

Investigating the Earthquake Cycle on Multiple Temporal and Spatial Scales using Satellites and Simulations

Thesis by
Oliver L. Stephenson

In Partial Fulfillment of the Requirements for the
Degree of
Doctor of Philosophy



CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2023
Defended August 2, 2022

© 2023

Oliver L. Stephenson
ORCID: 0000-0002-5509-090X

All rights reserved

To my family

*We shall not cease from exploration
And the end of all our exploring
Will be to arrive where we started
And know the place for the first time.*

—T. S. Eliot (1888-1965), *Little Gidding*

For Elasticity is the temper of matter to recover its place with vehemence

—Christopher Smart (1722-1771), *Jubilate Agno*

ACKNOWLEDGEMENTS

It is often said that it takes a village to raise a PhD student. In my case, it has taken several. Of course my academic “village” of advisors and collaborators have made this thesis possible, and I am profoundly grateful to them all. To Mark Simons, for encouraging me to apply to Caltech, teaching me to use my “geophysical sense of smell”, and tolerating my wayward focus. To Nadia Lapusta, for showing me how to represent the complexities of the world in computer models. To Zach Ross, for teaching me about machine learning. To Rob Clayton, for joining my thesis committee and taking me to deploy heavy seismometers in hot weather many years ago.

Yuan-Kai Liu and Yunjun Zhang worked closely with me on Chapter 1, and it was a joy to share ideas with them. Thea Ragon provided invaluable assistance in the modeling efforts of Chapter 2, while Cunren Liang gave lots of useful support on the ionosphere corrections used in Chapters 1 and 2. Tobias Koehne and Eric Zhan were a pleasure to work with for the damage mapping project presented in Chapter 3. Piyush Agram was a deep reservoir of knowledge on SAR. Valere Lambert and Stacy Larochelle gave insightful advice for the simulations in Chapter 4. Many others have helped, supported, educated, and inspired me during my time here, including Victor Tsai, Joe Kirschvink, Jean-Philippe Avouac, Paul Asimow, and Mike Gurnis. Thank you all.

To my Seismolab village, including office mates Vish, Erin, Ojashvi, Xiaolin, Zhe and Minyan, as well as Zhichao, Valere, Vasilije, Stacy, Celeste, Ethan, Olivia, Jorge, Jack, and everyone else on the third floor of South Mudd: thank you for creating such a warm and creative environment. My time at Caltech would not have been what it was without Donna Mireles, Rosemary Miller, Priscilla McLean, Kim Baker-Gatchalian, Sarah Gordon, Julie Lee, Carolina Oseguera, Scott Dugan, Cruz, and many other staff in the Geological and Planetary Sciences division. Thank you for supporting me in my whistling, and other endeavours. You make things happen, and I will miss you all.

To those who have set me on this path, I am profoundly grateful. To James Wilson, who nurtured my interest in the fundamentals of the Universe. To Graham Bone, for the worst mock interview of my life. To David Beckwith, who taught me how to think rigorously, and told me that I wouldn’t be able to understand everything. You may

be right, but I haven't stopped trying. Mike Payne and Alex Copley helped me on my academic path whilst at university, and Romain Jolivet and Bryan Riel supported me during my early work with InSAR. Jean-Francois Molinari and Fabian Barras kindled my interest in fracture mechanics during my stay at École polytechnique fédérale de Lausanne, where I was generously hosted by Jean-Philippe and Christine Dumont (vous êtes très gentils!).

Music is one of the great joys in my life, and I am thankful to have had a village of people with whom to share this joy. Nancy Sulahian, Sean Mullin, and everyone in the Caltech Glee Club and Chamber Choir gave me a chance to sing some of my favourite pieces, as well as explore new choral frontiers. Barb Catlin and the Caltech Jazz Improv group helped me with those thirds, sevenths and tritone substitutions when I decided to try something new. The chance to play with Secret Girlfriend, perhaps one of the hottest new bands on the LA music scene until COVID tragically killed our ambitions, was a real highlight of grad school. Thank you to Kat, Paul, Sean and Cory.

Another highlight was the experience of starting and running Caltech Letters, a platform for the research and ideas of the Caltech community. Working with such an amazing team to go from an idea, to an outlet that has helped scientists from across campus share their work with hundreds of thousands of people, has truly been one of the most rewarding experiences I have had on campus. Thank you to Aditi and Phelan for taking over leadership of the site, as well as Jeremy, Giuliana, Nicole, Vasilije, Ali, Namita, Heidi, and everyone else who has helped build the site into what it is today. Thanks also to Sean Carroll and Susanne Hall, for their guidance on how to share science with the world. I would never let a social gathering pass without promoting the site to anyone who would listen, and I will not waste this opportunity; find our website at <https://caltechletters.org>.

I believe that scientists have a responsibility to go beyond the ivory tower and play an active part in improving the world around them. Working with people in the GSC, CAH, SoC, C4BL, PTU, and probably some other acronyms that I have forgotten, including Peishi, Charles, Jane, Kriti, Ashay, Sarah, Dawna, Cora, Nate, and many others, helped me go beyond just talking about changing the world, and actually start doing it. In my journey into science policy, I am particularly indebted to Lucy Jones for the example that she has set through decades of “science activation”, for showing me the value of combining science with storytelling, and for her personal support. The AGU’s “Voices for Science” team, especially Liz Landau, gave me

the chance to build on my early policy work and convinced me that policy was the right next step for me. Greg Fletcher, and everyone else at the Caltech Y, have been a continual positive presence on campus. Thank you for your tireless efforts to help Caltech students make a world of difference.

Next, to my village of friends. To Stacy, who was the very model of discipline and self control in getting us to those early morning SCEC talks, along with Vasilije, Kat, Willy, and of course Panda: you were the ideal lockdown buddies. Thank you so much for the friendship, the support, and the tiramisu. To Steve and Celeste, who ably assisted me in the (unintended) consumption of an entire bottle of Scotch on my first night at Caltech, along with Peishi, Jeremy, Nicole, Katrina, Kelsey, Ben, Aida, Elyse, Shreyas, Valere, Sean, Arian, Konrad, Sarah, and many others with whom I have been so fortunate to have shared times both high and low during my time at Caltech: thank you. To my friends back in the UK, who I'm sure are beginning to wonder where I went, including Susie "bae" Wright, Lachlan, Paddy, Victoria, Katie, Pam, Richard, Michael, Will, Hugh, other Will, George, Catherine, Sam and Rachel: I have missed you all. My godparents, Clare, Howard, Rachael and Steve, have given me much support, advice and wisdom over the years. To Su, Peter, Emily, Alex and Ed: thank you for the many walks and adventures, and the battle that must rage eternally between Readers and Stephensons.

To Remaya, who is always in my corner. Thank you for your support, your love, your strength, your charm, your intelligence, your curiosity, and your kindness. It has been a privilege to share the highs and lows of the past few years with you. I'm so excited for what our next steps will bring.

The villages of Colebrooke, Devon (and, for one windswept year, Paible, Taransay) have made me who I am today. For my family, to whom I dedicate this thesis, I am filled with gratitude. Thank you to Roger, Rosemary, Felix and Jamie for your eternal love and support, even in the face of confusion about what it is I do exactly. Percaw.

This thesis was prepared on the traditional, ancestral, and unceded territory of the Gabrielino/Tongva peoples. The villages of Xaxaamonga (Hahamongna), Shevaanga, 'Akuuranga and many others once occupied these lands. Thank you to Tongva elder Julia Bogany for the chance to learn of the people who were here before we arrived, and will remain after we are gone.

ABSTRACT

The motion of the Earth's tectonic plates creates a gradual accumulation of stress at their boundaries, followed by a rapid release in earthquakes, a process known as the earthquake cycle. Studying this process is important because of the hazards earthquakes pose, but presents challenges due to the multi-scale nature of the problem—stresses build up over hundreds to thousands of years, while earthquakes break narrow fault zones in a matter of seconds. In this thesis, we combine a variety of techniques to study the earthquake cycle on multiple temporal and spatial scales, including satellite-based interferometric synthetic aperture radar (InSAR) to observe the slow deformation of the Earth over wide areas, and high-performance computational simulations to model faults during earthquakes. We begin by presenting a method for removing the signal of plate-tectonic motion in large-scale InSAR measurements, allowing for better observation of small ground deformations. We then use these corrections to study the Makran subduction zone, on the Iran-Pakistan border. Our InSAR-derived ground velocity map can resolve motions at the level of millimeters per year over an area of nearly one million square kilometers, and we use it to place constraints on the degree of coupling on the subduction megathrust. Next, we show how InSAR can be combined with deep learning techniques to rapidly map earthquake damage in all weather conditions, day and night. Such products will hopefully prove useful in future disaster response. Finally, we present computational simulations of dynamic earthquake ruptures with enhanced dynamic weakening due to thermal pressurization. We apply our simplified model to the creeping section of the San Andreas Fault, which is generally thought to be a barrier to earthquake rupture. Our results show how thermal pressurization can allow earthquakes to propagate partially or completely through the creeping section for a range of physically reasonable parameters. Our work illustrates how results from multiple fields can be combined to deliver new insights into the earthquake cycle and the hazards that it poses.

PUBLISHED CONTENT AND CONTRIBUTIONS

Stephenson, O. L., Liu, Y.-K., Yunjun, Z., Simons, M., Rosen, P., & Xu, X. (2022). The Impact of Plate Motions on Long-Wavelength InSAR-Derived Velocity Fields. *Geophysical Research Letters* (in review). <https://doi.org/10.1002/essoar.10511538.2>.

O.L.S. participated in the conception of the project, contributed to writing the code and analyzing the data, and led the writing of the manuscript. This work is adapted into Chapter 1.

Stephenson, O. L., Kohne, T., Zhan, E., Cahill, B. E., Yun, S. H., Ross, Z. E., & Simons, M. (2021). Deep Learning-Based Damage Mapping with InSAR Coherence Time Series. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–17. <https://doi.org/10.1109/TGRS.2021.3084209>.

O.L.S. participated in the conception of the project, contributed to writing the code and analyzing the data, and led the writing of the manuscript. This work is adapted into Chapter 3.

TABLE OF CONTENTS

Acknowledgements	v
Abstract	viii
Published Content and Contributions	ix
Table of Contents	ix
List of Illustrations	xii
List of Tables	xxviii
Introduction	1
0.1 Background	1
0.2 Thesis Overview	6
Chapter I: The Impact of Plate Motions on Long-Wavelength InSAR-Derived Velocity Fields	8
1.1 Introduction	8
1.2 The Reference Frame of InSAR Measurements	10
1.3 Data and Methods	13
1.4 Results	14
1.5 Discussion	15
1.6 Conclusions	20
Chapter II: Imaging the Makran Subduction Zone with Dense InSAR Time Series	21
2.1 Introduction	21
2.2 Data and Methods	23
2.3 Results	32
2.4 Discussion	48
2.5 Conclusions	55
Chapter III: Deep Learning-Based Damage Mapping with InSAR Coherence Time Series	56
3.1 Introduction	56
3.2 Background and Previous Work	59
3.3 Proposed Approach	63
3.4 Data	68
3.5 Results	69
3.6 Discussion	77
3.7 Conclusions	84
Chapter IV: Exploring the Effect of Thermal Pressurization on Earthquake Propagation through a Creeping Fault Using Simplified 2-D Models	86
4.1 Introduction	86
4.2 Methods	89
4.3 Analytical Models of Thermal Pressurization	96
4.4 Rupture Propagation Into the Creeping Section	102

4.5 Discussion	110
4.6 Conclusions	119
Conclusions and Outlook	122
Bibliography	128
Appendix A: Supplementary Materials for Chapter 1: The Impact of Plate Motions on Long-Wavelength InSAR-Derived Velocity Fields	153
A.1 Introduction	153
A.2 InSAR Processing Details	153
A.3 Further Details on InSAR Corrections and Residuals	155
A.4 Further Details on the Signal of Vertical and Horizontal Plate Motions in InSAR Measurements	162
A.5 The Impact of Plate Motion on 3D Velocity Fields Derived from InSAR	167
Appendix B: Supplementary Materials for Chapter 3: Deep Learning-Based Damage Mapping with InSAR Coherence Time Series	169
B.1 Phase variance and the logit transform	169
B.2 RNN Model and Training Details	169
B.3 Study Areas and Data	172
B.4 Training with a Variable Amount of Pre-Event Data	177
Appendix C: Some Thoughts on Productivity and Doing a PhD	180
C.1 My Approach	180
C.2 A Few Extra Pieces of Scattered Advice	186
C.3 Some Specific Tools That I Have Found Helpful	187
C.4 Recommended Reading	188

LIST OF ILLUSTRATIONS

<i>Number</i>	<i>Page</i>
1.1	
Illustration of how uniform horizontal and vertical motions result in ramps in InSAR-derived velocity measurements. (a) Satellite images acquired from ascending (ASC) and descending (DSC) orbital tracks, which have a varying LOS incidence angle across the track. The term “range” refers to the distance from the ground target to the satellite, with near range and far range the closest and furthest points from the satellite, respectively. Red and blue arrows represent the ground-to-satellite LOS vector in the near range for ASC and DSC tracks, respectively. Grey and black arrows represent plate motion in the reference frame of the satellite. For illustration purposes we assume the ASC and DSC tracks are parallel to each other but in opposite directions, and ignore Earth curvature. Figure not to scale. (b) Profile of the horizontal plate velocity projected into the LOS of the ASC and DSC tracks, against geographic distance along the ground. (c) Same as (b), except for vertical plate motion, resulting in opposite gradients in the LOS profiles. The observing geometry creates a small curvature in all profiles, which is exaggerated in the figure. For InSAR measurements, the LOS velocity is expressed relative to a point within the image, so each of these profiles would be vertically shifted to intersect with the x axis at the chosen reference point. . . .	12

- 1.2 Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 86 (ASC) over the Makran subduction zone. For plotting purposes, we remove the median value from each velocity field. Positive values represent apparent motion towards the satellite. Color bars are re-scaled between plots. Numbers in parentheses within the color bars refer to the 2nd and 98th percentiles of the velocity. “az” is the azimuth direction (satellite direction of motion), and “rg” the range direction (perpendicular to the satellite direction of motion). **(a)** No corrections applied. **(b)** Estimated ionosphere removed. **(c)** Tropospheric model removed. **(d)** SET model removed and DEM error (Topo) correction applied. **(e)** Plate motion correction applied. The positive signal around (28 °N, 62 °E) is post-seismic deformation from the 2013 Khash earthquake (Barnhart, Hayes, Samsonov, et al., 2014). **(f)** Applied ionospheric correction. **(g)** Applied tropospheric correction. **(h)** Applied SET correction. **(i)** Applied DEM error correction. Larger signals in the south may be bias from tropospheric residuals (Fattahi & Amelung, 2015). **(j)** Applied plate motion correction. 16
- 1.3 The impact of plate motion adjustments for InSAR tracks over the Makran subduction zone. For plotting purposes, we remove the median from each velocity field. **(a)** track 86 (ASC). After plate motion correction, the post-seismic signal from the 2013 Khash earthquake (Barnhart, Hayes, Samsonov, et al., 2014) can be more clearly seen in the south-east of the figure. **(b)** track 20 (DSC). **(c)** Location of tracks (a) and (b) and the velocity field of the Eurasian plate, used to correct the tracks. **(i)** Velocity before plate motion correction, but after other corrections have been applied. **(ii)** Velocity after plate motion correction. **(iii)** Applied plate motion correction. **(iv)** Across-track profile of the velocity before plate motion correction. The number below each profile is the gradient of the linear least squares fit to the profile. Note that profiles are plotted as a function of ground range, which increases with distance from the satellite. **(v)** Across-track profile of the velocity after plate motion correction. **(vi)** Across-track profile of the applied plate motion correction. 17

1.4	Same as Figure 1.3, but showing the tracks for the Gulf of Aqaba and Australia. (a) Gulf of Aqaba, track 87 (ASC). (b) Gulf of Aqaba, track 21 (DSC). (c) Location of tracks (a) and (b) and the velocity field of the Arabian plate, used to correct the tracks. (d) Australia, track 119 (DSC). (e) Australia, track 46 (DSC). (f) Location of tracks (d) and (e) and the velocity field of the Australian plate, used to correct the tracks. (i)-(vi) are as described in Figure 1.3.	18
2.1	Map of the Makran region. Earthquakes and faults are shown in red, coastlines and borders are shown in black. The large red arrow at 23°N indicated the direction of convergence of the Arabian plate under the Eurasian plate used in this study. S.H.: Strait of Hormuz. MZP: Minab-Zendan-Palami fault zone. SSZ: Sistan Suture Zone. O-N/Ch: Ornach-Nal/Chaman fault zone. Approximate locations of ruptures on the megathrust in 1765, 1851, and 1945 are taken from Byrne et al. (1992). Focal mechanisms for three earthquakes in 2013 (Minab, Khash, and Balochistan) are shown, along with the rupture trace for the Hoshab fault, which hosted the Balochistan earthquake. .	24
2.2	Example of errors in the ionosphere calculation (a) Velocity from a linear fit to the ionosphere correction time series for track 20 (descending). Note the large negative velocity around (30°N, 61.5 °E) as a result of unwrapping errors. There is also a subtle velocity ramp, of about 0.5 mm/yr, in the range direction, which is also seen in the ionosphere of other descending tracks (b) Average spatial coherence for all interferograms for track 20. Low coherence areas around (30°N, 61.5 °E) result in unwrapping errors which bias the ionospheric phase estimation. Filtering of the ionospheric phase then causes this error to affect surrounding areas, as can be seen in (a).	27
2.3	Megathrust geometry used in our forward models. The geometry is taken from Slab2 (Hayes et al., 2018), and we discretize the fault into 504 triangular patches. Black lines indicate coastlines and borders. .	31
2.4	Low coupling model. The shallowest patches are all assumed to be decoupled, although this makes minimal difference to the modeled on-shore deformation. Coupling is 0.5 at 13 km depth, then transitions linearly with depth to decoupled below 20 km. Profiles of the modeled velocities at 60°E are shown in Figure 2.19	32

2.5	High coupling model. The shallowest patches are all assumed to be decoupled, although this makes minimal difference to the on-shore deformation. The fault is fully coupled to 30 km depth, then transitions linearly with depth to decoupled below 40 km. Profiles at 60°E are shown in Figure 2.20.	33
2.6	Example of strong tropospheric fronts in the InSAR time series. All 2018 data from track 86 (asc) is shown, after ionosphere corrections but before other corrections have been applied. The first date (2018-01-11) is taken as the temporal reference, and the black point shows the spatial reference. Sharp tropospheric fronts can be seen between July and August. The large inset on the bottom right highlights a particularly strong front on 2018-08-27, with over 20 cm of apparent deformation over a few kilometers.	35
2.7	Impact of time series corrections for a point at (26°N, 60°E) relative to (32°N 60°E) for track 86a. Correction time series (grey) are subtracted from the displacement time series (black). Each time series is offset by 50 cm for display purposes. Corrections are applied cumulatively, meaning the time series at the bottom of the plot has had all the above corrections applied. Note that we do not apply plate motion corrections in the time series domain, so the final time series includes a secular trend due to plate motion. The final time series is shown on a larger scale in Figure 2.18(g).	36
2.8	Time series corrections for track 20d. Note the much weaker ionosphere for the descending track. The final time series is shown on a larger scale in Figure 2.18(h). Caption is otherwise the same as Figure 2.7	37
2.9	LOS velocity for ascending tracks after applying all corrections other than plate motion to every track, and merging. Plate motion creates a range-dependent velocity ramp in each track, resulting in a large ramp across the entire region. Track numbers are shown at the top of the figure. “az”: azimuth direction, the satellite direction of travel. “rg”: range direction, the look direction of the satellite. Positive velocities represent motion towards the satellite.	40
2.10	Same as Figure 2.9, except for descending tracks.	41
2.11	LOS velocity for ascending tracks after removing the effect of plate motion using a plate motion model.	41

