E., P. R. Cummins, and D. Lumley (2017), Retrieval of the P wave reflectivity response from autocorrelation of seismic noise: Jakarta Basin, Indonesia, *Geophysical Research Letters*, 44(2), 792–799, doi:10.1002/2016GL071363.
- [223] Scanlon, B. R., R. W. Healy, and P. G. Cook (2002), Choosing appropriate technique for quantifying groundwater recharge, *Hydrogeology Journal*, 10, 18–39, doi:10.1007/s10040-0010176-2.
- [224] Schimmel, M. (1999), Phase Cross-Correlations : Design , Comparisons , and Applications, *Bulletin of the Seismological Society of America*, 89(5), 1366–1378.
- [225] Schippkus, S., M. Garden, and G. Bokelmann (2020), Characteristics of the Ambient Seismic Field on a Large-N Seismic Array in the Vienna Basin, *Seismological Research Letters*, 91(5), 2803–2816, doi:10.1785/0220200153.
- [226] Schurr, D. P., J.-Y. Kim, K. G. Sabra, and L. J. Jacobs (2011), Damage detection in concrete using coda wave interferometry, *NDT & E International*, 44(8), 728–735, doi:10.1016/j.ndteint.2011.07.009.
- [227] Seats, K. J., J. F. Lawrence, and G. A. Prieto (2012), Improved ambient noise correlation functions using Welch’s method, *Geophysical Journal International*, 188(2), 513–523, doi:10.1111/j.1365-246X.2011.05263.x.
- [228] Segall, P. (2010), *Earthquake and Volcano Deformation*, vol. 48, 48–0287 pp., Princeton University Press, Princeton, doi:10.1515/9781400833856.

- [229] Sens-Schönfelder, C., and T. Eulenfeld (2019), Probing the in situ Elastic Nonlinearity of Rocks with Earth Tides and Seismic Noise, *Physical Review Letters*, 122(13), 138,501, doi: 10.1103/PhysRevLett.122.138501.
- [230] Sens-Schönfelder, C., and U. Wegler (2006), Passive image interferometry and seasonal variations of seismic velocities at Merapi Volcano, Indonesia, *Geophysical Research Letters*, 33(21), 1–5, doi:10.1029/2006GL027797.
- [231] Shangguan, Y., J. Li, Q. Liang, R. Alvarez, and I. McGraw (2019), Optimizing Speech Recognition For The Edge.
- [232] Shapiro, N. M., and M. Campillo (2004), Emergence of broadband Rayleigh waves from correlations of the ambient seismic noise, *Geophysical Research Letters*, 31(7), 8–11, doi: 10.1029/2004GL019491.
- [233] Shapiro, N. M., M. Campillo, L. Stehly, and M. H. Ritzwoller (2005), High-resolution surface-wave tomography from ambient seismic noise., *Science*, 307(5715), 1615–1618, doi: 10.1126/science.1108339.
- [234] Shelly, D. R., D. P. Hill, F. Massin, J. Farrell, R. B. Smith, and T. Taira (2013), A fluid-driven earthquake swarm on the margin of the Yellowstone caldera, *Journal of Geophysical Research E: Planets*, 118(9), 4872–4886, doi:10.1002/jgrb.50362.
- [235] Shi, W., J. Cao, Q. Zhang, Y. Li, and L. Xu (2016), Edge Computing: Vision and Challenges, *IEEE Internet of Things Journal*, 3(5), 637–646, doi:10.1109/JIOT.2016.2579198.
- [236] Shokouhi, P., A. Zoëga, and H. Wiggenhauser (2010), Nondestructive Investigation of Stress-Induced Damage in Concrete, *Advances in Civil Engineering*, 2010, 1–9, doi: 10.1155/2010/740189.