2.12	Estimated standard deviation of the time series velocity (σ_v) assuming uncorrelated Gaussian errors and after outlier removal for ascending tracks. Uncertainties are expressed relative to a reference point in each track (which has 0 uncertainty), with each reference point at a latitude of 29°N.	42
2.13	LOS velocity for descending tracks after removing the effect of plate motion using a plate motion model.	42
2.14	Uncertainties for the descending tracks. Caption is otherwise the same as Figure 2.12.	43
2.15	Location of the velocity profiles plotted in Figure 2.16, shown over the merged ascending track LOS velocities. Areas which do not unwrap in every interferogram have been masked, and are excluded from profiles, other than for (g), where the temporal coherence mask is used.	43
2.16	Profiles through ascending (blue) and descending (red) track LOS velocity fields. Locations of the profiles are shown in Figure 2.15, and profile distances are calculated from the letter, to the letter primed (e.g., A to A'). (a) Region-spanning profile. (b) South to north profile at 58°E (25.5°N to 32°N). (c) South to north profile at 60°E (25°N to 32°N). (d) South to north profile at 62.1°E (25°N to 32°N). The left side of the profile spans post-seismic deformation from the 2013 Khash earthquake. (e) South to north profile at 64.1°E (25°N to 32°N). The left side of the profile spans post-seismic deformation from the 2013 Balochistan earthquake. Note that profiles (b)-(e) are chosen to lie within individual tracks to avoid offsets at the track boundaries. (f) Profile spanning post-seismic deformation from the 2013 Balochistan earthquake. (g) Profile spanning rapid subsidence near the city of Rafsanjan, Iran, likely due to aquifer depletion. . . .	44

- 2.17 Locations of time series (a)-(h) plotted in Figure 2.18, plotted over ascending track velocities. All time series are plotted relative to a reference point. The location of the time series is shown with a black triangle, and the location of the reference point is shown with a white triangle. Note that all time series are relative to a point within the same track as the chosen point, and are plotted before the removal of the plate motion signal (which is only significant for points far from their reference, such as G and H. Tracks and exact locations are given in Table 2.2. 45
- 2.18 InSAR LOS displacement time series illustrating a variety of different temporal behaviors in our region. The locations of the time series and reference points are given in Figure 2.17 and Table 2.2. **(a)** post-seismic deformation from the 2013 Balochistan earthquake near the point of peak deformation. **(b)** Post-seismic deformation from the 2013 Balochistan earthquake across a portion of the fault creeping at the surface. The proximity to the reference point (1 km away) results in much lower tropospheric noise compared to (a). **(c)** post-seismic deformation from the 2013 Khash earthquake. **(d)** post-seismic deformation from the 2013 Minab earthquake. **(e)** Creep on the Chaman fault, and co-seismic offset from an earthquake on 2018-06-27 (Dalaison et al., 2021). **(f)** Rapid subsidence near the city of Rafsanjan, Iran, likely due to aquifer discharge (Motagh et al., 2017). **(g)** Time series near the coast for track 86 after corrections shown in Figure 2.7. A linear trend partly due to plate motion can be seen. **(h)** Same as (g), but for track 20, after corrections shown in Figure 2.8. 46
- 2.19 Profiles at 60°E for the low coupling model shown in Figure 2.4. **(a)** Profiles through the LOS velocity and estimated uncertainties ($2\sigma_v$) for ascending (asc) track 86 and descending (dsc) track 20. All velocities and uncertainties are relative to the reference point. Thin blue (asc) and red (dsc) lines show the running median of the InSAR data. Thick blue (asc) and red (dsc) lines show the predicted LOS velocity from the forward model. **(b)** Profile through the megathrust depth, with the color showing the coupling. **(c)** Plot of the coupling from (b). **(d)** Surface elevation profile. 49

2.20	Same as Figure 2.19, but with higher coupling on the megathrust (coupling as shown in Figure 2.5). The increased coupling results in larger variations in the predicted InSAR LOS velocity, which are not consistent with our observations.	50
2.21	Same as Figures 2.19 and 2.20, but with intermediate coupling on the megathrust. This coupling is at the limit of what is compatible with our data, given the conservative errors bounds we have chosen. . . .	51
3.1	Schematic of the existing Coherence Change Detection (CCD) method for damage mapping (Yun et al., 2015), presented for the town of Sarpol-e-Zahab, damaged during the November 2017 Iran-Iraq earthquake. A pre-event coherence image (x_{T-1}) is subtracted from the co-event coherence image (x_T) in order to calculate the coherence loss. The coherence loss is thresholded and plotted to produce a damage proxy map. Optical data from Google, CNES/Airbus, taken July 27th 2020.	63
3.2	Schematic of our proposed recurrent neural network (RNN) method presented for the town of Sarpol-e-Zahab, damaged during the November 2017 Iran-Iraq earthquake. The transformed coherence values (x) are used to train a recurrent neural network to make a Gaussian forecast of the co-event coherence with mean μ'_{T+1} and standard deviation σ'_{T+1} . The forecast is compared with the observed co-event coherence, x_{T+1} , to calculate the z-score, z (see Eq. 3.10). The z-score is thresholded and plotted to produce a damage proxy map. A more detailed illustration of the neural network architecture can be found in Fig. B.1. Optical data from Google, CNES/Airbus, taken July 27th 2020.	64
3.3	RNN DPM for the town of Amatrice, Italy, badly damaged during the 2016 M_w 6.2 central Italy earthquake. The center of the town, which was largely destroyed, is clearly highlighted by elevated damage proxy values towards the top left of the map. Z-score values below 4.93 (chosen from the $F_{0.5}$ score, Eq. 3.11) are masked, values above 12 are set to red as indicated by the color bar. Ground truth damage data are presented in Figure B.3. Optical imagery from Google, taken July 6th 2017.	70

- 3.4 Quality metrics for the CCD and RNN damage classification methods.
- (a) Precision-recall curves for Amatrice damage proxy maps using the CCD (blue line) and RNN (red line) methods. For a perfect classifier we can choose a threshold that gives precision and recall equal to one, indicated in the top right corner. A random classifier gives a constant precision, equal to the fraction of the data set that is truly damaged, with recall varying with the threshold, indicated by the grey horizontal line at the bottom of the plot. The larger area under the curve (AUC) for the RNN method indicates improved performance. The black crosses show the position of maximum $F_{0.5}$ score identified in figures (b) and (c). (b) $F_{0.5}$ score (see Eq. 3.11) for varying z-score damage thresholds using the RNN method. (c) $F_{0.5}$ score for varying coherence loss damage thresholds using the CCD method. 71
- 3.5 Comparison of CCD (a) and RNN (b) DPMs for Sarpol-e-Zahab, damaged in the November 12, 2017 M_w 7.3 Iran-Iraq earthquake. For each plot we mask values below a threshold, with the threshold chosen using the maximum value of the $F_{0.5}$ curve for the Amatrice data set (see Table 3.1). The upper threshold of the color scales are chosen such that both plots have the same number of points above the threshold. The white dashed lines highlight example areas where the CCD method gives false positive damage detection in regions outside of the city that are no longer classified as damaged by the proposed RNN method. The black dashed line shows an area over a rocky ridge where greater damage is shown by the RNN method compared to the CCD method. Ground truth damage data are presented in Figure B.4. Optical imagery from Google, CNES/Airbus, taken July 27th 2020. . . 75

- 3.6 Sarpol-e-Zahab RNN damage map along with coherence time series and Gaussian forecasts for four representative locations ((a)-(d)) around the city. The z-score indicates the number of standard deviations between the forecast and the ground truth, and the coherence is plotted in logit space (i.e., it has been transformed into an unbounded space, see Eq. 3.9). Note that the shape of the coherence chip on the ground changes depending on the topography due to the way SAR data is acquired. The “Coherence” plotted on the y-axis is the logit transform of the squared coherence (Eq. 3.9). Optical imagery from Google, CNES/Airbus, taken July 27th 2020. 76
- 3.7 Ridgecrest RNN damage proxy map with different thresholds. (a) All points with $z < 4.93$ are masked. Black lines indicate mapped surface ruptures from Ponti *et al.* (Ponti et al., 2020). The white dashed line indicates the approximate extent of the dry Searles Lakebed that saw substantial liquefaction (Zimmaro et al., 2020). Global CMT focal mechanisms are plotted for the M_w 6.4 and M_w 7.1 earthquakes (Dziewonski et al., 1981; Ekström et al., 2012). The inset shows a regional map with simplified Quaternary faults. (b) Same damage proxy map as (a) except with no masking and points with $z < 0$ plotted in black. This threshold allows less intense off-fault anomalous change to be more clearly seen. 78
- 4.1 Examples of M_w 6 and M_w 7-style events. Blue contours show accumulated inter-seismic slip every four years. Red lines show accumulated co-seismic slip every 0.5 seconds. (a) M_w 6-style event. A 25-km-long velocity-weakening patch is placed between two velocity-strengthening patches, the crustal plane thickness is $H_{seis} = 2$ km. The velocity-strengthening patch on the left has TP with $\Lambda = 0.34$ MPa/K, $w = 10$ mm and $\alpha_{hy} = 10^{-4}$ m²/s for this case; the parameters are varied in other simulations (Table 4.1). The TP with these parameters is not efficient enough for the M_w 6-style event to propagate into the creeping section. (b) M_w 7-style event. The velocity-weakening patch is expanded to 80 km long, and crustal plane thickness increased to $H_{seis} = 8$ km. TP parameters have the same values as in (a). The larger event now activates sufficient TP to propagate through the creeping section and hit the boundary of our model (not shown here), allowing the event to grow well beyond M_w 7. 95

- 4.2 Evolution of parameters with slip for a point within the creeping section (position -15 km) during a M_w 7-style event that fully ruptures the creeping section. TP parameters are $\Lambda = 0.34$ MPa/K, $w = 10$ mm, and $\alpha_{hy} = 10^{-5}$ m²/s. **(a)** Shear stress (τ) with slip. **(b)** Effective normal stress (σ_{eff}) with slip. **(c)** Coefficient of friction (τ/σ_{eff}) with slip. **(d)** State variable term ($\ln[V^*\theta/D_{RS}]$) with slip. **(e)** Slip velocity with slip. The dashed line indicates the mean slip velocity since the start of the event. Note the 15 m/s slip velocity limit. **(f)** Temperature change with slip. 97
- 4.3 Schematic evolution of shear stress, τ , with slip for a point in the creeping section of the fault that experiences TP due to incoming dynamic rupture. The quantities illustrated are used in the theoretical estimates of Section 4.3.1. Steady aseismic creep happens with the slip velocity v_{pl} and stress τ_{creep} . When an event propagates into the creeping section, there is an initial rapid increase of stress to τ_{peak} (the rate-and-state direct effect) and then a decay to τ_{seis} , controlled by the steady-state rate-and-state parameter. TP can then be activated due to the resulting shear heating, causing the stress to continue to decrease. Given sufficient slip before the rupture arrests, the stress could decrease back to τ_{creep} over a slip of δ_{drop} , and then even below this level, allowing for a positive stress drop in the creeping section. A simulation output of this evolution can be seen in Figure 4.2(a). . . 99
- 4.4 Fraction of the creeping section ruptured for a M_w 6-style event and varying TP parameters. The value of the TP coupling coefficient, Λ , increases from (a) to (d), resulting in increasingly efficient TP and greater rupture fractions. Points within black squares are those models for which the event fully ruptures the creeping section, indicating that TP is too efficient to be consistent with the observation of the 2004 M_w 6 event arresting in the creeping section. 104

4.5	Fraction of the creeping section ruptured for a M_w 7-style event and varying TP parameters. The size of the TP coupling coefficient, Λ , increases from (a) to (d), resulting in increasingly efficient TP and greater rupture fractions. Points within black boxes are those that fully ruptured the creeping section for M_w 6-style events (Figure 4.4), indicating that TP is too efficient to be consistent with the observation of the 2004 M_w 6 event arresting in the creeping section. Points within red boxes are those that arrested in the creeping section for M_w 6-style events, but fully rupture the creeping section for M_w 7-style events.	105
-----	--	-----

- 4.6 Variation in the slip δ_{drop}^* (top row) required for a positive stress drop, assuming the slip-on-a-plane solution, as we change TP parameters w , α_{hy} and Λ , and validity of the approximation (bottom row). The simulated rupture fraction values are also shown, for a M_w 7-style event. **(a)** Colorbar shows variation of δ_{drop}^* with w and α_{hy} , for $\Lambda = 0.1$ MPa/K. Each color change delineates an order of magnitude change in δ_{drop}^* , which only varies with α_{hy} given the assumption of an infinitesimally thin fault in the corresponding analytical solution. The approximation is expected to work for faults that are thin enough compared to diffusion time scales; this regime is highlighted with bright colors (see panel (c)), with the rest of the panel muted. The dashed black and red lines show where $\delta_{drop}^* = \delta_{event}$ for the M_w 6-style and M_w 7-style events, respectively. Note that, within the region of the validity of the analytical approximation (bright colors), the red dashed line predicts the transition from rupture propagation over the creeping section to rupture arrest. **(b)** Same as (a), except with Λ increased by an order of magnitude to 1 MPa/K. This increased efficiency of TP results in δ_{drop}^* decreasing by two orders of magnitude for each value of α_{hy} . **(c)** Validity for the slip-on-a-plane approximation with $\Lambda = 0.1$ MPa/K. The colorbar shows variation of δ_d/δ_{drop}^* . The thick black contour delineates $\delta_{drop}^* = 10\delta_d$. The blue area to the left and top of this line satisfies $\delta_{drop}^* > 10\delta_d$, the range for which we take the slip-on-a-plane fault model to be a good approximation, as marked in (a). **(d)** Same as (c), except with $\Lambda = 1$ MPa/K. This increased efficiency of TP results in a reduced δ_{drop}^* , as shown in (b). This decrease causes the validity bound for the slip-on-a-plane approximation to shift to the left, reducing the number of our fault models that are well approximated by the slip-on-a-plane fault regime. 109

- 4.7 Same as Figure 4.6, except for the adiabatic, undrained fault solution (i.e. replacing δ_{drop}^* with $\delta_{drop}^{(c)}$). Note that the order of magnitude change in Λ between (a) and (b) causes an order of magnitude change in $\delta_{drop}^{(c)}$, compared to two orders of magnitude for δ_{drop}^* (shown in Figure 4.6). As the adiabatic, undrained fault solution assumes no transfer of fluid or heat out of the fault zone, $\delta_{drop}^{(c)}$ depends only on the fault zone width. The thick black contour in (c) and (d) now delineates $\delta_{drop}^{(c)} = 0.1\delta_d$; models to the right of this line satisfy $\delta_{drop}^{(c)} < 0.1\delta_d$, the range for which we take the adiabatic, undrained solution to be a good approximation to our models. As we go from (c) to (d), the increasing efficiency of TP results in a shorter $\delta_{drop}^{(c)}$, reducing the diffusion length scales and making more parameter combinations well approximated by the adiabatic, undrained regime. Similar to Figure 4.6, the red dashed line in (a) and (b) shows where $\delta_{drop}^{(c)} = \delta_{event}$ and marks a transition in the fraction of the creeping section ruptured. . . . 111
- 4.8 Logarithm of the fraction of the creeping section ruptured against the logarithm of the ratio between the total slip for the event in the absence of TP, δ_{event} , and the estimated slip required for a positive stress drop due to TP, δ_{drop} . For each combination of TP parameters, δ_{drop} is estimated for the slip-on-a-plane solution (δ_{drop}^*) and the adiabatic, undrained solution ($\delta_{drop}^{(c)}$). Only points for estimates that are expected to be good are plotted, defined by $\delta_{drop}^{(c)} < 0.1\delta_d$ and $\delta_{drop}^* > 10\delta_d$, as established based on Figures 4.6 and 4.7. Points with estimates of δ_{drop}^* outside of these validity bounds are plotted in Figure 4.9. The color of the symbols gives the log of the ratio between the hydrothermal diffusion distance, δ_d , and δ_{drop} , for the relevant end-member model. Values increasingly greater than 1 indicate that the fault model is increasingly closer to the adiabatic, undrained regime over the slip δ_{drop} , while values less than -1 indicate increasing closeness to the slip-on-a-plane fault regime. The line $\delta_{event} = \delta_{drop}$ captures the transition from ruptures arresting in the creeping region to increasingly large fraction of the creeping section being ruptured in our models. 112
- 4.9 Same as Figure 4.8, except for estimates of δ_{drop} that are outside of the validity bounds we choose (i.e. $\delta_{drop}^{(c)} > 0.1\delta_d$ and $\delta_{drop}^* < 10\delta_d$). 113

A.1	Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 20 (DSC) over the Makran subduction zone. Panels are as described in Figure 1.2. Note the substantial difference in the long-wavelength ionospheric correction between ASC (Figure 1.2(f)) and DSC tracks ((f)). The burst discontinuities due to the ionosphere can be clearly seen in (f) (Liang et al., 2019).	161
A.2	Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 87 (ASC) over the Gulf of Aqaba. Panels are as described in Figure 1.2.	162
A.3	Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 21 (DSC) over the Gulf of Aqaba. Panels are as described in Figure 1.2.	163
A.4	Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 46 (DSC) over western Australia. Panels are as described in Figure 1.2.	164
A.5	Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 119 (DSC) over western Australia. Panels are as described in Figure 1.2.	165
A.6	Size of velocity ramp in range for the 250 km wide Sentinel-1 swath for varying horizontal and vertical bulk motion. These calculations assume a constant plate velocity (i.e., no rotation), ignore the slight curvature in the across-track ramps (Figure 1), and take the LOS geometry from the ASC and DSC Sentinel-1 tracks over Aqaba (87 and 21, respectively). The LOS vector points from the ground to the satellite. Heading angles are measured clockwise from north. “LOS heading angle” is the heading angle of the horizontal projection of the LOS vector. “Heading angle of horizontal motion” is the heading angle of the bulk plate motion. A positive cross-track ramp means that LOS velocity increases with increasing satellite range. (a) Size of velocity ramp for horizontal motion in an ASC track. (b) Size of velocity ramp for horizontal motion in a DSC track. (c) Size of velocity ramp for vertical velocity motion. For vertical motion, the ramp has the same variation with range for ASC and DSC tracks. . . .	166

A.7	Impact of accounting for plate motion on the calculation of horizontal and vertical velocities from ASC and DSC tracks. The calculation is performed for the overlapping region between tracks 86 (ASC) and 20 (DSC) over the Makran subduction zone, with horizontal motion assumed to be purely east-west. The black square represents the reference point, which is assumed to be stationary with respect to the satellite. (a) Calculated eastward motion before plate motion correction, but after other corrections have been applied. Note the apparent north-south velocity gradient. (b) Calculated vertical motion before plate motion correction. Note the apparent east-west velocity gradient. (c), (d) Same as (a) and (b), respectively, but with the plate motion correction applied to each track before calculating the horizontal and vertical terms. Long wavelength velocity gradients have been substantially reduced.	168
B.1	Graphical model for our recurrent neural network for a series of time steps. We use gated recurrent units for f_ϕ and a feed-forward neural network for g_ψ . For each pixel the hidden state (h) summarizes the coherence information (x) up to that point, and is used to forecast the mean (μ') and standard deviation (σ') of the logit-transformed squared coherence. We initialize the hidden state h_0 as all zeros. The final hidden state, h_T , is used to forecast the co-event coherence distribution parameters ($\mu'_{T+1}, \sigma'_{T+1}$).	172
B.2	Schematic of the different coherence calculation methods for (a) $\mathcal{D}_t$, and (b) $\mathcal{D}_f$. The small squares represent SLC pixels and the large, numbered, squares represent coherence chips, here represented as three by three SLC pixels for illustration purposes.	174
B.3	Ground truth data used for quantifying damage classification performance for the town of Amatrice, damaged during the August 24, 2016 M_w 6.2 central Italy Earthquake. Building polygons and damage levels are supplied by the Copernicus Emergency Management Service (Copernicus Emergency Management Service, 2016), then simplified as described in Section B.3. Note that the optical imagery in this figure was taken a little under a year after the earthquake, and structures have been built that do not appear in the damage assessment. Optical imagery for the figure is from Google, taken July 6th 2017.	175

- B.4 Location of 683 potentially damaged buildings manually mapped from optical satellite imagery by the United Nations Institute for Training and Research (United Nations Institute for Training and Research, 2017) for the town of Sarpol-e-Zahab, damaged during the November 12, 2017 M_w 7.3 Iran-Iraq earthquake. Optical imagery for the figure is from Google, CNES/Airbus, taken July 27th 2020. . . . 176
- B.5 Impact of varying the amount of training data. **(a)** RNN Precision-recall area under curve (PR AUC) for the Amatrice data set, with varying amounts of pre-event training and forecasting coherence data (red line). Only data acquired on or after the indicated date are used in training and forecasting. The PR AUC for the CCD method is also presented for comparison (horizontal dashed blue line). Training that uses all data is indicated at the top left. Training that only uses two pre-event coherence images is indicated on the right hand side. **(b)** Mean of the logit transform of the squared coherence against the date of the first SAR acquisition in that coherence image. Values are presented up to the final pre-event coherence image. The co-event coherence is not used in training, so its mean is not presented here. Generally SAR images in each coherence image are separated by 12 or 24 days. The date of the August 2016 central Italy earthquake is indicated by “EQ” on the right of the both plots. 179

LIST OF TABLES

<i>Number</i>	<i>Page</i>
2.1 Summary of Sentinel-1 SAR Data used. ASC: Ascending track. DSC: Descending track. Aqn. no.: Number of acquisitions used for interferogram formation. Excl. no.: Number of acquisitions excluded using the 2 median absolute deviation threshold for velocity calculation. All tracks use Sentinel-1A acquisitions apart from track 159, which combines 1A and 1B. Tracks are listed from west to east. Each track is 250 km wide and adjacent tracks have a roughly 50 km overlap.	30
2.2 Details of the time series plotted in Figure 2.18. “Track” refers to the Sentinel-1 track, with “a” and “d” referring to ascending and descending tracks respectively. Time series are taken at the point given by “Location” and the InSAR reference point is set at “Reference.” “Distance” refers to the distance between the time series location and the reference point.	45
3.1 Optimum threshold and corresponding precision and recall values for both methods using the Amatrice data set, selected using the maximum value of the $F_{0.5}$ score, along with the area under the precision-recall curves (PR AUC).	73
4.1 Summary of simulation parameters. Parameters that we vary in this study are indicated in bold at the bottom, along with their ranges. DW: dynamic weakening. VW: velocity-weakening. VS: velocity-strengthening.	121
A.1 Summary of Sentinel-1 SAR data used. ASC: Ascending track. DSC: Descending track. Aqn. no.: Number of acquisitions. Makran and Aqaba use only Sentinel-1A data, Australia uses only Sentinel-1B. The plates used for plate motion corrections are listed in the final column and are taken from Table 1 of Altamimi et al. (2017).	158
B.1 Summary of SAR data used for each study area. N_{aqn} : number of pre-event acquisitions used, $N_{t,f}$: number of pixels in the training ($\mathcal{D}_t$) and forecasting ($\mathcal{D}_f$) sets, respectively, $E_{t,f}$: spatial extent of the training and forecasting regions, Track: Sentinel-1 orbital track and direction of flight.	177

INTRODUCTION

0.1 Background

Why do earthquakes happen? The Greek philosopher Aristotle (384-322 BCE) thought the culprit was the escape of evaporating water from the Earth, while for the ancient Chinese, who made the first written records of earthquakes, such events were the sign of an imbalance between heaven and Earth, and a harbinger of the ruling dynasty's fall from power (Needham, 1959). In 132 CE, the Chinese polymath Zhang Heng (78-139 CE) created the first seismoscope, a device that allowed the direction of distant earthquakes to be measured, even when seismic waves were imperceptible. The seismoscope could help the government dispatch relief to the affected areas; however, the underlying cause of earthquakes remained elusive.

In 1755, the Portuguese city of Lisbon was hit by a devastating earthquake and tsunami, triggering political and philosophical upheavals across Europe. The event was notable in the history of seismology for the detailed surveying of the shaking and damage, directed by the Portuguese chief minister, the Marquis of Pombal (1699-1782).¹ While interpreted as a sign of God's wrath, the disaster also prompted more naturalistic explanations, such as German philosopher Emmanuel Kant's (1724-1804) proposal that earthquakes were produced by the reaction of sulphur and iron with water in caverns beneath the Earth's surface (Reinhardt & Oldroyd, 1983).

It was not until 1910 that a key element in our modern understanding of earthquakes fell into place. In 1906, a rupture along the San Andreas Fault, in California, badly damaged the city of San Francisco. Harry Fielding Reid (1859-1944), professor of geology at Johns Hopkins University, found that the ground surface along the San Andreas Fault had deformed in the 50 years before the earthquake. Reid proposed that this deformation, or "strain," stored energy in the rocks beneath the surface, which was then released in the devastating earthquake: the "elastic-rebound theory" (Reid, 1910).

Reid's theory allowed him to suggest how future earthquakes could be predicted, by measuring where and by how much the Earth was deforming:

"As strains always precede the rupture and as the strains are sufficiently great to be easily detected before the rupture occurs, in order to foresee

¹Technically he did not gain the "Marquis of Pombal" title until 1769.

tectonic earthquakes it is merely necessary to devise a method of determining the existence of the strains; and the rupture will in general occur in the neighborhood of the line where the strains are greatest, or along an older fault-line where the rock is weakest” (Reid (1910), p. 31).

Reid went on to suggest a method for how this deformation could be measured:

“To measure the growth of strains, we should build a line of piers, say a kilometer apart, at right angles to the direction which a geological examination of the region, or past experience, indicates the fault will take when the rupture occurs; a careful determination from time to time, of the directions of the lines joining successive piers, their differences of level, and the exact distance between them, would reveal any strains which might be developing along the region the line of piers crosses” (Reid (1910), p. 31).

Reid was only able to speculate as to what was driving the deformation of the Earth, suggesting that “flows below the surface may have been the origin of the forces we have been considering” (Reid (1910), p. 27). In 1912, German meteorologist Alfred Wegener proposed “continental drift,” the idea that the present continents had once been part of a single land mass, and had slowly drifted apart. By the 1960s, this idea had developed into the modern theory of plate tectonics: the brittle outer layer of the Earth (the lithosphere²) is divided up into plates that slowly move past each other at a few centimeters per year (McKenzie & Parker, 1967). At their boundaries, the plates are stuck together, and so slowly bend as they move, storing up energy which is then released in earthquakes. The combination of the steady forcing from plate motion, with regular elastic rebound during earthquakes, suggests the idea of an “earthquake cycle”, a potentially periodic repetition of stress build up (the “inter-seismic” phase) followed by stress release (the “co-seismic” phase).

The picture of the Earth storing up elastic energy in the lithosphere, which is then released in earthquakes along faults (planes of weakness in the Earth), is complicated by more detailed observations of fault behavior. Faults do not just slip during earthquakes, they can also move slowly, at a speed of millimeters per

²In this introduction, I will use the terms “crust” and “lithosphere” interchangeably. This is technically incorrect—the lithosphere is the rigid outer layer of our planet, while the crust is the upper part of the lithosphere that is distinguished by its chemical composition. Sue me ☺

year, in a process known as “creep.” These observations can be explained by dividing faults into “velocity-strengthening” and “velocity-weakening” regions (e.g., Scholz (1998)). Velocity-strengthening sections of faults become stronger as the stresses build up and cause the fault to slip more quickly. Such strengthening behavior should make the fault unable to host earthquakes; fast slip on the fault causes a dramatic increase in strength, which rapidly slows any rupture. Within this model, velocity-strengthening sections should deform only by creep and act as barriers to earthquakes. Velocity-weakening sections, on the other hand, are “seismogenic”—they can rupture in earthquakes. With a build up in stress, the fault begins to move, growing weaker as it does. Such weakening allows it to slip faster, creating more weakening in a feedback loop that results in an earthquake. After this sudden slip, the fault becomes locked, once again storing up energy to be released in a future earthquake.

Measurements of the Earth’s surface allow us to identify regions of faults that are slowly creeping, and those that are locked (e.g., Burgmann (2000)). For example, we can divide the San Andreas Fault into locked sections in the north (where the 1906 earthquake occurred) and south, separated by a 140-km-long creeping section (e.g., Jolivet et al. (2015)). These measurements are easiest where faults are well exposed on land, as they are for the San Andreas, and are more challenging for offshore faults, such as the gigantic “megathrust” faults found at subduction zones, where one of the Earth’s plates slides under the other. Nevertheless, if we are able to identify the locked and creeping areas of faults, as well as the rate at which energy is being stored in the crust around the fault, such measurements can offer some indication of where, and how big, an earthquake might be (e.g., Avouac (2015)).

There are, however, numerous factors that further complicate this model. For example, not all energy that is used to deform the rock is necessarily recovered during the earthquake—some amount can be used in creating permanent crustal deformation, both during the inter-seismic and co-seismic phases, as well as during a “post-seismic” phase of deformation triggered by the earthquake (e.g., Avouac (2015), Johnson (2013), Perfettini and Avouac (2004), Shen et al. (1994), and K. Wang et al. (2021)). The straightforward division of faults into velocity-strengthening and velocity-weakening areas, leading to creeping and seismogenic behavior, respectively, is also being increasingly challenged by observations and experiments showing that apparently velocity-strengthening materials can dramatically weaken when driven at fast slip rates, a process known as “dynamic weakening” (Di Toro

et al., 2011; Tsutsumi & Shimamoto, 1997; Tullis, 2007). If velocity-strengthening patches can undergo dynamic weakening, then the idea that they represent barriers to earthquake rupture is in doubt (Noda & Lapusta, 2013).

Because of such complexities, understanding the earthquake cycle requires studying the Earth over many scales of time and space. The stresses in the Earth are built up by planetary scale motions of the Earth's lithosphere, which can load faults for hundreds to thousands of years between major earthquakes. While the interseismic loading can span many human lifetimes, the co-seismic rupture is orders of magnitude faster—the rupture propagates at kilometers per second, with meters of slip accumulating in just a few seconds (Heaton, 1990). During this rupture, dynamic weakening can allow faults to go from apparent barriers to earthquake rupture to dramatically weaker in a fraction of a second, and the width of the fault zone that slips can be tiny, as small as micrometers by some estimates (e.g., Chester and Chester (1998) and Platt et al. (2014)).

To do such multi-scale work, geoscientists must combine many fields of study. Geodesy, or the measurement of the shape of Earth's surface, can tell us about the deformation of the Earth over decades. In 1910, Reid was limited to the results of ground-based triangulation surveys of fixed locations before and after the earthquake. The development of space-based geodetic techniques, such as Global Navigation Satellite Systems (GNSS) and Interferometric Synthetic Aperture Radar (InSAR) now allow us to measure motions of millimeters per year, with dense temporal sampling, on a global scale (Blewitt, 2015; Simons & Rosen, 2015). These techniques can provide high resolution imaging of each stage of the earthquake cycle, from inter-seismic strain accumulation, to co-seismic offsets and post-seismic deformation (e.g., Avouac (2015), Fialko et al. (2001), Hsu et al. (2006), and Simons et al. (2011)).

Geodesy alone can only give us information over the time scales of the surveys. Given that the interval between large earthquakes on particular faults can be centuries, it is necessary to look to other techniques to gain a fuller understanding of the history of each fault—i.e., when and how big its previous earthquakes were. Large earthquakes can cause several meters of slip and alter the rocks and soil surrounding the fault. The field of paleoseismology searches for signs of these processes to create histories of previous earthquakes on major faults, potentially going back thousands of years (e.g., Sieh (1978b) and Sieh et al. (1989)). Geological investigations of fault

zones can be used to understand the total amount of slip on the fault, and how that slip was accommodated, over millions of years (e.g., Chester and Chester (1998))

The work described above is primarily observational—it relies on recording and understanding natural phenomena, rather than controlling them. Experiments present an alternative approach; taking samples, potentially obtained directly from fault zones, and measuring how they respond under different conditions. For example, scientists can adjust the pressure and water saturation conditions of a fault sample, and see how this affects its frictional strength at different sliding velocities, searching for signs of dynamic weakening (French et al., 2014). The controlled conditions of the laboratory allows for measurements on the scale of micrometers and microseconds (e.g., Rubino et al. (2022)), providing insights that are impossible to obtain directly from faults buried deep beneath the Earth’s surface.

The combination of observation and experimentation presents a messy, tangled picture of the physical processes acting in the Earth. It is the job of modeling to take this picture and reduce it to understandable parameters that give some insight into the physical processes that are acting. A classic example is Newton’s second law, which takes a series of observations of the world, from pushing a glass across a table, to the orbit of the Earth around the Sun, and unifies them under one theoretical model: $F = ma$. This equation tells us that the acceleration of a body (a) is equal to the net force we apply (F), divided by its mass (m), and allows us to explain how bodies move in response to forces. In a similar fashion, Reid’s elastic rebound theory reduced observations of ground deformation along the San Andreas fault into a model of energy stored in and then released from an elastic crust.

Just as an experimentalist measures how their samples respond to different conditions in the laboratory, models can be experimented with by varying their parameters and calculating the output, then comparing the results with real-world observations. When models are sufficiently complex, this experimentation can only be done with the aid of computer simulations. For example, when studying the earthquake cycle, computer simulations can be used to understand how the addition of dynamic weakening to velocity-weakening and velocity-strengthening patches could affect fault behavior over thousands of years, and compare these outputs with observed earthquakes (Cubas et al., 2015; Noda & Lapusta, 2013).

0.2 Thesis Overview

In this thesis, we draw on insights from a wide range of observational, experimental, computational and theoretical studies to better understand different components of the earthquake cycle. We focus our attention on geodetic measurements from InSAR, computational simulations of earthquake rupture, and theoretical understanding of these simulations.

We begin by discussing long-wavelength signals in InSAR measurements of ground velocity (Chapter 1). InSAR is a powerful technique for measuring ground deformation over wide areas; however, the data often contains contaminating signals that limit the ability of InSAR alone to measure deformation over hundreds of kilometers. This issue is generally solved by combining InSAR with GNSS measurements to constrain long-wavelength deformation. We demonstrate that the movement of the Earth’s tectonic plates creates long-wavelength signals in InSAR measurements that could potentially be misinterpreted as deformation, and show how to remove them using plate motion models.

Using the plate-motion corrections, we apply InSAR data to measure ground deformation over the Makran subduction zone, a 1000 km long megathrust where the Arabian plate slides beneath the Eurasian plate on the Iran-Pakistan border (Chapter 2). The approach is similar in spirit to Reid’s 1910 proposal to understand the earthquake cycle by measuring strain accumulation, but we have the advantage of 110 years of progress in geodesy. We are able to resolve on-going post-seismic deformation, co-seismic offsets, and rapid ground subsidence due to aquifer discharge over an area of nearly one million square kilometers using 7.5 years of InSAR data, without the aid of GNSS. Of particular interest is the degree of “coupling” on the subduction zone—i.e., how much of the motion between the plates is stored as elastic energy that could be released in future earthquakes. Our measurements do not allow us to unambiguously resolve the degree of coupling, but we use simple elastic models to place constraints on the strength of coupling and compare the results to previous studies.

In Chapter 3 we follow in the footsteps of Zhang Heng and the Marquis of Pombal by developing a method to rapidly map earthquake damage over wide areas using InSAR satellites. Using large amounts of InSAR data, combined with modern deep-learning techniques, our method can locate collapsed buildings within hours to days after major disasters, which could help to direct emergency response, particularly in

remote areas. The use of InSAR means that our method works day and night, and in all weather conditions.

For the final chapter, we shift our attention to high-performance computational simulations of earthquake rupture (Chapter 4). Building on work by Noda and Lapusta (2013) and Cubas et al., 2015, we explore the conditions under which dynamic weakening can allow creeping faults to rupture in earthquakes. We focus on a dynamic weakening process known as “thermal pressurization” (Sibson, 1973), and use our model to simulate the behavior of the creeping section of the San Andreas Fault as we vary the efficiency of weakening. We find that thermal pressurization could allow for ruptures to propagate through the creeping section within our simplified model, potentially allowing ruptures to grow much larger than would be expected if the creeping section was a barrier to earthquake rupture.

We finish with some conclusions and suggestions for future directions of research.

Chapter 1

THE IMPACT OF PLATE MOTIONS ON LONG-WAVELENGTH INSAR-DERIVED VELOCITY FIELDS

1.1 Introduction

Interferometric Synthetic Aperture Radar (InSAR) is an active imaging technique for measuring ground displacements that occur between repeat passes of an imaging platform, such as a satellite (e.g., Hanssen (2001), Massonnet and Feigl (1998), and Simons and Rosen (2015)). InSAR deformation measurements are generally expressed relative to a single point, or ensemble of points, within the imaged area, usually assumed to be stable through time (e.g., Mahapatra et al. (2018)). While InSAR has been used extensively for measuring large amplitude (> 1 cm), deformation over short-wavelengths (< 100 km) (e.g., Massonnet et al. (1993) and Merryman Boncori (2019)), other signals present in the data challenge our ability to measure deformation at the scale of millimeters per year over hundreds of kilometers.

InSAR observations at long wavelengths are the combination of motion of the Earth's surface, changes in the atmosphere, and measurement and processing errors. The Earth motion signals comprise the surface deformation of interest, e.g., from tectonic strain, volcanic activity or subsidence (Amelung et al. (1999), Massonnet and Feigl (1998), and Massonnet et al. (1993)), along with solid Earth tides (SET) (X. Xu & Sandwell, 2020), and ocean tidal loading (Dicaprio et al., 2008). Atmospheric signals come from propagation delay through the ionosphere (Z.-W. Xu et al., 2004) and troposphere (Tarayre & Massonnet, 1996). Error sources include the satellite orbits (Massonnet & Feigl, 1998), local oscillator drift (Marinkovic & Larsen, 2015), phase unwrapping (Biggs et al., 2007) and topography (Berardino et al., 2002).

These effects can obscure small amplitude, long-wavelength signals in InSAR due to local tectonic processes, such as surface deformation from interseismic loading (e.g., Fournier et al. (2011) and Parizzi et al. (2021)). Thus, it is common to not interpret long-wavelength signals from InSAR alone, instead removing them by empirically fitting 2D polynomial functions, known as “ramps”, to the data (e.g., Fialko (2006) and Jolivet et al. (2015)), or combining InSAR velocities with Global Navigation Satellite System (GNSS) measurements in order to constrain the long-wavelength deformation (e.g., Neely et al. (2020), Parizzi et al. (2020), Weiss et al.

(2020), and X. Xu et al. (2021)). Such approaches are limiting when we wish to measure large-scale deformation in regions of sparse GNSS coverage (Chaussard et al., 2016; Neely et al., 2020).

The quality of InSAR data and correction methods have substantially increased over the last several years. The European Space Agency's (ESA) Sentinel-1 satellites have been regularly acquiring data for significant portions of the planet since late 2014. Sentinel-1 offers the advantages of improved orbital controls and uncertainties, reducing the noise contribution from satellite orbits (Fattahi & Amelung, 2014), as well as unrestricted data access. Split-band processing now allows for the estimation of the ionospheric signal directly from the InSAR data (Fattahi, Simons, et al., 2017; Gomba et al., 2016; Liang et al., 2019), and higher quality weather models have improved the correction of the tropospheric phase (Doin et al., 2009; Jolivet et al., 2011; Z. Li et al., 2005). Techniques for removing the SET (X. Xu & Sandwell, 2020) and ocean tidal loading signals (Dicaprio et al., 2008; Yu et al., 2020) have also been developed, among other correction methods. After corrections, there may still be long-wavelength residuals in multi-year Sentinel-1 time series, including from the troposphere, which can contribute up to 5 mm/yr over 150 km (Parizzi et al., 2021), and orbital errors, contributing around 0.5 mm/yr over 100 km for Sentinel-1 (Fattahi & Amelung, 2014).

In this work, we focus on the contribution of coherent uniform motion of Earth's tectonic plates to the long-wavelength component of InSAR-derived velocity fields. The satellite line-of-sight (LOS) vector varies systematically in the satellite range direction (i.e., across the satellite track), causing a changing sensitivity to ground deformation with range. Bulk motion of tectonic plates in the satellite frame of reference, coupled with this LOS variation, can create quasi-linear gradients in InSAR-derived velocity fields, resulting in ramps, predominantly in the satellite range direction. This effect has been noted before, e.g., by Bähr et al. (2012), Bähr (2013) and Parizzi et al. (2020). Here, we demonstrate that plate motion creates ramps of several millimeters per year, across the 250 km track width, in six multi-year Sentinel-1 InSAR time series. After other corrections have been applied, plate motion is the dominant long-wavelength signal in our data, and we show that this signal can be straightforwardly compensated for using plate motion models. This adjustment is not currently part of several open-source InSAR time series analysis packages (e.g., Agram et al. (2013), Hooper et al. (2012), Morishita et al. (2020),

and Yunjun et al. (2019)), and we provide an implementation of the method in the MintPy package (Yunjun et al., 2019).

1.2 The Reference Frame of InSAR Measurements

Quantifying ground deformation using InSAR requires a precise measurement of the satellite orbit ephemerides (Fattahi & Amelung, 2014; Peter, 2021). For Sentinel-1 the orbit is measured with respect to the International Terrestrial Reference Frame (ITRF) (Peter, 2021), an Earth-centered, Earth-fixed reference frame in which there is no net rotation of the Earth's surface (Altamimi et al., 2016). Observations of absolute ground motion relative to the satellite are therefore also in ITRF (Bähr et al., 2012; Lazecky & Hooper, 2022).