- [237] Silver, P. G., T. M. Daley, F. Niu, and E. L. Majer (2007), Active source monitoring of cross-well seismic travel time for stress-induced changes, *Bulletin of the Seismological Society of America*, 97(1 B), 281–293, doi:10.1785/0120060120.
- [238] Skempton, A. W. (1954), The Pore-Pressure Coefficients A and B, *Géotechnique*, 4(4), 143–147, doi:10.1680/geot.1954.4.4.143.
- [239] Smail, R. A., A. H. Pruitt, P. D. Mitchell, and J. B. Colquhoun (2019), Cumulative deviation from moving mean precipitation as a proxy for groundwater level variation in Wisconsin, *Journal of Hydrology X*, 5(June), 100,045, doi:10.1016/j.hydroa.2019.100045.
- [240] Smith, K. R. C. J. R. J. Z. H. F. T. L. Z., R.G. (2017), Estimating the permanent loss of groundwater storage in the southern San Joaquin Valley, California, *Water Resources Research*, 53(3), 2133–2148, doi:10.1002/2016WR019861.
- [241] Snieder, R. (2004), Extracting the Green’s function from the correlation of coda waves: a derivation based on stationary phase., *Physical review. E, Statistical, nonlinear, and soft matter physics*, 69(4 Pt 2), 046,610, doi:10.1103/PhysRevE.69.046610.
- [242] Snieder, R., A. Grêt, H. Douma, and J. Scales (2002), Coda Wave Interferometry for Estimating Nonlinear Behavior in Seismic Velocity, *Science*, 295(5563), 2253–2255, doi:10.1126/science.1070015.
- [243] Snieder, R., M. Miyazawa, E. Slob, I. Vasconcelos, and K. Wapenaar (2009), A Comparison of Strategies for Seismic Interferometry, *Surveys in Geophysics*, 30(4-5), 503–523, doi:10.1007/s10712-009-9069-z.
- [244] Soergel, D., H. A. Pedersen, L. Stehly, L. Margerin, and A. Paul (2020), Coda-Q in the 2.5–20 s period band from seismic noise: application to the greater Alpine area, *Geophysical Journal International*, 220(1), 202–217, doi:10.1093/gji/ggz443.

- [245] Stewart, C. A., D. Hancock, D. Stanzione, G. Turner, T. M. Cockerill, N. Merchant, J. Taylor, M. Vaughn, I. Foster, E. Skidmore, S. Tuecke, and N. I. Gaffney (2015), Jetstream: A self-provisioned, scalable science and engineering cloud environment, *ACM International Conference Proceeding Series*, 2015-July, doi:10.1145/2792745.2792774.
- [246] Strauss, J. A., Q. Kong, S. Pothan, S. Thompson, R. F. Mejia, S. Allen, S. Patel, and R. M. Allen (2020), MyShake Citizen Seismologists Help Launch Dual-Use Seismic Network in California, *Frontiers in Communication*, 5(May), 1–7, doi:10.3389/fcomm.2020.00032.
- [247] Sumiejski, L. E., and P. M. Shearer (2012), Temporal Stability of Coda Q-1 in Southern California, *Bulletin of the Seismological Society of America*, 102(2), 873–877, doi:10.1785/0120110181.
- [248] Sun, Q., C. Miao, Q. Duan, H. Ashouri, S. Sorooshian, and K. Hsu (2018), A Review of Global Precipitation Data Sets: Data Sources, Estimation, and Intercomparisons, *Reviews of Geophysics*, 56(1), 79–107, doi:10.1002/2017RG000574.
- [249] Swain, D. L. (2015), A tale of two California droughts: Lessons amidst record warmth and dryness in a region of complex physical and human geography, *Geophysical Research Letters*, 42(22), 9999–10,003, doi:10.1002/2015GL066628.
- [250] Swain, D. L., M. Tsiang, M. Haugen, D. Singh, A. Charland, B. Rajaratnam, and N. S. Diffenbaugh (2014), The Extraordinary California Drought of 2013 / 2014 : Character , Context , and the Role of Climate Change, *Bulletin of the American Meteorological Society*, (September), 3–7.
- [251] Swain, D. L., D. E. Horton, D. L. Swain, D. E. Horton, D. Singh, and N. S. Diffenbaugh (2016), Trends in atmospheric patterns conducive to seasonal precipitation and temperature