However, it is not possible for InSAR to record absolute motions due to the 2π ambiguity in the interferometric phase (e.g., Massonnet and Feigl (1998)). Instead, displacement measurements are generally expressed relative to a reference point within the imaged region, assumed to be stationary. Velocities can then be obtained from functional fits to displacement time series, with inferred velocities also expressed relative to this point.

Selecting the reference point is not equivalent to expressing the InSAR velocities in a reference frame moving with that point (Bähr, 2013; Bähr et al., 2012). We must therefore consider how velocities in ITRF appear in the InSAR deformation field. We represent the 3D ITRF secular velocity field of the Earth's surface as:

$$\mathbf{v}(\mathbf{x}) = \mathbf{v}_p(\mathbf{x}) + \mathbf{v}_d(\mathbf{x}). \quad (1.1)$$

$\mathbf{v}_p(\mathbf{x})$ is the velocity field due to the strain-free motion of the relevant rigid plate in ITRF, and $\mathbf{v}_d(\mathbf{x})$ is the velocity due to internal deformation of the plate, for example due to tectonic, volcanic, or hydrological processes.

Defining the LOS unit vector pointing from the ground to the satellite as $\hat{\mathbf{l}}(\mathbf{x})$, the LOS projection of the 3D velocity field, minus the InSAR reference velocity, can be written as:

$$\mathbf{v}_l(\mathbf{x}) = \mathbf{v}(\mathbf{x}) \cdot \hat{\mathbf{l}}(\mathbf{x}) - \mathbf{v}(\mathbf{x}') \cdot \hat{\mathbf{l}}(\mathbf{x}'), \quad (1.2)$$

where the reference is at point $\mathbf{x}'$. $\mathbf{v}_l(\mathbf{x})$ is the secular velocity that will be measured by the satellite, assuming all other signals and noise can be neglected.

In ITRF, $\mathbf{v}_l(\mathbf{x})$ has a contribution from the plate motion, which we can write as:

$$\mathbf{v}_{l,p}(\mathbf{x}) = \mathbf{v}_p(\mathbf{x}) \cdot \hat{\mathbf{l}}(\mathbf{x}) - \mathbf{v}_p(\mathbf{x}') \cdot \hat{\mathbf{l}}(\mathbf{x}'). \quad (1.3)$$

The second term, from the reference, is constant, while the first term depends on the spatial variation of $\mathbf{v}_p(\mathbf{x})$ and $\hat{\mathbf{l}}(\mathbf{x})$. The LOS vector $\hat{\mathbf{l}}(\mathbf{x})$ can vary substantially over an image swath. For Sentinel-1, the incidence angle (the angle between the LOS and the vertical) varies approximately from 29° in the near range to 46° in the far range over the 250 km width of the imaging swath (for data acquired in Interferometric Wideswath mode). The range-dependent variation in $\hat{\mathbf{l}}(\mathbf{x})$ implies a changing sensitivity to components of the 3D deformation field across the track, with sensitivity to horizontal motion increasing and vertical motion decreasing as we move from near range to far range. This range-dependent sensitivity causes uniform plate motions to appear as velocity ramps in the range direction when projected into the satellite LOS (Figure 1.1).

The plate velocity, $\mathbf{v}_p(\mathbf{x})$, also varies over an image swath. The motion of a rigid plate on Earth's surface can be represented by a rotation rate about an axis, known as an Euler pole (McKenzie & Parker, 1967). Given the angular velocity of a chosen plate, $\boldsymbol{\Omega}$, we can write the velocity of any point, $\mathbf{x}$, on that plate as $\mathbf{v}_p(\mathbf{x}) = \boldsymbol{\Omega} \times \mathbf{x}$, where $\times$ is the cross product. Thus, the velocity field due to rigid plate motion varies with distance from the plate's Euler pole. This variation in $\mathbf{v}_p(\mathbf{x})$ also contributes to the long-wavelength LOS velocity field.

Because of the effect of plate motions, InSAR velocity measurements should not generally be considered to be in a local reference frame, despite the use of a local reference point. Choosing a reference point within an InSAR image offsets InSAR velocity measurements from the LOS projection of ITRF velocities by an unknown constant (Equation 1.1), but does not remove the long-wavelength gradients that can be induced by plate motion (Equation 1.3). If ITRF plate motion is negligible when projected to the LOS, and does not vary substantially over the InSAR track, or the satellite LOS variation across the track is small, then $\mathbf{v}_{l,p}(\mathbf{x}) \approx 0$. Choosing a reference point that is stable with respect to the plate is then approximately equivalent to putting the InSAR velocities into the reference frame of that plate; however, this should not be generally assumed.

Several authors have investigated the reference frame of InSAR observations, generally in the context of using GNSS to put InSAR measurements into a terrestrial reference frame (e.g., Johnston et al. (2021) and Mahapatra et al. (2018)). The influence of plate motion on InSAR velocities has been noted by Bähr et al. (2012) and Bähr (2013), who term it the *reference frame effect*. Bähr et al. (2012) present this phenomenon in terms of a temporally increasing correction to the interferometric

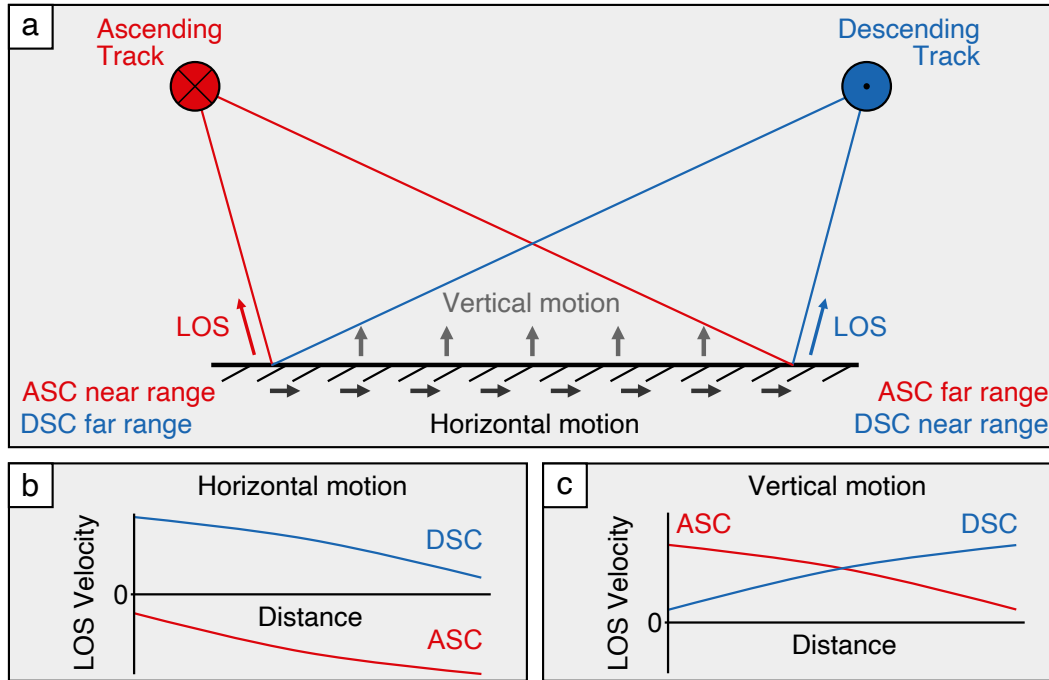


Figure 1.1: Illustration of how uniform horizontal and vertical motions result in ramps in InSAR-derived velocity measurements. **(a)** Satellite images acquired from ascending (ASC) and descending (DSC) orbital tracks, which have a varying LOS incidence angle across the track. The term “range” refers to the distance from the ground target to the satellite, with near range and far range the closest and furthest points from the satellite, respectively. Red and blue arrows represent the ground-to-satellite LOS vector in the near range for ASC and DSC tracks, respectively. Grey and black arrows represent plate motion in the reference frame of the satellite. For illustration purposes we assume the ASC and DSC tracks are parallel to each other but in opposite directions, and ignore Earth curvature. Figure not to scale. **(b)** Profile of the horizontal plate velocity projected into the LOS of the ASC and DSC tracks, against geographic distance along the ground. **(c)** Same as (b), except for vertical plate motion, resulting in opposite gradients in the LOS profiles. The observing geometry creates a small curvature in all profiles, which is exaggerated in the figure. For InSAR measurements, the LOS velocity is expressed relative to a point within the image, so each of these profiles would be vertically shifted to intersect with the x axis at the chosen reference point.

baseline, while Bähr (2013) notes that this can also be framed in terms of the varying satellite LOS causing differing sensitivity to plate motion (the approach taken here). Parizzi et al. (2020) used plate motion models to adjust their LOS velocity fields after merging InSAR with GNSS. Authors have also noted the impact of plate motions on SAR geolocation accuracy (Cong et al. (2012)). Our focus here is to demonstrate that plate motions can explain a significant fraction of observed residual long-wavelength surface velocities, after other corrections have been applied, and without combining InSAR data with GNSS.

1.3 Data and Methods

1.3.1 Data Processing

We present several examples using InSAR data from the ESA’s Sentinel-1 satellites, taken from ascending (ASC) and descending (DSC) tracks covering the Makran subduction zone (Iran), the Gulf of Aqaba (at the northern end of the Red Sea), and western Australia. For each track, we process at least 5 years of data using the InSAR Scientific Computing Environment (ISCE) (Fattahi, Agram, et al., 2017; Rosen et al., 2012). After forming the interferogram networks, we create deformation time series using MintPy (Yunjun et al., 2019).

Before examining residual signals due to plate motion, we apply corrections for the ionosphere, troposphere, SET, and digital elevation model (DEM) error. We use split-band processing to correct for the ionosphere (Liang et al., 2019), PyAPS and the ERA5 weather model to mitigate the tropospheric delay (Hersbach et al., 2020; Jolivet, Agram, et al., 2014), the method of Fattahi and Amelung (2013) for DEM error correction and PySolid to correct for SET (Milbert, 2018; Yunjun et al., 2022). Further details of our data and processing are presented in Appendix A (Sections A.2, A.3, and Table A.1).

1.3.2 Adjusting InSAR Measurements for Plate Motion

After all other corrections have been applied, we can then observe and account for the signal of plate motion. InSAR observations of ground motion are generally used to study regional deformation, rather than plate translations or rotations. For such purposes, a useful reference frame is one that moves with the plate in which we are trying to measure strain. Translating into this reference frame requires us to remove the signal of plate motion in the satellite’s frame of reference, i.e., ITRF (Bähr, 2013; Parizzi et al., 2020).

GNSS networks can be used to connect InSAR measurements to ITRF (e.g., Johnston et al. (2021) and Mahapatra et al. (2018)), which can then be transformed into a reference frame moving with the chosen plate. In the absence of sufficient GNSS coverage, we can estimate the transformation into the plate’s frame of reference using the following steps:

1. Choose an InSAR reference point, $\mathbf{x}'$, that is stable with respect to the plate
2. Find the velocity field of the plate within ITRF, i.e., $\mathbf{v}_p(\mathbf{x})$
3. Project that velocity field into the satellite LOS direction
4. Subtract the LOS velocity of the reference point, $\mathbf{v}_{l,p}(\mathbf{x}')$, from the projected plate velocity to compute $\mathbf{v}_{l,p}(\mathbf{x})$, which is then removed from the InSAR velocity map.

Note that, after these steps, InSAR-derived velocities are still expressed relative to a reference point, meaning that deformation and other signals seen at the reference point will still affect the entire scene.

We use the geodetically constrained ITRF plate motion model of Altamimi et al. (2017) to estimate the plate velocity field. For each study region, we identify our reference plate (Table A.1), then use the modeled angular velocity of the plate to calculate horizontal velocities for our observation region. We then project these velocities into the LOS direction and remove them from the velocity map.

1.4 Results

1.4.1 The Importance of Removing Other Signals for Revealing Plate Motion

We expect plate motion to contribute below 8 mm/yr across the 250 km width of the Sentinel-1 tracks for our chosen regions (Figure A.6), making it important to remove other signals to show what fraction of the residual velocity can be explained by plate motion. For ASC tracks in the Makran and Gulf of Aqaba, ionosphere corrections have a particularly large effect on the long-wavelength velocity signal (e.g., contributing a 25 mm/yr ramp along track 86 for the Makran, Figure 1.2), with DSC tracks showing substantially less ionospheric signal. ASC tracks are acquired at dusk—a period of greater ionosphere activity than dawn, when DSC tracks are acquired. This impact is still notable in C-band Sentinel-1 data, even though it suffers much less from ionospheric effects than L-band (Fattahi, Simons, et al., 2017; Liang et al., 2019).

We find that troposphere corrections have a less significant impact on the long-wavelength velocity signal than the ionosphere for ASC tracks, and a comparable effect for DSC tracks. Corrections for the SET have a small effect on the long-wavelength secular velocity, contributing below 0.5 mm/yr over several hundred kilometers. The range of DEM error corrections is less than ± 0.5 mm/yr in our results and has a minimal contribution to the long-wavelength velocity field. We show the impact of the above corrections for all tracks in Figures 1.2 and A.1-A.5, and present more details in Section A.3.

1.4.2 The Impact of Accounting for Plate Motion

After applying the suite of corrections we are left with residual velocity ramps in all of our tracks, predominantly in the range direction. We present the results of plate motion adjustments for several tracks in Figures 1.3 and 1.4. Our results show that accounting for plate motion removes a significant fraction of the residual velocity ramp in every case, reducing the across-track ramps from 4-7 mm/yr/track to below 1.5 mm/yr/track. For our data the plate motion signal is comparable to the troposphere in its effect on the long-wavelength velocity field.

The proximity of the Arabian plate Euler pole to the Gulf of Aqaba study area results in the plate velocity field varying appreciably within the tracks (Altamimi et al., 2017). This variation causes an additional LOS velocity ramp along the track, with an opposite direction for the ascending and descending tracks. Figures 1.4 (a) and (b) show how this along-track gradient can be clearly seen in the data, and is well corrected for by the plate motion model. We do not see similar along-track ramps for Australia and Makran, which is consistent with the plate motion velocity field.

1.5 Discussion

Other authors have previously noted that plate motion will affect InSAR velocity measurements (e.g., Bähr et al. (2012)). However, the narrower variation of the satellite LOS angle for earlier satellites, more limited data, and the presence of other significant long-wavelength signals, has made the signal difficult to isolate. The quality of recently available data and correction methods, and the wide swath of Sentinel-1, allow us to show the plate motion signal and the clear impact of accounting for it. Our results from the Gulf of Aqaba illustrate that plate rotation is an important part of the correction.

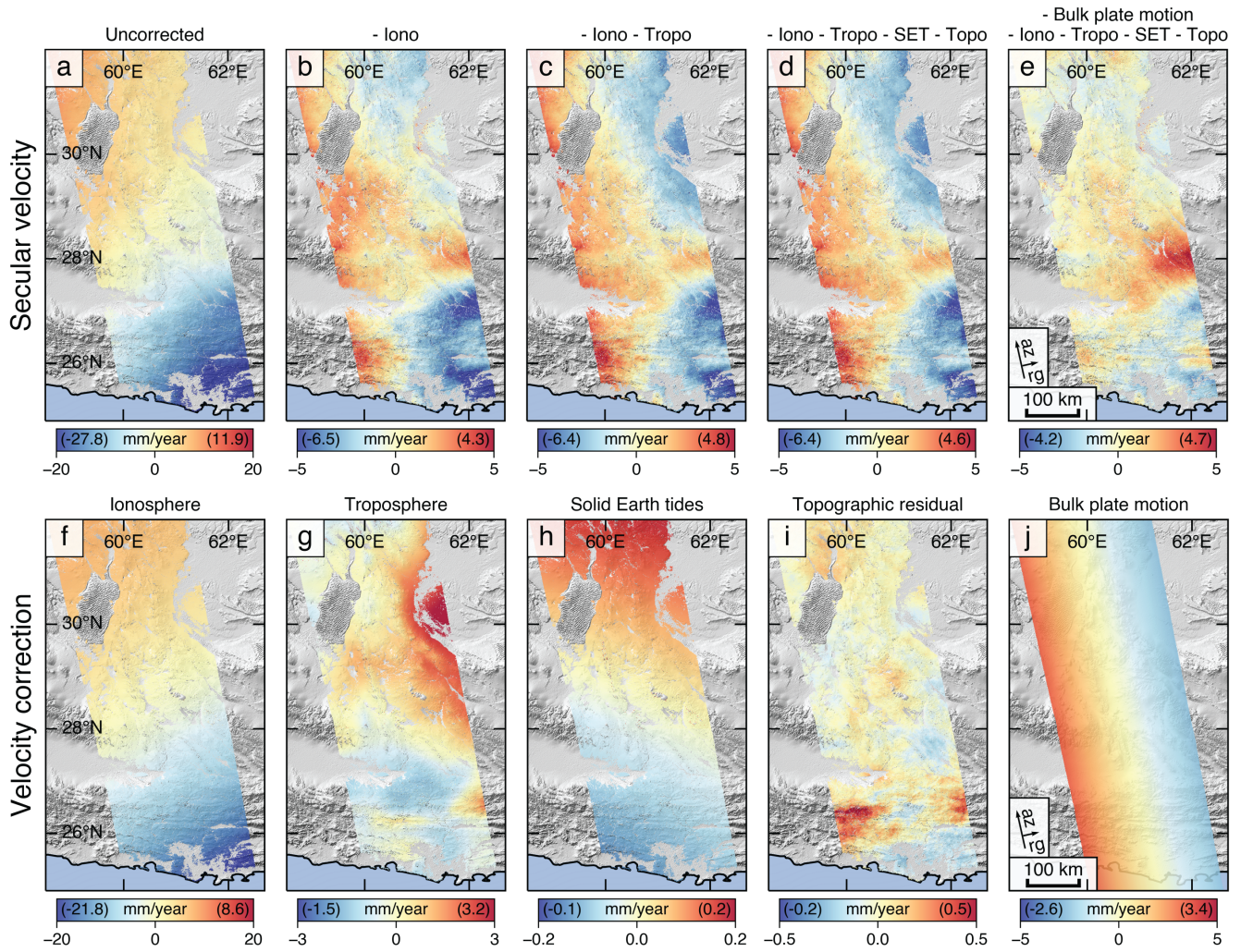


Figure 1.2: Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 86 (ASC) over the Makran subduction zone. For plotting purposes, we remove the median value from each velocity field. Positive values represent apparent motion towards the satellite. Color bars are re-scaled between plots. Numbers in parentheses within the color bars refer to the 2nd and 98th percentiles of the velocity. “az” is the azimuth direction (satellite direction of motion), and “rg” the range direction (perpendicular to the satellite direction of motion). (a) No corrections applied. (b) Estimated ionosphere removed. (c) Tropospheric model removed. (d) SET model removed and DEM error (Topo) correction applied. (e) Plate motion correction applied. The positive signal around (28 °N, 62 °E) is post-seismic deformation from the 2013 Khash earthquake (Barnhart, Hayes, Samsonov, et al., 2014). (f) Applied ionospheric correction. (g) Applied tropospheric correction. (h) Applied SET correction. (i) Applied DEM error correction. Larger signals in the south may be bias from tropospheric residuals (Fattahi & Amelung, 2015). (j) Applied plate motion correction.

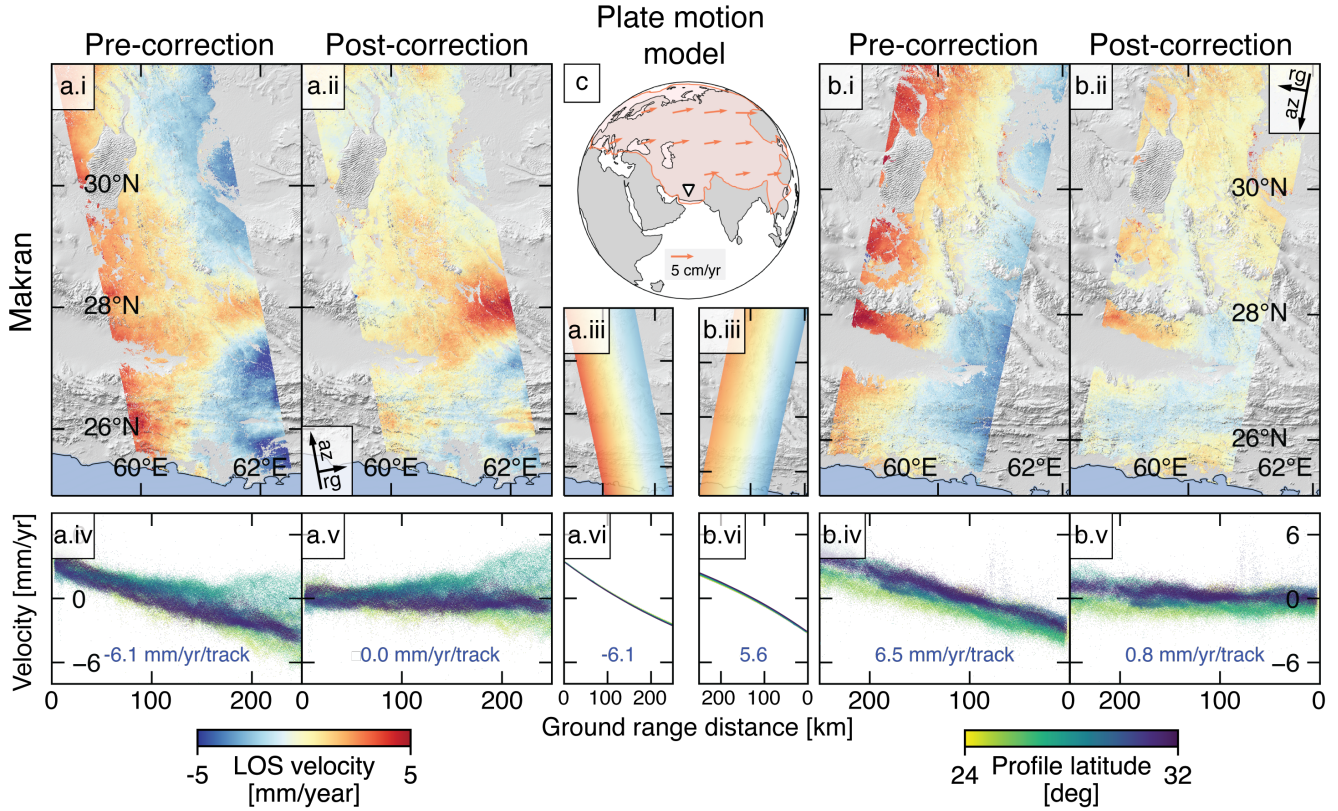


Figure 1.3: The impact of plate motion adjustments for InSAR tracks over the Makran subduction zone. For plotting purposes, we remove the median from each velocity field. **(a)** track 86 (ASC). After plate motion correction, the post-seismic signal from the 2013 Khash earthquake (Barnhart, Hayes, Samsonov, et al., 2014) can be more clearly seen in the south-east of the figure. **(b)** track 20 (DSC). **(c)** Location of tracks (a) and (b) and the velocity field of the Eurasian plate, used to correct the tracks. **(i)** Velocity before plate motion correction, but after other corrections have been applied. **(ii)** Velocity after plate motion correction. **(iii)** Applied plate motion correction. **(iv)** Across-track profile of the velocity before plate motion correction. The number below each profile is the gradient of the linear least squares fit to the profile. Note that profiles are plotted as a function of ground range, which increases with distance from the satellite. **(v)** Across-track profile of the velocity after plate motion correction. **(vi)** Across-track profile of the applied plate motion correction.

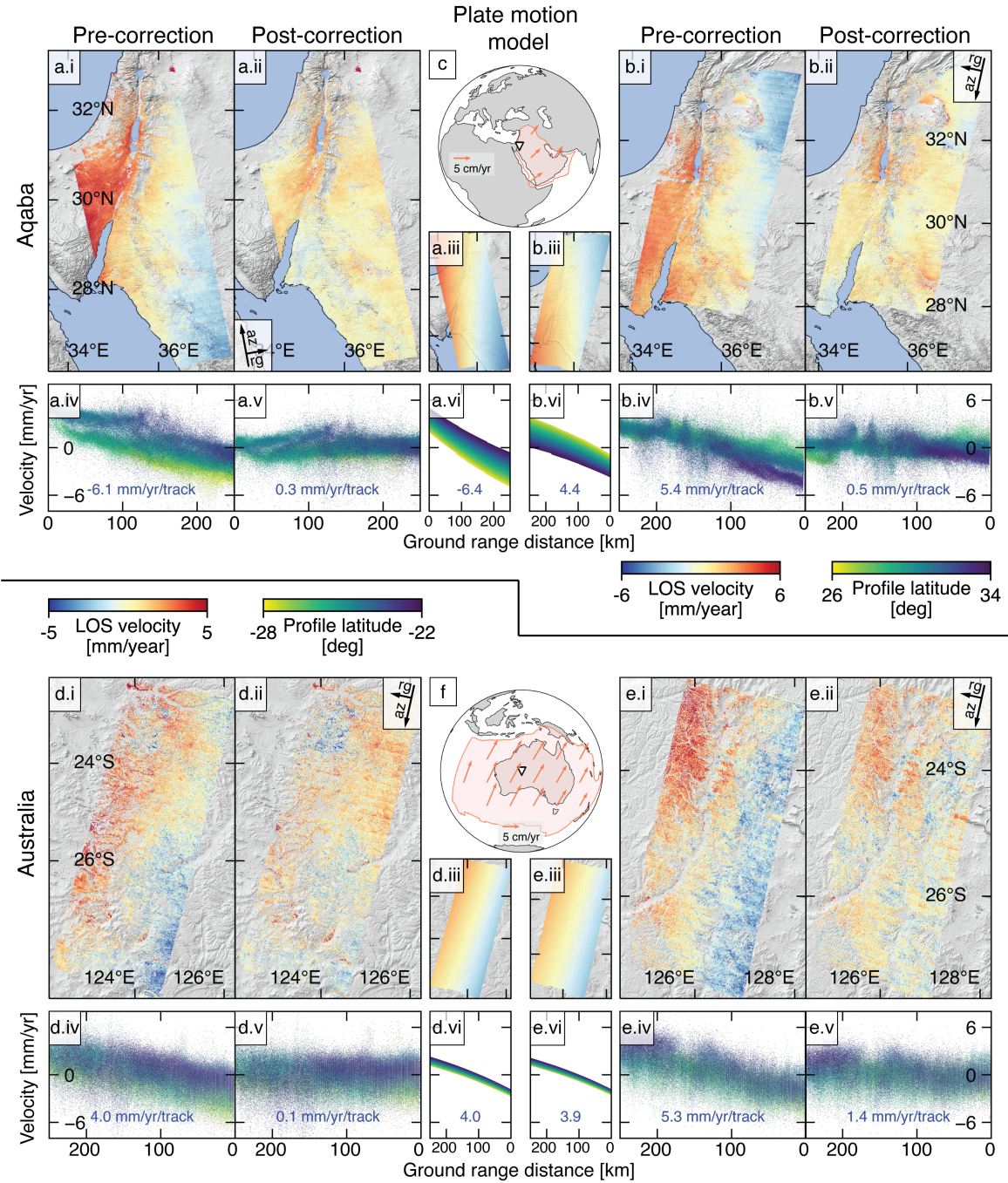


Figure 1.4: Same as Figure 1.3, but showing the tracks for the Gulf of Aqaba and Australia. **(a)** Gulf of Aqaba, track 87 (ASC). **(b)** Gulf of Aqaba, track 21 (DSC). **(c)** Location of tracks (a) and (b) and the velocity field of the Arabian plate, used to correct the tracks. **(d)** Australia, track 119 (DSC). **(e)** Australia, track 46 (DSC). **(f)** Location of tracks (d) and (e) and the velocity field of the Australian plate, used to correct the tracks. **(i)-(vi)** are as described in Figure 1.3.

After adjusting for plate motion, remaining long-wavelength signals could be due to the incomplete removal of some signals (predominantly the troposphere (Fattahi & Amelung, 2015; Parizzi et al., 2021)), sources that we have not corrected for (e.g., ocean tidal loading (Dicaprio et al., 2008) and orbital errors (Fattahi & Amelung, 2014)), or actual strain accumulation in the lithosphere—the signal that InSAR measurements often target. See Section A.3 for more details on the contributors to long-wavelength residuals.

Deficiencies in the plate motion model, or motion of the InSAR track reference point relative to the assumed plate, could also create long-wavelength residuals. Motion relative to the plate will be of particular importance in areas of diffuse plate boundary deformation, where it is not possible to choose a reference point that is stable with respect to the rigid plate. This could be the case for tracks covering the Makran subduction zone and the Gulf of Aqaba, both of which span plate boundary zones. In these situations, plate motion models may not fully account for the impacts of bulk motion, and using local GNSS measurements to put InSAR measurements into a local terrestrial reference frame could be necessary (Bähr, 2013).

These results emphasise the importance of accounting for the reference frame before interpreting long-wavelength InSAR-derived velocity fields. When using InSAR for studies of tectonic deformation, the most natural reference frame is one that is fixed to a stable region within the scene, so that we can interpret velocity gradients in terms of tectonic strain rather than strain-free translation and rotation. There are several situations in which failing to account for the reference frame could bias the results:

1. Combining multiple tracks to estimate 3D deformation (Fialko et al., 2001; Wright et al., 2004). In Section A.5 and Figure A.7, we show how plate motion can bias estimates of the 3D velocity field when we use an overlapping ascending and descending track to calculate horizontal and vertical velocities.
2. Modeling InSAR signals. If the long-wavelength signals in an InSAR velocity field are being modeled, and the model is assumed to not be rotating or translating, then a velocity ramp from plate motion may be modeled as strain accumulation and bias the results (e.g., changing the locking depth in a subduction zone model).
3. Comparisons between GNSS and InSAR. Both data sets must be in the same reference frame (Parizzi et al., 2020). If the GNSS are in a local reference

frame, the InSAR and GNSS velocities will diverge at long wavelengths due to the signal of plate motion in the InSAR.

Studies which removed ramps from InSAR-derived velocities to account for orbital errors may have inadvertently removed the impact of plate motion in their observations as well, reducing the biases we outlined above.

In this work, our primary focus is on the impact of horizontal plate motions, and we have not considered the contribution of long-wavelength vertical velocities. Horizontal plate motions in ITRF are generally at the scale of centimeters per year (Altamimi et al., 2017), with long-wavelength vertical motions, for example due to post-glacial rebound, significantly smaller at millimeters per year (e.g., Lau et al. (2020) and Riddell et al. (2020)). If an InSAR track is taken within a region that is experiencing constant vertical motion, this motion will also create a velocity ramp in the satellite LOS velocity field, but with ASC and DSC tracks having opposite gradients (Figure 1.1(c)). However, the amplitude of vertical velocities will result in smaller velocity gradients across the satellite track than those caused by horizontal motion (Section A.4, Figure A.6).

1.6 Conclusions

We have illustrated how InSAR velocity measurements are sensitive to tectonic plate motion in the satellite reference frame. This motion will induce ramps in the InSAR velocity fields, predominantly in the satellite range direction, of up to several millimeters per year. In all of our multi-year time series, plate motion was the dominant long-wavelength signal after ionospheric and tropospheric corrections were applied. We have presented a simple adjustment method, which uses plate motion models to remove the plate motion signal from the InSAR velocity field. This adjustment substantially reduces long-wavelength ramps in multiple InSAR tracks from three different regions of the Earth. Routinely accounting for plate motion in InSAR could reduce biases when constraining long-wavelength tectonic strain induced by local geophysical phenomena. This adjustment is likely to be particularly useful where GNSS is not available to constrain the long-wavelength deformation. The signal of plate motion in InSAR data could also be used to improve plate motion models, which may be helpful where GNSS observations are sparse but high-quality InSAR data are available.

Chapter 2

IMAGING THE MAKRAN SUBDUCTION ZONE WITH DENSE INSAR TIME SERIES

2.1 Introduction

Interferometric synthetic aperture radar (InSAR) is increasingly being used to measure small deformation signals over wide areas, generally with the aid of the Global Navigation Satellite Systems (GNSS), such as the Global Positioning System (GPS), to help constrain the long-wavelength deformation (e.g., Weiss et al. (2020) and X. Xu et al. (2021), and see Chapter 1). There are many obfuscating contributors to InSAR observations that must be accounted for in order to maximize our ability to discern long-wavelength, small-amplitude deformation without introducing GNSS. The improvements in InSAR correction methods, including corrections for plate motion we presented in Chapter 1, as well as the dramatic increase in the amount of data available in the last five years, give us the opportunity to test the ability of InSAR alone to constrain very long-wavelength deformation.

In this chapter, we test the ability of temporally-dense InSAR time series to constrain deformation in the Makran subduction zone, an enigmatic megathrust that is one of the least studied on the planet. The Makran subduction zone accommodates motion of the Arabian plate under the Eurasian plate at a rate of about 3 cm/yr, over a 1000 km long fault along the southern coasts of Iran and Pakistan. Above the megathrust lies a large accretionary prism, the sub-aerial portion of which is one of the largest in the world (Fruehn et al., 1997). To the east, the subduction zone is bounded by the left-lateral Chaman-Ornach-Nal fault, accommodating motion between the Indian and Eurasian plates, while in the west the Minab-Zendan-Palami fault zone marks the transition region between the subduction zone and the Zagros collision belt (Figure 2.1, and see Nemati (2019)).

The eastern end of the subduction zone has hosted several large events, including in 1765, 1851, and most recently a M_w 8.1 earthquake and associated tsunami in 1945 (Byrne et al., 1992). The number of casualties for this event is often cited as 4000; however, this number is based on early estimates, with later numbers closer to 300 (Hoffmann et al., 2013). Based on historic seismicity, the subduction zone seems to be bifurcated, with large ruptures limited to the eastern end and no

major ruptures conclusively attributed to the western end of the subduction zone in historical records (Byrne et al., 1992; Musson, 2009). However, paleotsunami evidence suggests that the western Makran may have hosted significant earthquakes, including in 1008 and 1542 (e.g., Hoffmann et al. (2020), Prizomwala et al. (2018), Rajendran et al. (2021), and Shah-hosseini et al. (2011)).

Previous efforts using geodesy to constrain potential coupling on the megathrust include use of GNSS (Frohling & Szeliga, 2016; Khorrami et al., 2019; Penney et al., 2017) and InSAR (Lin et al., 2015). These studies generally find significant coupling on the megathrust at both the eastern and western ends, albeit with substantial uncertainty due to the limited availability of geodetic data. Results from thermal modeling of the subduction zone have also indicated the potential for a wide seismogenic zone (Khaledzadeh & Ghods, 2022; Smith et al., 2013). These studies combine a calculated temperature distribution with assumptions about the temperature range over which subduction zones are seismogenic to infer that the Makran could host earthquakes as large as M_w 9 for ruptures along the entire length of the fault. Such earthquakes, and associated tsunamis, present a substantial danger to the growing population bordering the Arabian sea (e.g., (Qiu et al., 2022; Salah et al., 2021)). It is therefore important to better understand the behavior of the megathrust and constrain the potential size of future ruptures.

In our work, we used data from the European Space Agency’s Sentinel-1 satellites to study an area of nearly one million square kilometers over the Makran subduction zone. We use images from six ascending and six descending tracks, totalling over 2000 acquisitions and spanning 7.5 years (from October 2014 to April 2022) to examine the ability of InSAR data to measure deformation over wide areas, without the aid of GNSS. We apply corrections, including for the ionosphere, troposphere, solid Earth tides, DEM error and plate motion, and show their importance for observing small signals over long wavelengths.

The resulting InSAR-derived velocity field allows us to observe numerous deformation processes, including post-seismic deformation in the accretionary prism from the 2013 M_w 6.1 Minab and M_w 7.7 Balochistan earthquakes (Avouac et al., 2014; Jolivet, Duputel, et al., 2014; Penney et al., 2015), and post-seismic deformation from the 2013 M_w 7.7 Khash earthquake, an intraslab earthquake that occurred at 80 km depth (Barnhart, Hayes, Briggs, et al., 2014). We also measure creep and coseismic offsets along the Chaman fault (Dalaison et al., 2021) and rapid subsidence due to aquifer depletion (Motagh et al., 2017).

The use of ascending and descending tracks provides two components of the deformation field improving our ability to discern signals of locking on the megathrust. Post-seismic deformation at the eastern end of the subduction zone, and high tropospheric noise, do not allow us to unambiguously resolve such a signal. However, by comparing our velocity measurements with simple forward models, we can place limits on the degree of coupling on the fault. Our results suggest that scenarios where the western end of the megathrust is strongly coupled north of the coast are unlikely to be consistent with InSAR observations, under the assumptions of our simple linear elastic forward model. Coupling may therefore be less strong than some previous studies have suggested, although further work is needed to robustly constrain coupling on the megathrust.

2.2 Data and Methods

2.2.1 Interferogram Processing

We begin with Sentinel-1 Single Look Complex (SLC) data from the Alaska Satellite Facility. We use all available data from the beginning of Sentinel-1 acquisitions in October 2014 until early April 2022, resulting in approximately 7.5 years of data in each track. We process 12 tracks in total, six ascending (ASC) and six descending (DSC) from the coast up to 32°N . The details of each track are shown in Table 2.1.

We process the SLC data using the InSAR Scientific Computing Environment (ISCE) (Fattahi, Agram, et al., 2017; Rosen et al., 2012). All tracks are processed using the topsStack processing chain. We take 5 looks in azimuth and 20 looks in range before unwrapping, but do not apply filtering. Each acquisition is used to form interferograms with the next three acquisitions, resulting in roughly 500-550 interferograms per track (although the combination of Sentinel-1 A/B pairs for track 159 allows us to form 786 interferograms). We unwrap each full interferogram (from the coast to 32°N), at once, rather than splitting the track into frames. Details of unwrapping, ionosphere corrections, azimuth misregistration estimation and phase unwrapping are as described in Chapter 1, with some modifications described below. Note that, as we use the topsStack code, we do not take account of burst discontinuities introduced by the ionosphere, as discussed in Appendix A and Liang et al. (2019).

As a brief recap, ionosphere corrections rely on the dispersive nature of the ionosphere, allowing us to form interferograms at higher and lower frequency bands and estimate the ionospheric phase screen from the difference between the two interfer-

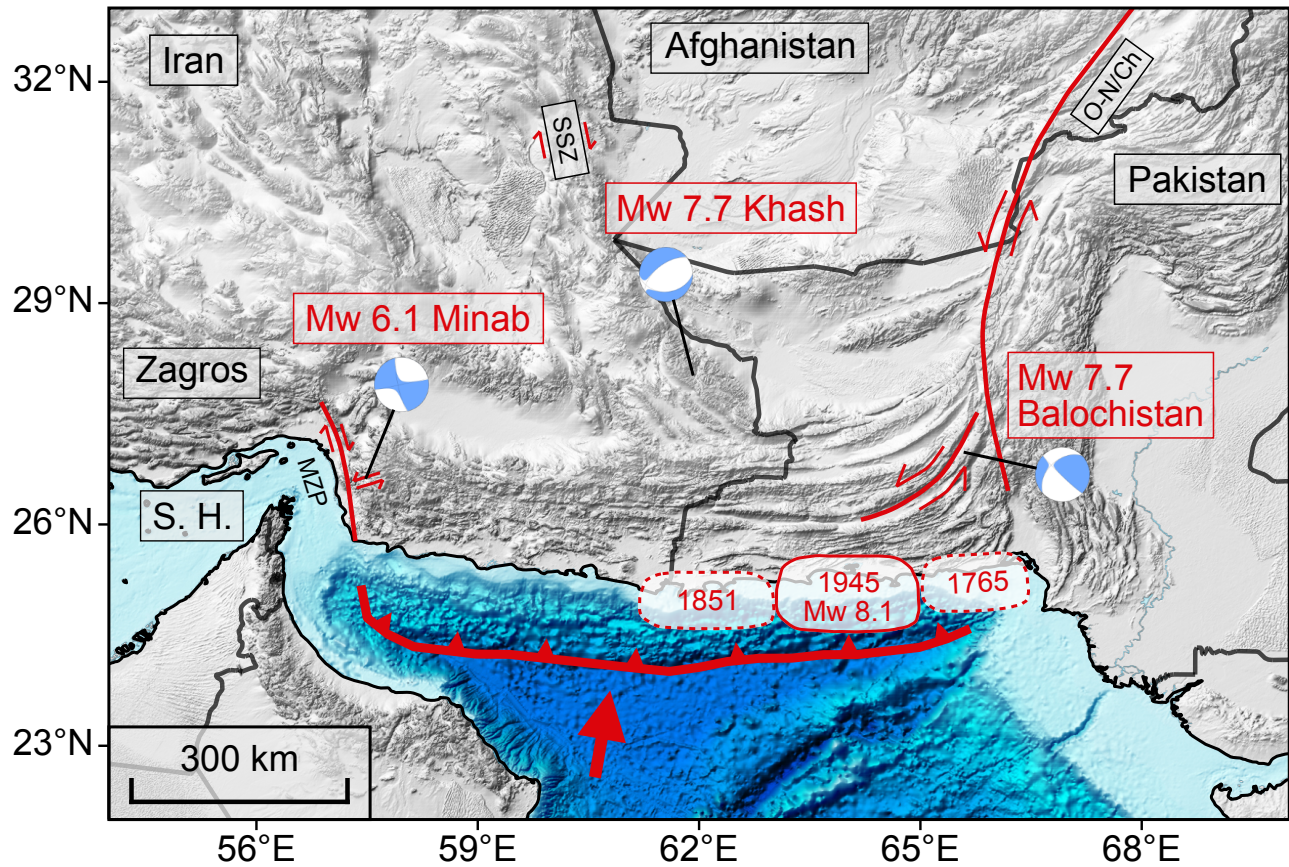


Figure 2.1: Map of the Makran region. Earthquakes and faults are shown in red, coastlines and borders are shown in black. The large red arrow at 23°N indicated the direction of convergence of the Arabian plate under the Eurasian plate used in this study. S.H.: Strait of Hormuz. MZP: Minab-Zendan-Palami fault zone. SSZ: Sistan Suture Zone. O-N/Ch: Ornach-Nal/Chaman fault zone. Approximate locations of ruptures on the megathrust in 1765, 1851, and 1945 are taken from Byrne et al. (1992). Focal mechanisms for three earthquakes in 2013 (Minab, Khash, and Balochistan) are shown, along with the rupture trace for the Hoshab fault, which hosted the Balochistan earthquake.

ograms Liang et al., 2019. Forming sub-band interferograms increases the noise in each interferogram, making it necessary to strongly filter the interferograms before unwrapping (unwrapping refers to removing the 2π phase ambiguity inherent to interferometry Chen and Zebker, 2002).