- extremes in California seasonal precipitation and temperature extremes, (April), 1–14, doi: 10.1126/sciadv.1501344.
- [252] Swain, D. L., B. Langenbrunner, J. D. Neelin, and A. Hall (2018), Increasing precipitation volatility in twenty-first-century California, *Nature Climate Change*, 8(5), 427–433, doi: 10.1038/s41558-018-0140-y.
- [253] Sweet, J. R., K. R. Anderson, S. Bilek, M. Brudzinski, X. Chen, H. DeShon, C. Hayward, M. Karplus, K. Keranen, C. Langston, F. C. Lin, M. B. Magnani, and R. L. Woodward (2018), A community experiment to record the full seismic wavefield in Oklahoma, *Seismological Research Letters*, 89(5), 1923–1930, doi:10.1785/0220180079.
- [254] Taira, T., F. Brenguier, and Q. Kong (2015), Ambient noise-based monitoring of seismic velocity changes associated with the 2014 Mw 6.0 South Napa earthquake, (September 2010), 6997–7004, doi:10.1002/2015GL065308.Received.
- [255] Takano, T., T. Nishimura, H. Nakahara, Y. Ohta, and S. Tanaka (2014), Seismic velocity changes caused by the Earth tide: Ambient noise correlation analyses of small-array data, *Geophysical Research Letters*, 41(17), 6131–6136, doi:10.1002/2014GL060690.
- [256] Tallaksen, L. (1995), A review of baseflow recession analysis, *Journal of Hydrology*, 165(1-4), 349–370, doi:10.1016/0022-1694(95)92779-D.
- [257] Talwani, P., L. Chen, and K. Gahalaut (2007), Seismogenic permeability, ks, *Journal of Geophysical Research: Solid Earth*, 112(7), 1–18, doi:10.1029/2006JB004665.
- [258] Taylor, R. G., B. Scanlon, P. Döll, M. Rodell, R. van Beek, Y. Wada, L. Longuevergne, M. Leblanc, J. S. Famiglietti, M. Edmunds, L. Konikow, T. R. Green, J. Chen, M. Taniguchi, M. F. P. Bierkens, A. MacDonald, Y. Fan, R. M. Maxwell, Y. Yechieli, J. J. Gurdak, D. M.

- Allen, M. Shamsudduha, K. Hiscock, P. J.-F. Yeh, I. Holman, and H. Treidel (2012), Ground water and climate change, *Nature Climate Change*, 3(4), 322–329, doi:10.1038/nclimate1744.
- [259] Teng, H., and G. Branstator (2017), Causes of extreme ridges that induce california droughts, *Journal of Climate*, 30(4), 1477–1492, doi:10.1175/JCLI-D-16-0524.1.
- [260] Theis, C. (1940), The Source of Water Derived from Wells, *Civil Engineering*, pp. 277–280.
- [261] Theis, C. V. (1935), The relation between the lowering of the Piezometric surface and the rate and duration of discharge of a well using groundwater storage, *Transactions American Geophysical Union*, 16(2), 519–524, doi:10.1029/TR016i002p00519.
- [262] Thomas, R., J. McAuliffe, M. Lam, R. Giordano, J. Regier, K. Pamnany, D. Schlegel, J. Revels, S. Howard, Prabhat, A. Noack, and K. Fischer (2019), Cataloging the visible universe through Bayesian inference in Julia at petascale, *Journal of Parallel and Distributed Computing*, 127, 89–104, doi:10.1016/j.jpdc.2018.12.008.
- [263] Tibuleac, I. M., and D. von Seggern (2012), Crust-mantle boundary reflectors in Nevada from ambient seismic noise autocorrelations, *Geophysical Journal International*, 189(1), 493–500, doi:10.1111/j.1365-246X.2011.05336.x.
- [264] Tolman, C. F., and J. F. Poland (1940), Ground-water, salt-water infiltration, and ground-surface recession in Santa Clara Valley, Santa Clara County, California, *Transactions, American Geophysical Union*, 21(1), 23, doi:10.1029/TR021i001p00023.
- [265] Toupin, R. A., and B. Bernstein (1961), Sound Waves in Deformed Perfectly Elastic Materials. Acoustoelastic Effect, *The Journal of the Acoustical Society of America*, 33(2), 216–225, doi:10.1121/1.1908623.

- [266] Trugman, D. T., M. T. Page, S. E. Minson, and E. S. Cochran (2019), Peak Ground Displacement Saturates Exactly When Expected: Implications for Earthquake Early Warning, *Journal of Geophysical Research: Solid Earth*, 124(5), 4642–4653, doi:10.1029/2018JB017093.
- [267] Tsai, V. C. (2010), The relationship between noise correlation and the Green's function in the presence of degeneracy and the absence of equipartition, *Geophysical Journal International*, 182(3), 1509–1514, doi:10.1111/j.1365-246X.2010.04693.x.
- [268] Tsai, V. C. (2011), A model for seasonal changes in GPS positions and seismic wave speeds due to thermoelastic and hydrologic variations, *Journal of Geophysical Research: Solid Earth*, 116(4), 1–9, doi:10.1029/2010JB008156.
- [269] van Dinther, C., L. Margerin, and M. Campillo (2021), Laterally varying scattering properties in the North Anatolian Fault Zone from ambient noise cross-correlations, *Geophysical Journal International*, 225(1), 589–607, doi:10.1093/gji/ggaa606.
- [270] Ventosa, S., M. Schimmel, and E. Stutzmann (2019), Towards the Processing of Large Data Volumes with Phase Cross-Correlation, *Seismological Research Letters*, doi:10.1785/0220190022.
- [271] Viens, L., M. A. Denolle, N. Hirata, and S. Nakagawa (2018), Complex Near-Surface Rheology Inferred From the Response of Greater Tokyo to Strong Ground Motions, *Journal of Geophysical Research: Solid Earth*, pp. 5710–5729, doi:10.1029/2018JB015697.
- [272] Wada, Y., L. P. Van Beek, C. M. Van Kempen, J. W. Reckman, S. Vasak, and M. F. Bierkens (2010), Global depletion of groundwater resources, *Geophysical Research Letters*, 37(20), 1–5, doi:10.1029/2010GL044571.