After unwrapping and calculating the ionospheric phase screen we inspect all ionosphere pairs and remove the small number of pairs that have large unwrapping errors. For tracks 57, 93, and 166, the tracks contain land on the southern side of the Strait of Hormuz. These disconnected land masses result in two areas of the ionospheric phase screen that are correctly unwrapped; however, they have an unknown offset between them as interferograms cannot be unwrapped across water. These separate areas are known as disconnected components, and can result in biases when computing the ionospheric phase that dominate the overall error budget. We therefore mask out the disconnected areas before calculating the ionosphere for these tracks.

The same ionosphere issue arises whenever we have disconnected components in the unwrapped ionospheric phase estimate (Liang et al., 2019). Such unwrapping errors can be caused by a small area of low noise (high coherence) surrounded by a narrow area of high noise (low coherence), which is in turn surrounded by high coherence (for more details on coherence, see Chapter 3). The high coherence areas unwrap individually; however, the low coherence zone between them can result in an offset between the two areas in a similar fashion to the water discussed above. Such unwrapping errors in small areas do not present significant issues for normal interferogram processing, as we can just mask out the low coherence areas and the small high coherence areas after time series processing. However, the need to strongly filter the ionosphere can cause large unwrapping errors to bleed in to adjacent areas in the ionospheric phase.

An example of how these errors affect the ionospheric velocity correction can be seen in Figure 2.2. The error here is caused by sand dunes, which have low coherence in interferograms due to their rapid change in surface properties, surrounding a small patch of higher coherence which is not connected (via higher coherence paths) to the rest of the scene. These unwrapping errors affect a significant number of our ionospheric phase estimates and result in a large error in the ionospheric correction to the InSAR velocity. In this case, the error is greater than 5 mm/yr over more than 2,500 km², shown in Figure 2.2(a). The size of the area affected is related to the width of the spatial filter, the extent of the low coherence area, and the size of the isolated high coherence area. For narrower filters, wider low coherence areas, and

smaller high coherence areas, we would not expect to see the error extending in to the surrounding region.

Such unwrapping errors could be corrected in a similar fashion to the masking we used to address errors across the Strait of Hormuz, described above. Unfortunately, unlike with large landmasses separated by water, identifying the locations of ionosphere unwrapping errors caused by small disconnected patches of high coherence is challenging in advance of computing the ionospheric phase screen for a large number of pairs. Such corrections require processing large volumes of data, visually inspecting the processed data, and then hand-selecting the areas to be masked out before reprocessing the data, taking additional time and computational resources. In the future, automatically dealing with such unwrapping errors in the ionospheric phase screen will be important for reliable ionospheric corrections when processing large volumes of data.

Unwrapping errors affect a few specific regions of our study area, but we also find that the descending ionosphere corrections seem to add a subtle velocity ramp, with each track showing an approximately 0.5 mm/yr/track gradient in the satellite range direction. It is unclear what the origin of these ramps is; however, the fact that they appear to be very similar for every descending track suggests that they may be an artefact of the processing, and not a genuine feature of the ionosphere. For the descending tracks the expected ionospheric signal is small (e.g., Liang et al. (2019)), and as we see several errors due to unwrapping, as well as range ramps, we therefore do not apply the ionosphere corrections in the descending track for the final velocity calculations. Ascending tracks have much more substantial ionosphere, causing the actual ionospheric signal to be far greater than the errors and making ionospheric corrections necessary. The size of the ascending ionospheric signal results in it being hard to see if the ionospheric corrections have a systematic range-dependent ramp, as we see for the descending tracks. We do find one region with particularly prominent unwrapping errors at the southern end of ascending track 42, which we mask in the final velocity map. We use the temporal coherence of the ionosphere time series to define the mask, resulting in a smooth masking boundary that can be seen in the bottom right of Figure 2.11 (Yunjun et al., 2019). Other areas may also have errors due to issues with ionosphere unwrapping, although we have not systematically explored this contribution to the error budget.

Descending track 122 contains unwrapping errors in the interferogram phase related to low coherence areas around (29.7°N, 62°E). We deal with these by applying the

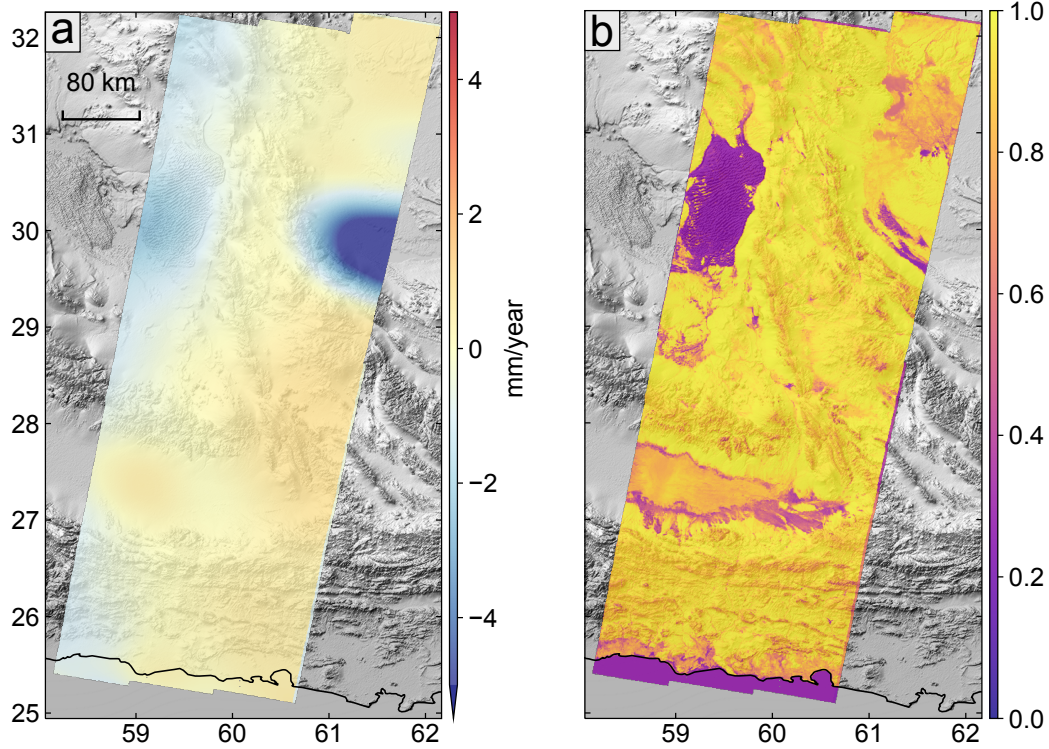


Figure 2.2: Example of errors in the ionosphere calculation **(a)** Velocity from a linear fit to the ionosphere correction time series for track 20 (descending). Note the large negative velocity around (30°N, 61.5 °E) as a result of unwrapping errors. There is also a subtle velocity ramp, of about 0.5 mm/yr, in the range direction, which is also seen in the ionosphere of other descending tracks **(b)** Average spatial coherence for all interferograms for track 20. Low coherence areas around (30°N, 61.5 °E) result in unwrapping errors which bias the ionospheric phase estimation. Filtering of the ionospheric phase then causes this error to affect surrounding areas, as can be seen in (a).

bringing unwrapping error correction method described by Yunjun et al. (2019). This method substantially reduces the number of pairs with unwrapping errors, but does not eliminate them entirely. We therefore visually inspect the interferograms and remove the remaining pairs with obvious unwrapping errors.

Ascending track 159 has a roughly year-long gap in acquisitions from March 2017 to March 2018, after which the addition of Sentinel-1B acquisitions gives a 6-day repeat period. As described by Liang et al. (2019) (Section III.F), ionospheric phase estimation for Sentinel-1 A/B pairs requires removing an empirical ramp from the phase. The error introduced by this ramp is unclear, so when we form pairs for ionospheric phase estimation we limit the number of cross A/B pairs. Specifically

we form a small number of cross A/B pairs at the beginning of the Sentinel-1 B acquisitions (March 2018) and at the end (December 2021) to ensure the network is fully connected, but other than that only form A/A and A/B pairs for ionosphere estimation. We have not systematically explored how network choice affects the quality of the final ionosphere estimation.

An issue also arises when estimating the ionospheric phase for pairs of acquisitions that have different starting ranges, which can result in phase discontinuities between adjacent subswaths (Liang et al. (2019), Section III.E). These errors can be estimated and removed empirically; however, this potentially introduces additional errors into the ionospheric phase estimation. To minimise the number of pairs with different starting ranges used in the processing, we exclude any groups of acquisitions with the same starting range that have fewer than 10 acquisitions before forming interferogram networks.

2.2.2 Time Series Processing

Time series processing is also as described in Chapter 1 (using the SBAS methodology (Berardino et al., 2002), implemented in MintPy (Yunjun et al., 2019)), with one modification to deal with outliers, described below. After inverting the interferogram network and applying ionosphere, troposphere and solid Earth tides corrections in the temporal domain, we fit a linear function to the time series and calculate the root mean square of the residuals, using a temporal coherence threshold of 0.7 to mask unreliable pixels (Yunjun et al., 2019). (Unless otherwise stated, a temporal coherence threshold of 0.7 is used for masking in all velocity maps). We then compute the median absolute deviations (MAD) of the residual RMS as described in Yunjun et al. (2019), and exclude dates that have an RMS value of more than 2 MAD when estimating the DEM error (Fattahi & Amelung, 2013). This approach results in less bias from large residual tropospheric signals in the estimation of the DEM error. When estimating the final linear velocity, after removing the DEM error, we similarly exclude 2 MAD outliers to reduce the bias from the noisiest scenes.

All processing up to this point is performed in radar coordinates. After calculating the linear velocity for each track, we geocode the velocity on to a grid of 0.007 by 0.007 degrees, resulting in pixel sizes of approximately 780 meters in the latitude direction, and a variable width in the longitude direction, from 710 meters at 24°N to 660 meters at 32°N. This additional averaging reduces spatial decorrelation noise.

We then apply the plate motion correction in the geocoded domain, as described in Chapter 1. For plate motion removal we use the Eurasian plate from the ITRF2014 plate motion model (Altamimi et al., 2017).

After applying all corrections we combine the velocity fields of adjacent tracks by calculating an offset between the means of the overlap between the tracks, similar to the approach taken by Fattahi and Amelung (2016). Note that this approach ignores the difference in the satellite line-of-sight (LOS) vector in the overlapping region (Shirzaei, 2015). As we will show in Section 2.3.2, this approach to merging will create substantial biases in the final velocity field when we do not account for plate motion.

We perform the merging separately for ascending and descending tracks. We then take profiles through the merged ascending and descending track velocity fields. When taking profiles we mask out any point that is not correctly unwrapped in every interferogram (i.e., areas identified as unreliably unwrapped by the SNAPHU algorithm (Chen & Zebker, 2002)). This masking approach removes pixels where coherence is more variable, particularly areas covered by sand or agriculture, and is more aggressive than a temporal coherence threshold of 0.7. Using this mask allows us to focus on the more reliable pixels, at the expense of losing some spatial coverage.

As well as calculating a velocity from our InSAR time series, we also wish to estimate the uncertainty of this velocity. Under the assumption that residuals are uncorrelated in time and normally distributed, the velocity uncertainty for each pixel is the standard error of linear regression (e.g., Equation 10 of Fattahi and Amelung (2015)):

$$\sigma_v = \sqrt{\frac{\sum_{i=1}^N (d_i - \hat{d}_i)^2}{(N - 2) \sum_{i=1}^N (t_i - \bar{t})^2}}, \quad (2.1)$$

where N is the number of acquisitions used in the linear regression, d_i is the InSAR observation at epoch i , $\hat{d}_i$ is the predicted value at i from the linear fit, t_i is the time at i and $\bar{t}$ is the mean of all t_i values. This quantity does not capture the uncertainty due to temporally correlated noise, or systematic biases in the time series (Fattahi & Amelung, 2015). It also does not give us the covariance between different pixels, although we would expect residual troposphere to lead to a spatial covariance in the velocity field (e.g., Lohman and Simons (2005)).

Table 2.1: Summary of Sentinel-1 SAR Data used. ASC: Ascending track. DSC: Descending track. Aqn. no.: Number of acquisitions used for interferogram formation. Excl. no.: Number of acquisitions excluded using the 2 median absolute deviation threshold for velocity calculation. All tracks use Sentinel-1A acquisitions apart from track 159, which combines 1A and 1B. Tracks are listed from west to east. Each track is 250 km wide and adjacent tracks have a roughly 50 km overlap.

Track	Direction	Start date	End date	Aqn. no.	Excl. no.
57	ASC	20141021	20220325	185	55
159	ASC	20141016	20220401	264	69
86	ASC	20141023	20220327	188	54
13	ASC	20141123	20220322	186	48
115	ASC	20141025	20220329	188	56
42	ASC	20141008	20220324	188	62
166	DSC	20141029	20220402	173	41
93	DSC	20141012	20220328	177	38
20	DSC	20141007	20220323	181	43
122	DSC	20141014	20220330	183	51
49	DSC	20141021	20220325	178	48
151	DSC	20141016	20220401	174	49

2.2.3 Forward Modeling

To compare our InSAR results with possible signals from coupling on the subduction megathrust, we construct simple forward models of the subduction zone and use different coupling distributions to predict the surface velocities that would be observed by the ascending and descending satellites. These velocities can then be compared to the InSAR observations. Given the uncertainties in our data, and small size of the potential signal, we do not invert for a coupling distribution in this work.

We use the megathrust geometry from Slab2 (Hayes et al., 2018) and discretize the fault into 504 triangular patches (Figure 2.3). For the shallowest part of the fault we make an approximation as to how the fault comes to the sea floor. As this area is more than 80 km from the coast, our onshore deformation predictions are not sensitive to the assumptions that we make here. We use the backslip technique of Savage (1983) and equations of Okada (1985) to model the surface deformation due to coupling on the megathrust embedded in a linear elastic half space, implemented in the Classic Slip Inversion software package (<https://github.com/jolivet/csi>, Jolivet et al. (2015)). We set a convergence rate of 30 mm/yr at an angle of 10° , the same as the values used by Lin et al. (2015). The coupling value is defined as $1 - v_{patch}/v_{plate}$, where v_{patch} is the creep rate of the fault patch and v_{plate} is the

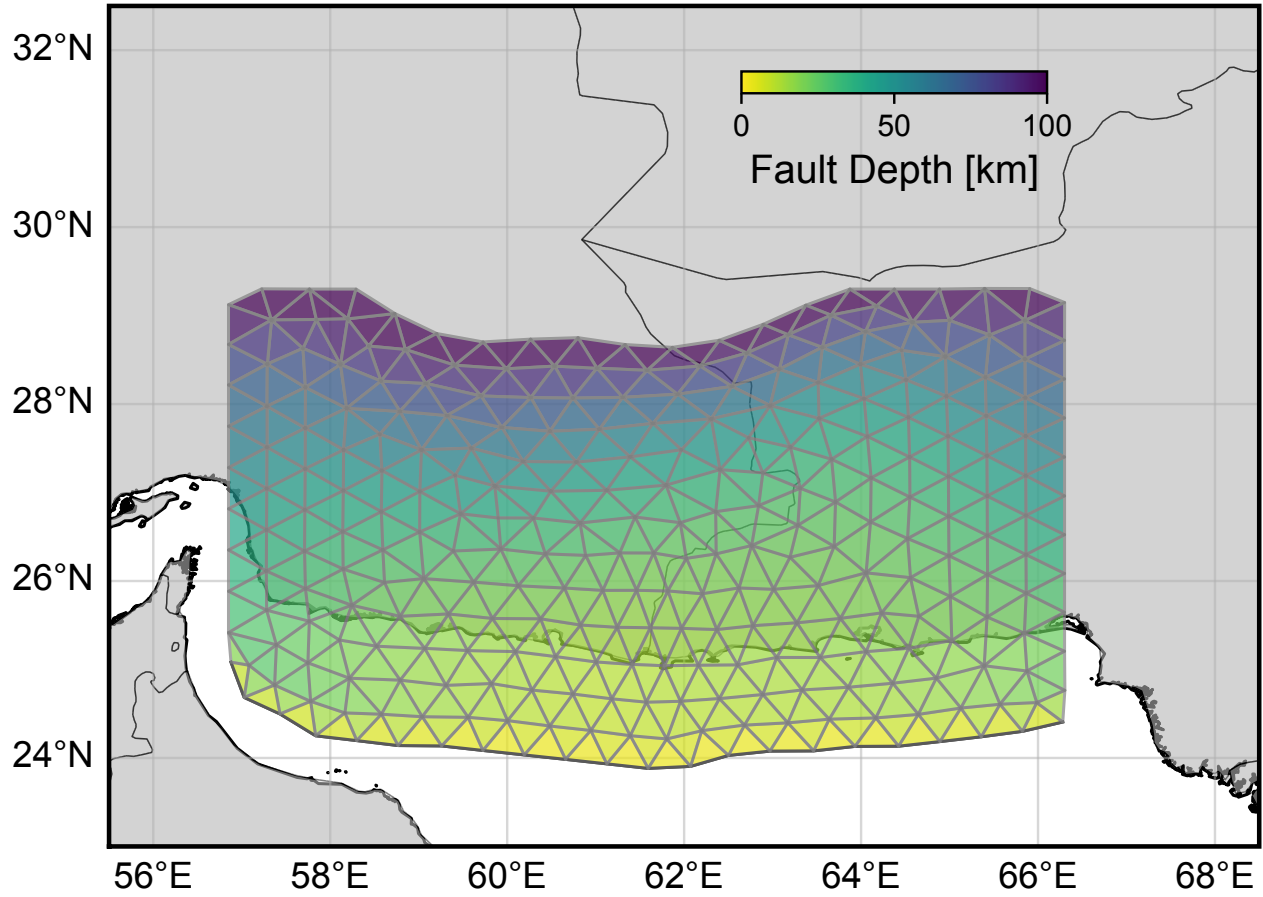


Figure 2.3: Megathrust geometry used in our forward models. The geometry is taken from Slab2 (Hayes et al., 2018), and we discretize the fault into 504 triangular patches. Black lines indicate coastlines and borders.

convergence velocity of the plate (i.e 30 mm/yr). A coupling value of 1 means that the fault is fully locked, with elastic deformation accumulating at the convergence rate, while a value of 0 means the fault is sliding without accumulating strain. For a given coupling distribution we project the modeled 3-D velocity field into the ascending and descending line of sight, and then compare the prediction with the InSAR observations.

In this work we present results from two end member coupling distributions: low coupling, where the fault starts at a coupling value of 0.5 at 13 km depth and is fully decoupled below 20 km, and high coupling, where the fault is fully coupled to a depth of 30 km, transitioning to decoupled below 40 km depth. These distributions are illustrated in Figures 2.4 and 2.5, respectively. In both cases we assume that

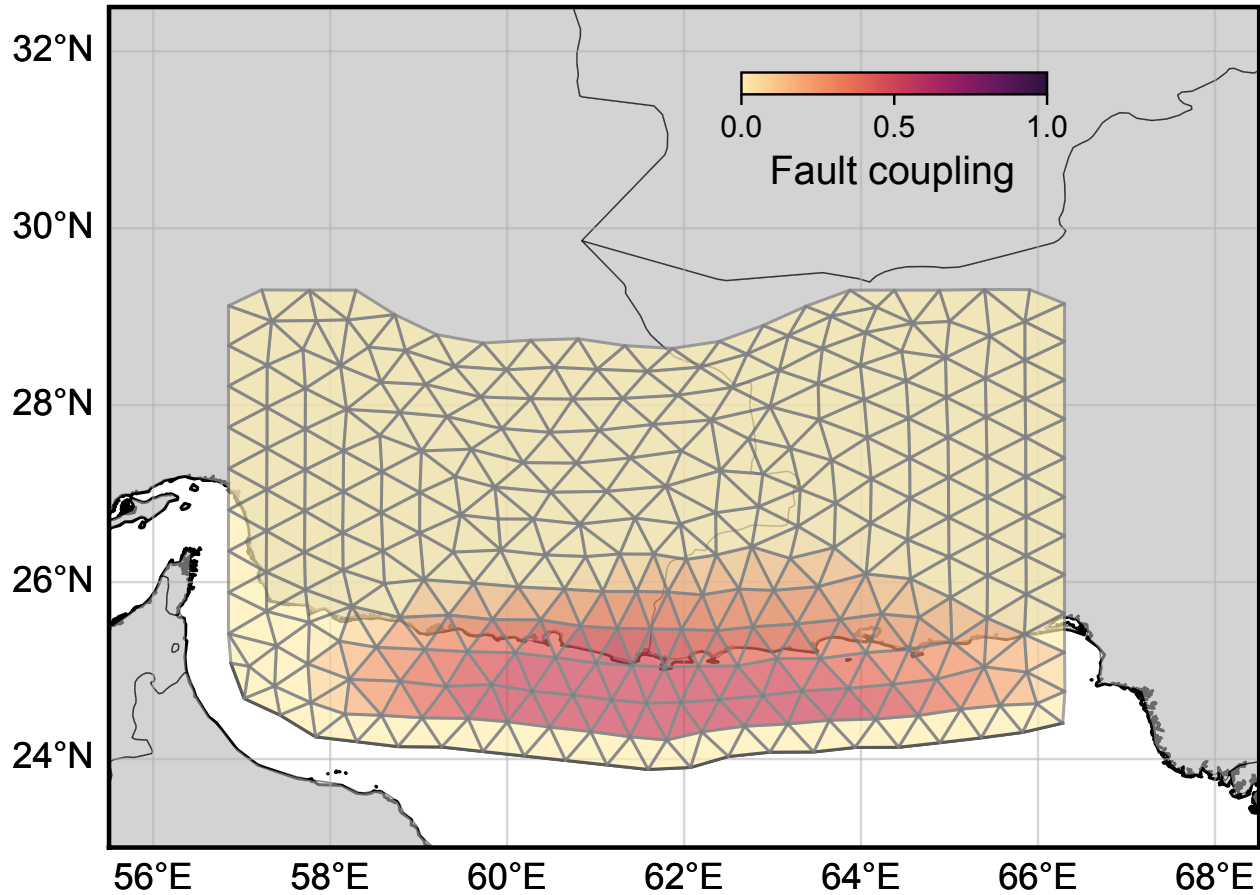


Figure 2.4: Low coupling model. The shallowest patches are all assumed to be decoupled, although this makes minimal difference to the modeled on-shore deformation. Coupling is 0.5 at 13 km depth, then transitions linearly with depth to decoupled below 20 km. Profiles of the modeled velocities at 60°E are shown in Figure 2.19

fault patches that come to the surface have zero coupling, but this assumption has minimal impact on our predicted velocities on the land.

2.3 Results

2.3.1 The Impact of Time Series Corrections

In Chapter 1, we illustrated the impact of each stage of the corrections on the long-wavelength InSAR-derived velocity fields. We apply the same corrections in this chapter, but also remove outliers before calculating the velocity. In Figures 2.7 and 2.8, we show the cumulative impact of corrections for an ascending and descending

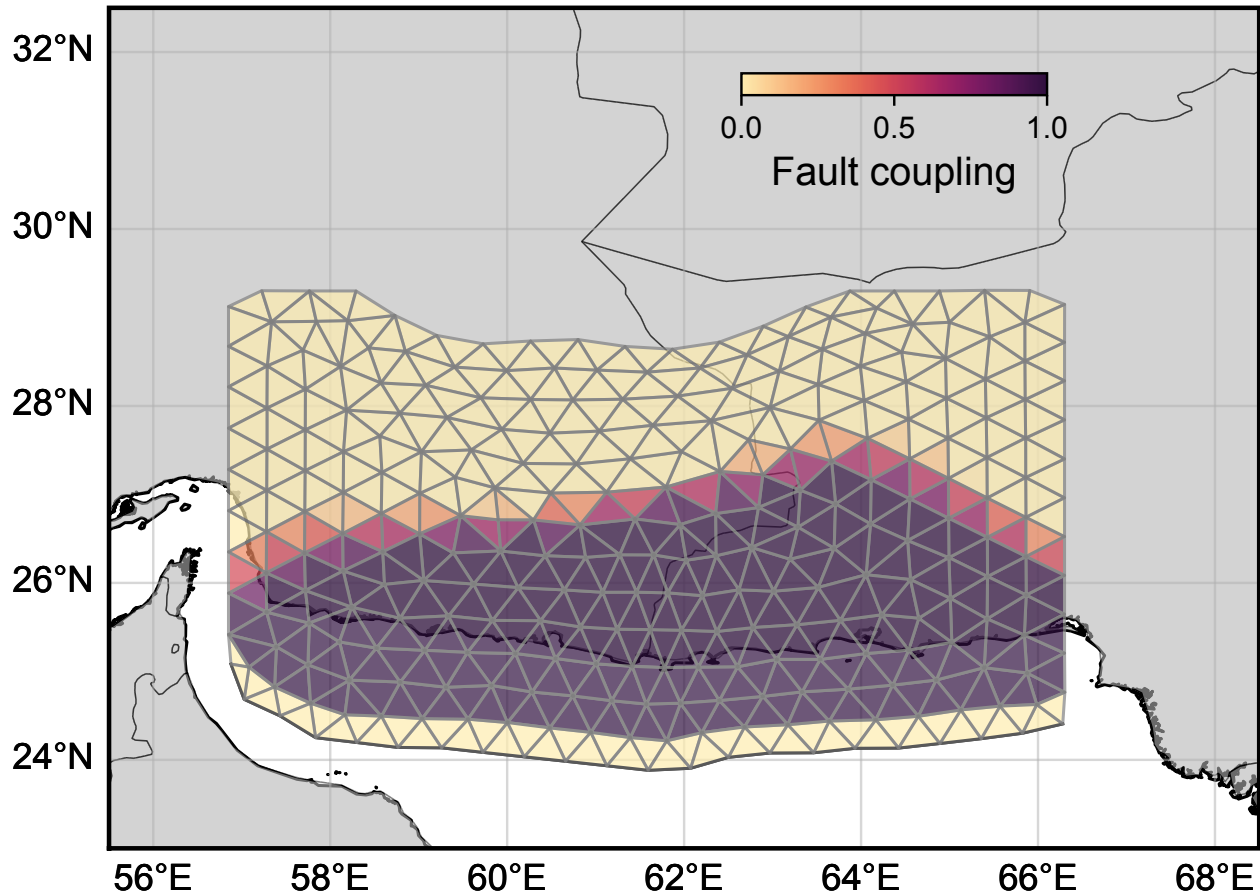


Figure 2.5: High coupling model. The shallowest patches are all assumed to be decoupled, although this makes minimal difference to the on-shore deformation. The fault is fully coupled to 30 km depth, then transitions linearly with depth to decoupled below 40 km. Profiles at 60°E are shown in Figure 2.20.

track on InSAR time series near the Makran coast, relative to a point 667 km to the north.

The initial time series shows scatter of tens of centimeters for both ascending and descending tracks, with the ascending track having much greater noise at the beginning of the time series due to ionosphere activity. Ionospheric corrections substantially reduce the variance and secular trend in the time series for the ascending track, while making a much smaller difference for the descending track. Troposphere corrections, using the ERA5 weather model, reduce the time series variance for both ascending and descending tracks. The solid Earth tide correction also reduces the variance; however, it has very little impact on the overall secular trend in the time series. The DEM error correction has the least impact on the time series.

After the above corrections have been applied, the signal of plate motion is still present in the time series (recall that we correct for plate motion in the velocity domain, rather than correcting the time series). This signal can be seen by examining the linear trend for tracks 86 and 20, shown in Figures 2.18(g) and (h) respectively. As the two tracks have opposite orbital directions, the change in satellite range between the reference point and time series location is positive for track 20 and negative for track 86, resulting in the relative LOS plate motion velocity being opposite for each track.

The final time series still show substantial scatter (e.g., around 10 cm for the ascending track in Figure 2.18(g), and somewhat less for the descending track in Figure 2.18(h)), likely due to residual troposphere (e.g., Parizzi et al. (2021)). Fattahi and Amelung (2015) showed that tropospheric noise in the region varies seasonally, and is particularly large near the coast, consistent with our results. We find sharp tropospheric fronts in our time series, particularly affecting the summer acquisitions, likely associated with the seasonal monsoon (Figure 2.6). Sharp changes in the troposphere are not well corrected by tropospheric models, and the residual tropospheric noise results in the coastal areas having the highest velocity uncertainties (Figures 2.12 and 2.14). There is a noticeable difference between the velocity uncertainties calculated for the ascending (Figures 2.12) and descending (Figures 2.14) tracks. We attribute this difference to a more energetic troposphere for the dusk-acquired ascending tracks compared to the dawn-acquired descending tracks.

To reduce the influence of large residual tropospheric signals we exclude outliers in the time series fit (Section 2.2.2). Our outlier threshold is aggressive, rejecting over 20 percent of the data for each track (Table 2.1). The changes in velocity due to outlier removal can be as high as ± 2 mm/yr, illustrating the substantial uncertainties, predominantly due to tropospheric residuals.

2.3.2 Plate Motion and Merging Tracks

After applying all time series corrections (but not correcting for plate motion), we calculate the linear velocity and merge the tracks as described in Section 2.2.2. We show the results of this merging for the ascending tracks in Figure 2.9 and descending tracks in Figure 2.10. The across-track ramp induced by plate motion combines to create a ramp in the range direction of around 30 mm/yr over 1200 km.

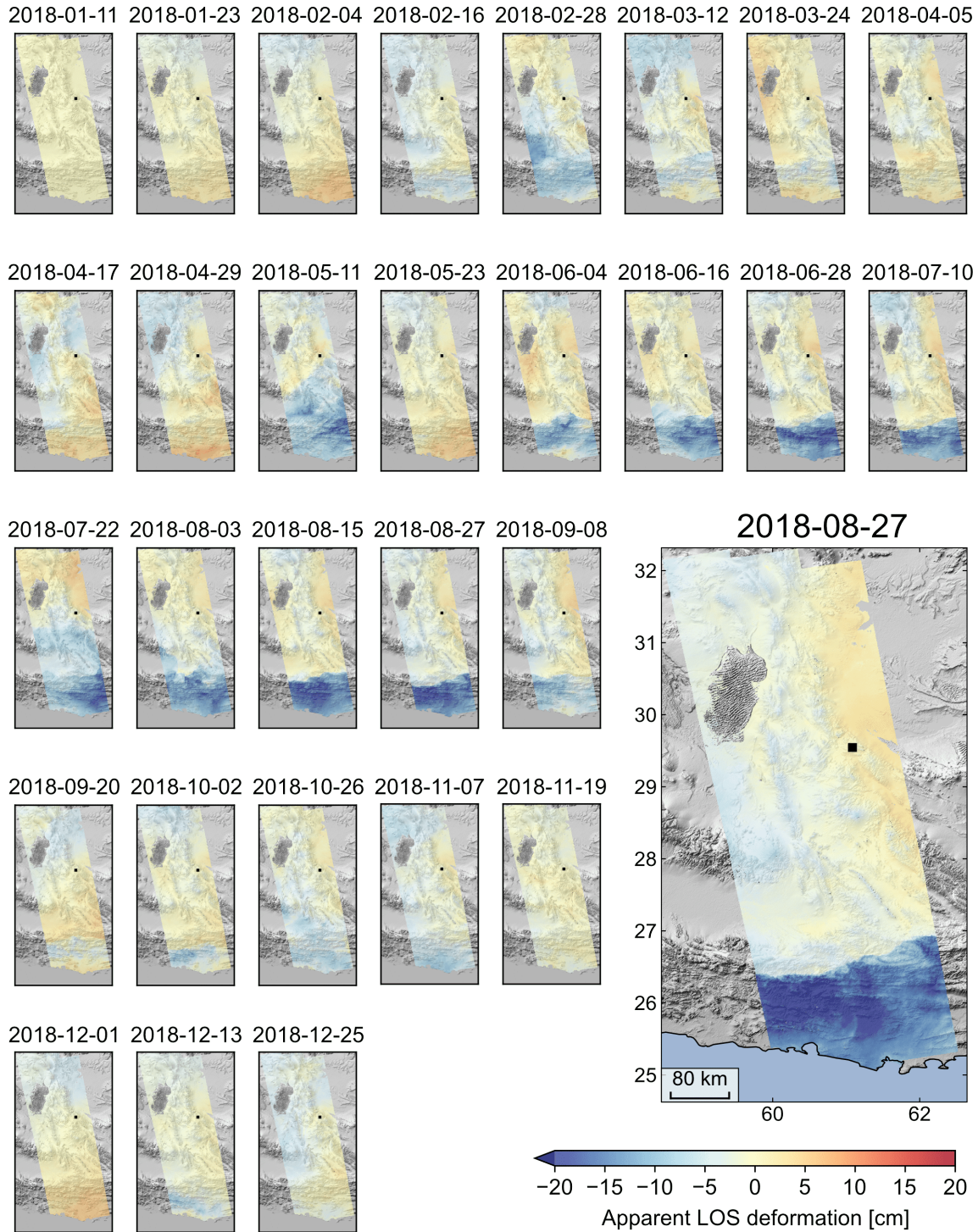


Figure 2.6: Example of strong tropospheric fronts in the InSAR time series. All 2018 data from track 86 (asc) is shown, after ionosphere corrections but before other corrections have been applied. The first date (2018-01-11) is taken as the temporal reference, and the black point shows the spatial reference. Sharp tropospheric fronts can be seen between July and August. The large inset on the bottom right highlights a particularly strong front on 2018-08-27, with over 20 cm of apparent deformation over a few kilometers.

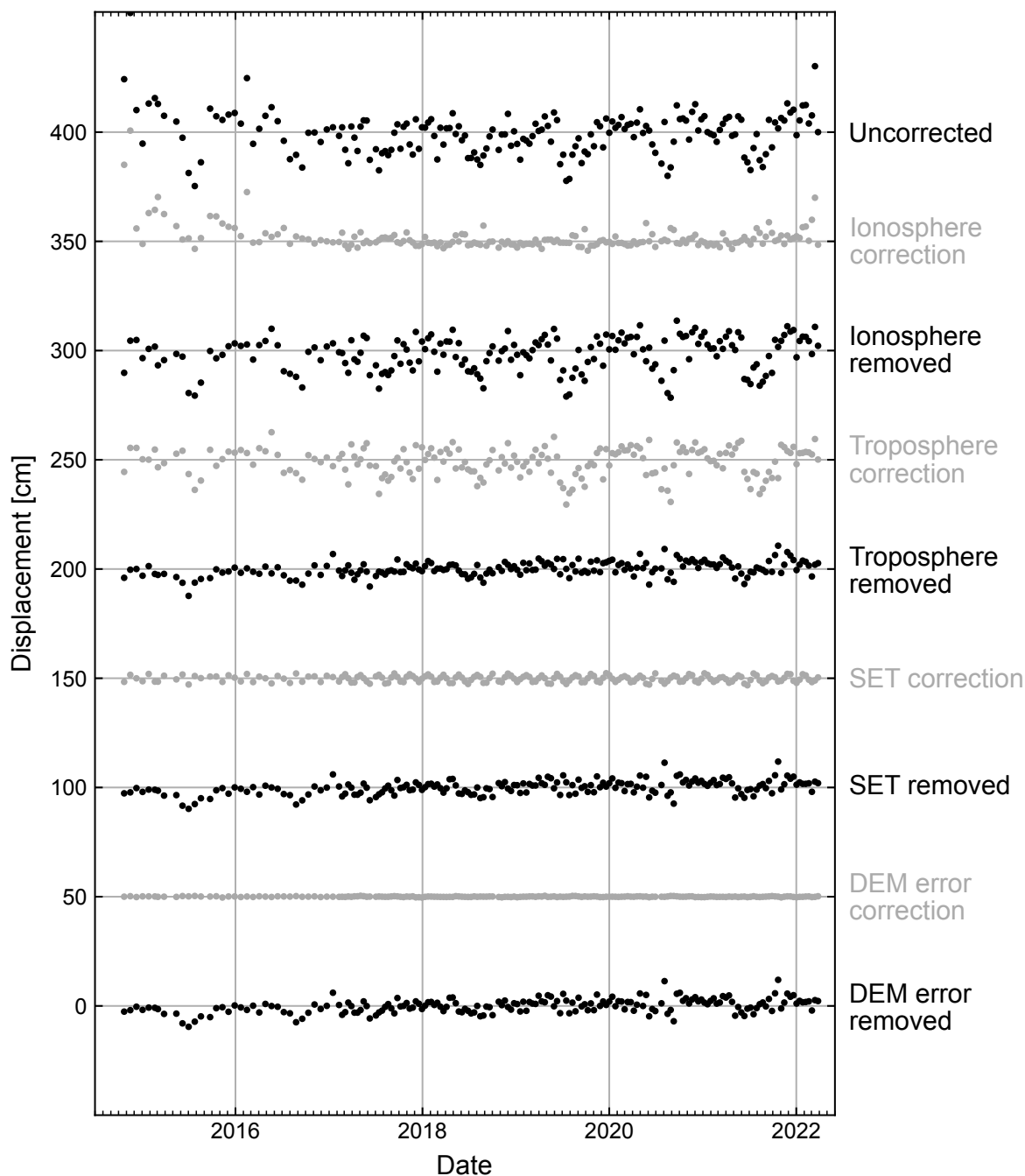


Figure 2.7: Impact of time series corrections for a point at (26°N , 60°E) relative to (32°N 60°E) for track 86a. Correction time series (grey) are subtracted from the displacement time series (black). Each time series is offset by 50 cm for display purposes. Corrections are applied cumulatively, meaning the time series at the bottom of the plot has had all the above corrections applied. Note that we do not apply plate motion corrections in the time series domain, so the final time series includes a secular trend due to plate motion. The final time series is shown on a larger scale in Figure 2.18(g).