- [273] Wang, C. Y., and M. Manga (2015), New streams and springs after the 2014 Mw6.0 South Napa earthquake, *Nature Communications*, 6(May), 1–6, doi:10.1038/ncomms8597.
- [274] Wang, C.-Y., and M. Manga (2021), *Water and Earthquakes*.
- [275] Wang, H. F., X. Zeng, D. E. Miller, D. Fratta, K. L. Feigl, C. H. Thurber, and R. J. Mellors (2018), Ground motion response to an ML 4.3 earthquake using co-located distributed acoustic sensing and seismometer arrays, *Geophysical Journal International*, 213(3), 2020–2036, doi:10.1093/gji/ggy102.
- [276] Wang, Q. Y., F. Brenguier, M. Campillo, A. Lecointre, T. Takeda, and Y. Aoki (2017), Seasonal Crustal Seismic Velocity Changes Throughout Japan, *Journal of Geophysical Research: Solid Earth*, 122(10), 7987–8002, doi:10.1002/2017JB014307.
- [277] Wang, Y., P. Getreuer, T. Hughes, R. F. Lyon, and R. A. Saurous (2016), Trainable Frontend For Robust and Far-Field Keyword Spotting, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, (1), 5670–5674, doi:10.1109/ICASSP.2017.7953242.
- [278] Wapenaar, K. (2004), Retrieving the elastodynamic Green’s function of an arbitrary inhomogeneous medium by cross correlation, *Physical Review Letters*, 93(25), 1–4, doi:10.1103/PhysRevLett.93.254301.
- [279] Wapenaar, K., D. Draganov, R. Snieder, X. Campman, and A. Verdel (2010), Tutorial on seismic interferometry: Part 1 — Basic principles and applications, *Geophysics*, 75(5), 195–75, doi:10.1190/1.3457445.
- [280] Warden, P. (2018), Speech Commands: A Dataset for Limited-Vocabulary Speech Recognition.

- [281] Warden, P., and D. Situnayake (2020), *TinyML*, 504 pp.
- [282] Weaver, R., and O. Lobkis (2002), On the emergence of the Green's function in the correlations of a diffuse field: pulse-echo using thermal phonons, *Ultrasonics*, 40(1-8), 435–439, doi:10.1016/S0041-624X(02)00156-7.
- [283] Weaver, R. L. (1982), On diffuse waves in solid media, *The Journal of the Acoustical Society of America*, 71(6), 1608–1609, doi:10.1121/1.387816.
- [284] Weaver, R. L., and O. I. Lobkis (2001), Ultrasonics without a Source: Thermal Fluctuation Correlations at MHz Frequencies, *Physical Review Letters*, 87(13), 134,301, doi:10.1103/PhysRevLett.87.134301.
- [285] Webb, S. C. (2007), The Earth's 'hum' is driven by ocean waves over the continental shelves, 445(February), 754–757, doi:10.1038/nature05536.
- [286] Wegler, U., and C. Sens-Schönfelder (2007), Fault zone monitoring with passive image interferometry, *Geophysical Journal International*, 168(3), 1029–1033, doi:10.1111/j.1365-246X.2006.03284.x.
- [287] Wegler, U., H. Nakahara, C. Sens-Schönfelder, M. Korn, and K. Shiomi (2009), Sudden drop of seismic velocity after the 2004 Mw 6.6 mid-Niigata earthquake, Japan, observed with Passive Image Interferometry B06305, *Journal of Geophysical Research: Solid Earth*, 114(6), 1–11, doi:10.1029/2008JB005869.
- [288] Wen, Y., A. Behrangi, H. Chen, and B. Lambriksen (2018), How well were the early 2017 California Atmospheric River precipitation events captured by satellite products and ground-based radars?, *Quarterly Journal of the Royal Meteorological Society*, 144(S1), 344–359, doi:10.1002/qj.3253.

- [289] Wessel, P., J. F. Luis, L. Uieda, R. Scharroo, F. Wobbe, W. H. F. Smith, and D. Tian (2019), The Generic Mapping Tools Version 6, *Geochemistry, Geophysics, Geosystems*, 20(11), 5556–5564, doi:10.1029/2019GC008515.
- [290] Winkler, K. W., and X. Liu (1996), Measurements of third-order elastic constants in rocks, *The Journal of the Acoustical Society of America*, 100(3), 1392–1398, doi:10.1121/1.415986.
- [291] Winkler, K. W., and A. Nur (1982), Seismic attenuation: Effects of pore fluids and frictional-sliding, *GEOPHYSICS*, 47(1), 1–15, doi:10.1190/1.1441276.
- [292] Xiao, M., A. Koppa, Z. Mekonnen, B. R. Pagán, S. Zhan, Q. Cao, A. Aierken, H. Lee, and D. P. Lettenmaier (2017), How much groundwater did California’s Central Valley lose during the 2012-2016 drought?, *Geophysical Research Letters*, pp. 1–8, doi:10.1002/2017GL073333.
- [293] Yamamura, K., O. Sano, H. Utada, Y. Takei, S. Nakao, and Y. Fukao (2003), Long-term observation of in situ seismic velocity and attenuation, *Journal of Geophysical Research: Solid Earth*, 108(B6), 1–15, doi:10.1029/2002JB002005.
- [294] Yang, Y., M. H. Ritzwoller, A. L. Levshin, and N. M. Shapiro (2007), Ambient noise Rayleigh wave tomography across Europe, *Geophysical Journal International*, 168(1), 259–274, doi:10.1111/j.1365-246X.2006.03203.x.
- [295] Yeats, R. S. (2004), Tectonics of the san gabriel basin and surroundings, Southern California, *Bulletin of the Geological Society of America*, 116(9-10), 1158–1182, doi:10.1130/B25346.1.
- [296] Yoon, C. E., O. O’Reilly, K. J. Bergen, and G. C. Beroza (2015), Earthquake detection through computationally efficient similarity search, *Science Advances*, 1(11), e1501057, doi:10.1126/sciadv.1501057.

- [297] Yu, E., A. Bhaskaran, S.-l. Chen, Z. E. Ross, E. Hauksson, and R. W. Clayton (2021), Southern California Earthquake Data Now Available in the AWS Cloud, *Seismological Research Letters*, doi:10.1785/0220210039.
- [298] Yuan, C., J. Bryan, and M. Denolle (2021), Numerical comparison of time-, frequency- and wavelet-domain methods for coda wave interferometry, *Geophysical Journal International*, 226(2), 828–846, doi:10.1093/gji/ggab140.
- [299] Yukutake, H., and T. Nakajima (1988), In situ measurements of elastic wave velocity in a mine, and the effects of water and stress on their variation, *Tectonophysics*, 149, 165–175.
- [300] Zeng, X., and C. H. Thurber (2016), A Graphics Processing Unit Implementation for Time–Frequency Phase-Weighted Stacking, *Seismological Research Letters*, 87(2A), 358–362, doi:10.1785/0220150192.
- [301] Zhang, Y., O. Abraham, F. Grondin, A. Loukili, V. Tournat, A. L. Duff, B. Lascoup, and O. Durand (2012), Study of stress-induced velocity variation in concrete under direct tensile force and monitoring of the damage level by using thermally-compensated Coda Wave Interferometry, *Ultrasonics*, 52(8), 1038–1045, doi:10.1016/j.ultras.2012.08.011.
- [302] Zhang, Y., N. Suda, L. Lai, and V. Chandra (2017), Hello Edge: Keyword Spotting on Microcontrollers, pp. 1–14.

THIS THESIS WAS TYPESET using L^AT_EX, originally developed by Leslie Lamport and based on Donald Knuth's T_EX. The body text is set in 11 point Egenolff-Berner Garamond, a revival of Claude Garamont's humanist typeface. A template that can be used to format a PhD thesis with this look and feel has been released under the permissive MIT (X11) license, and can be found online at github.com/suchow/Dissertate.