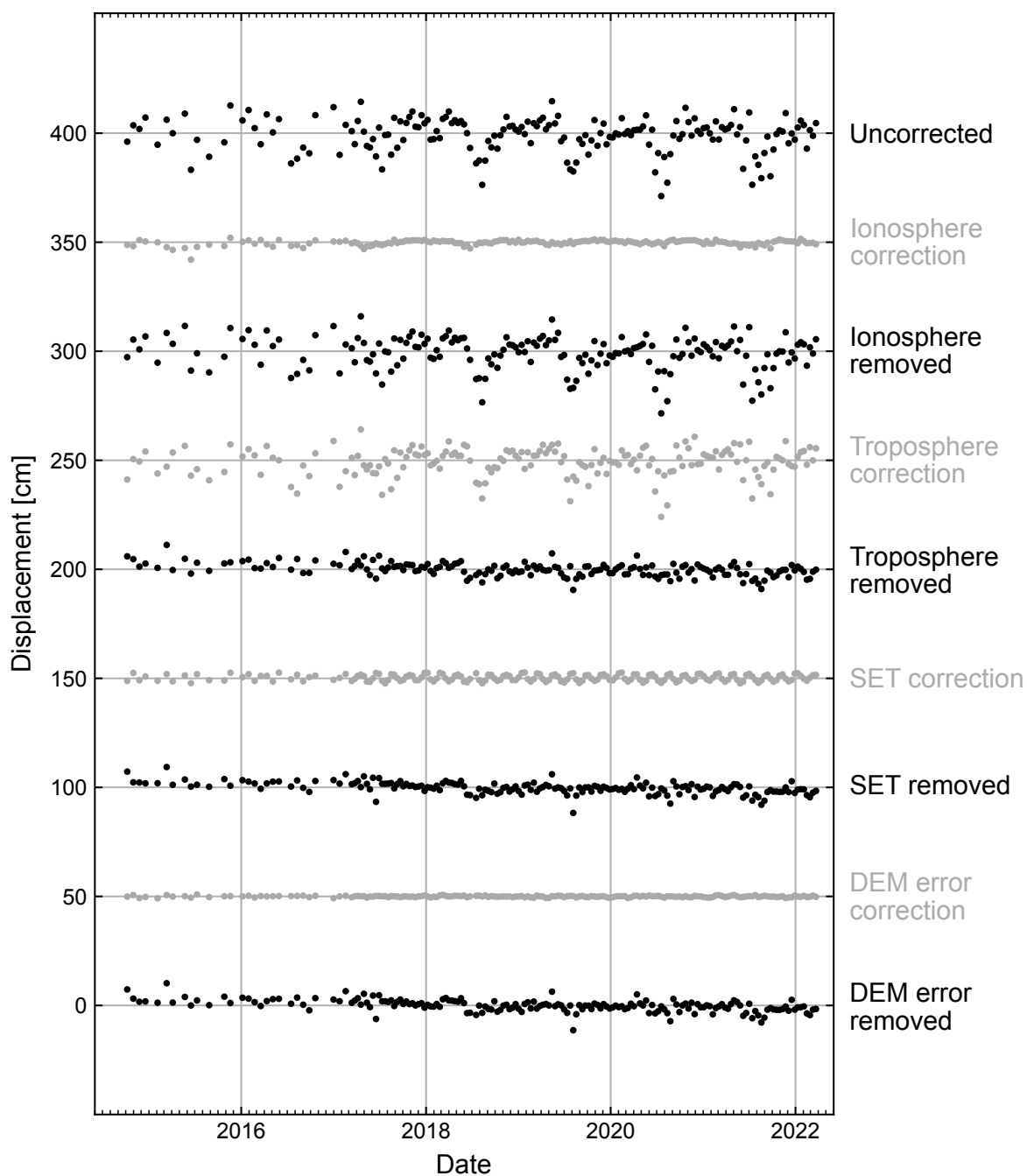


Figure 2.8: Time series corrections for track 20d. Note the much weaker ionosphere for the descending track. The final time series is shown on a larger scale in Figure 2.18(h). Caption is otherwise the same as Figure 2.7

This ramp clearly illustrates the limitations of neglecting the LOS difference in the overlap between the tracks when merging tracks without correcting for plate motion.

2.3.3 Exploring the Merged Velocity Fields

After removing the signal of plate motion from each track, we repeat the merging process and plot the results for the ascending (Figure 2.11) and descending (Figure 2.13) tracks, along with their uncertainties (Figures 2.12 and 2.14). The merged velocity fields no longer show large cross-track ramps, and allow us to discern the details of ground deformation at the level of millimeters per year over an area of nearly one million square kilometers.

There are noticeable discontinuities in the velocity between the tracks of up to several millimeters per year. These discontinuities are particularly prevalent at the southern end of the tracks. We attribute these discontinuities to two causes: (1) different noise realisations between adjacent tracks, and (2) differing sensitivities to tectonic deformation at the boundary between tracks. (1) is the likely cause of the discontinuity around (26°N, 60°E) in Figure 2.11. (2) can clearly be seen around (27°N, 64°E) in Figure 2.11, where a change of 15 degrees in the LOS incidence angle at the track boundary, coupled with predominantly westward motion from post-seismic deformation of the 2013 Balochistan earthquake (discussed below), results in a clear discontinuity of around 2 mm/yr.

We plot profiles through the velocity fields and example deformation time series in Figures 2.16 and 2.18 respectively. The locations of these profiles are illustrated in 2.15, and locations of the time series are given in Table 2.2. Figure 2.16(a) shows a profile through the ascending and descending merged velocity fields. The dominant feature is post-seismic deformation from the 2013 M_w 2013 Balochistan earthquake on the right of the profile (e.g., Jolivet, Duputel, et al. (2014), Lv et al. (2022), and Peterson et al. (2018)), which is also visible in profiles (e) and (f) and can be seen in Figures 2.11 and 2.13 as the dominant feature of the velocity field around (26-28°N, 64-66°E). Ascending and descending tracks have differing sensitivities to east-west motion, resulting in different velocity profiles from the ascending and descending tracks. This post-seismic deformation has been variously attributed to after slip on the Hoshab fault combined with viscoelastic relaxation of the accretionary prism (Peterson et al., 2018), or aseismic slip on the megathrust (Lv et al., 2022). In Figures 2.18(a) and (b), we illustrate logarithmic post-seismic deformation time series from Balochistan. The comparison between these two

profiles illustrates the spatial correlation of tropospheric noise (e.g., Emardson et al., 2003), with the time series showing deformation relative to a point 83 km away (Figure 2.18(a)) showing much greater scatter than deformation relative to a point 1 km away (Figure 2.18(b)).

A more subtle feature in profile 2.16(a), at 700-800 km along the profile, we attribute to post-seismic deformation from the 2013 Khash earthquake, a M_w 7.7 intraslab earthquake that occurred at 80 km depth (Barnhart, Hayes, Briggs, et al., 2014). Profile 2.16(d) gives a clearer picture of the velocity field from Khash post-seismic deformation, with a negative velocity lobe between 100 and 300 km along the profile and a positive velocity lobe between 300 and 400 km. These two lobes of the deformation can clearly be seen in Figures 2.11 and 2.13 around (28°N, 62°E). In Figure 2.18(c) we show a deformation time series from the area of peak LOS deformation in the ascending tracks. While the deformation is substantially smaller than for Balochistan post-seismic, we can see a logarithmic deformation profile with a cumulative LOS offset of 5 cm over 7.5 years.

Another region of post-seismic deformation is caused by the 2013 M_w 6.1 Minab earthquake, in the transition zone between the Makran subduction zone and the Zagros mountains (Penney et al., 2015). In Figure 2.18(d) we show a LOS deformation time series across the fault. Previous modeling work has attributed this motion to after slip (Plattner et al., 2021), and not viscoelastic relaxation. The deformation seems to have stopped by about mid-2020, although the substantial tropospheric noise makes it challenging to resolve small deformations without several years of data.

In Figure 2.18(e), we show a time series of LOS deformation across the Chaman fault, an 850 km long fault running through Afghanistan and Pakistan that accommodates motion between the Eurasian and Indian plates through both aseismic creep and co-seismic deformation (Figure 2.1). The time series exhibits both behaviors, showing slow creep and a co-seismic offset from an earthquake on June 27th 2018 (Dalaison et al., 2021).

Our velocity maps show numerous areas of rapid subsidence (< -2 cm/yr). These areas are most clearly visible in the western portion of Figures 2.9 and 2.10, where they can be seen despite the presence of the plate motion ramp. Rapid subsidence is correlated with areas of agricultural land use, and previous studies have attributed this motion to aquifer depletion due to over-extraction of water (e.g., Motagh et al. (2017)). In Figure 2.18(f), we present an example deformation time series, near the

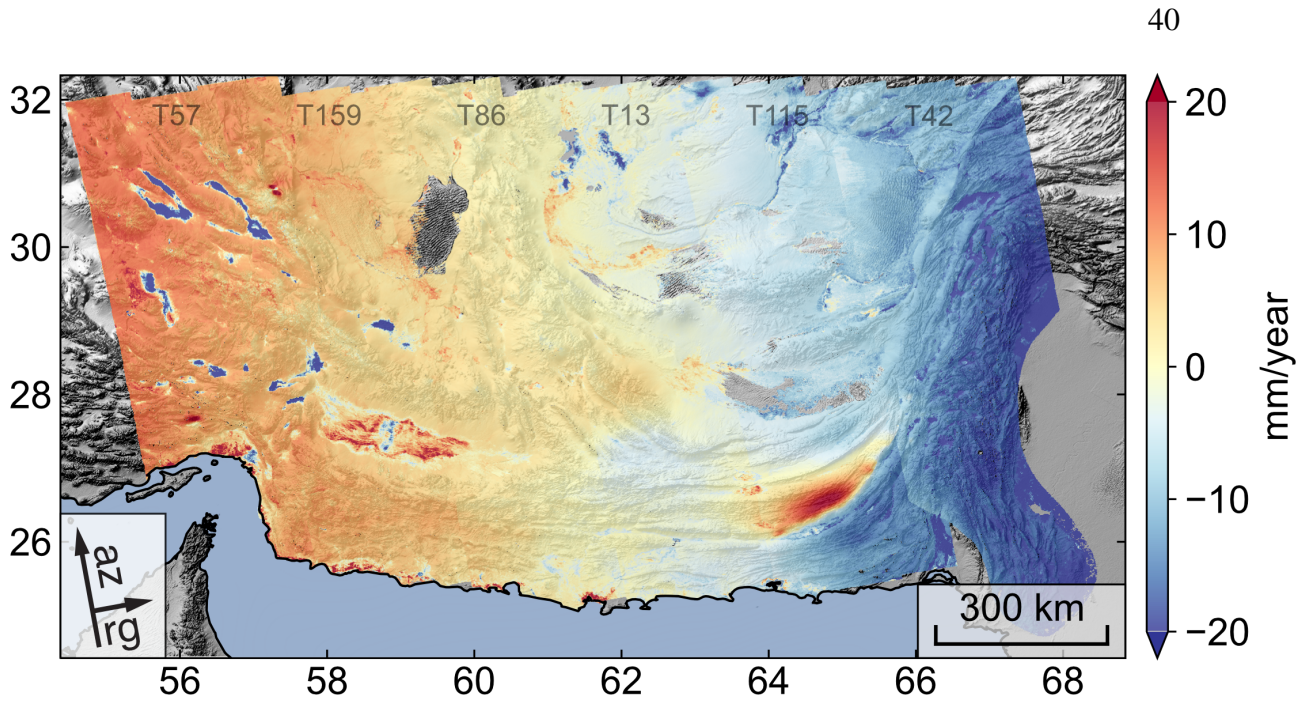


Figure 2.9: LOS velocity for ascending tracks after applying all corrections other than plate motion to every track, and merging. Plate motion creates a range-dependent velocity ramp in each track, resulting in a large ramp across the entire region. Track numbers are shown at the top of the figure. “az”: azimuth direction, the satellite direction of travel. “rg”: range direction, the look direction of the satellite. Positive velocities represent motion towards the satellite.

city of Rafsanjan, Iran, with a cumulative LOS offset of over 80 cm during the 7.5 year observation period.

All of our velocity maps rely on linear fits to our time series, but the data in Figure 2.18 shows deformation that is not linear in time. Behavior such as co-seismic steps and post-seismic logarithms will bias a purely linear fit. An example is the 2017 M_w 6.3 Pasani earthquake, which occurred on the eastern megathrust and caused 2-4 centimeters of vertical ground deformation (Yang et al., 2022). The linear rate around (25.5°N, 63°E) is biased by this offset. We choose a linear fit to illustrate the broad deformation trends across the region; however, a better velocity field could be calculated by modeling and removing earthquakes in the time series, and choosing different functional forms appropriate to the specific processes being studied.

2.3.4 Forward Models

To place constraints on the behavior of the subduction zone, we take four north-south profiles through our ascending and descending velocity fields, starting at the coast,

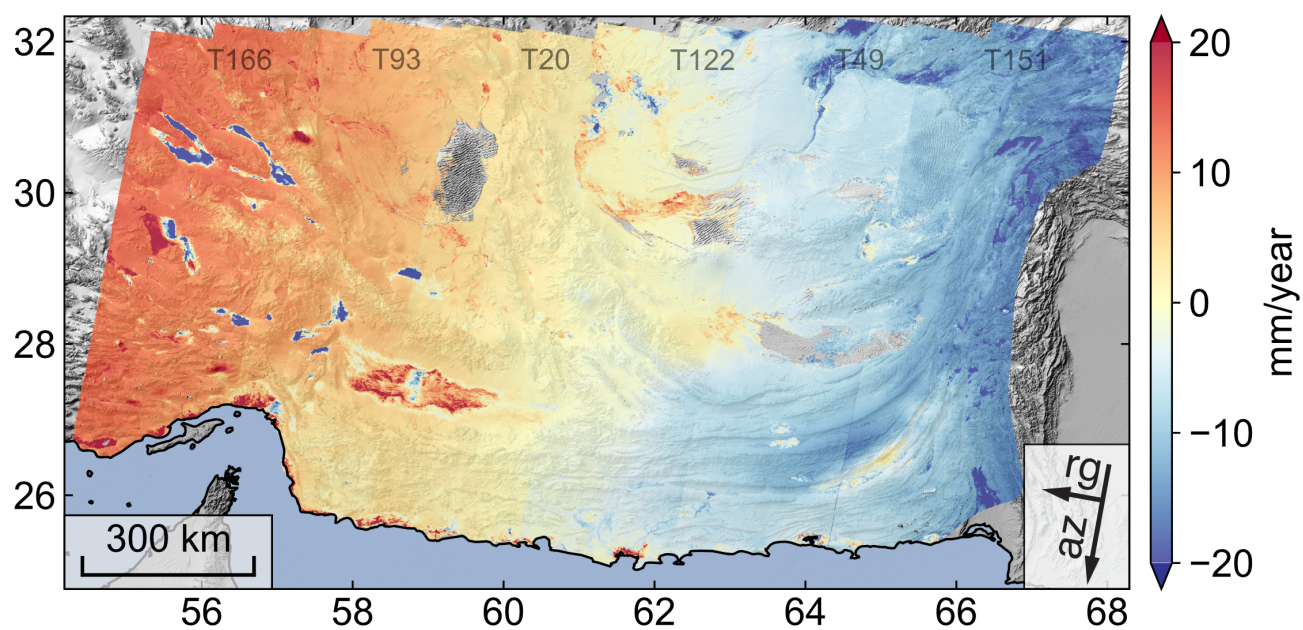


Figure 2.10: Same as Figure 2.9, except for descending tracks.

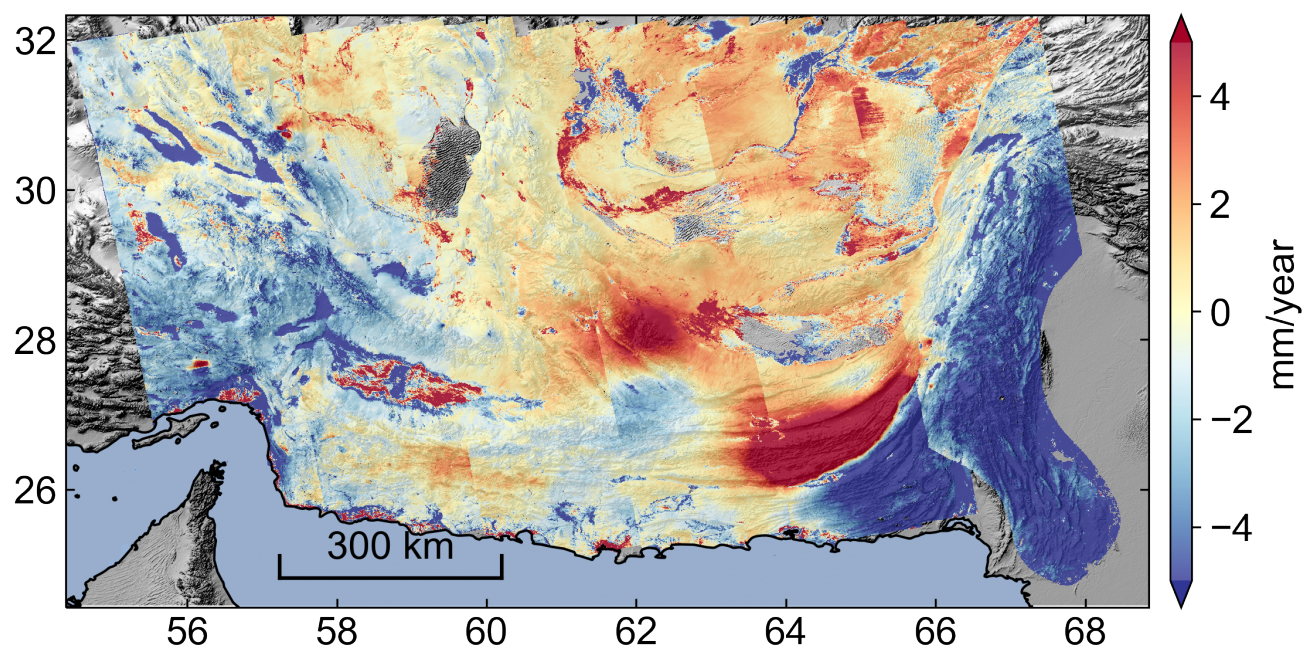


Figure 2.11: LOS velocity for ascending tracks after removing the effect of plate motion using a plate motion model.

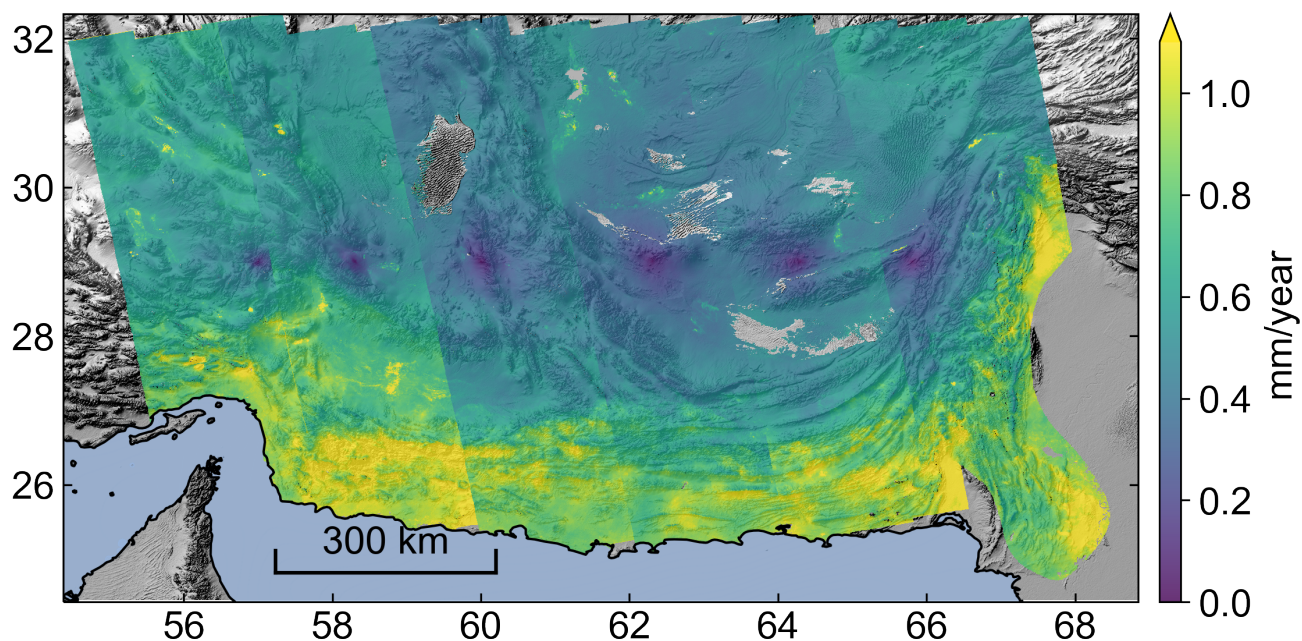


Figure 2.12: Estimated standard deviation of the time series velocity (σ_v) assuming uncorrelated Gaussian errors and after outlier removal for ascending tracks. Uncertainties are expressed relative to a reference point in each track (which has 0 uncertainty), with each reference point at a latitude of 29°N.

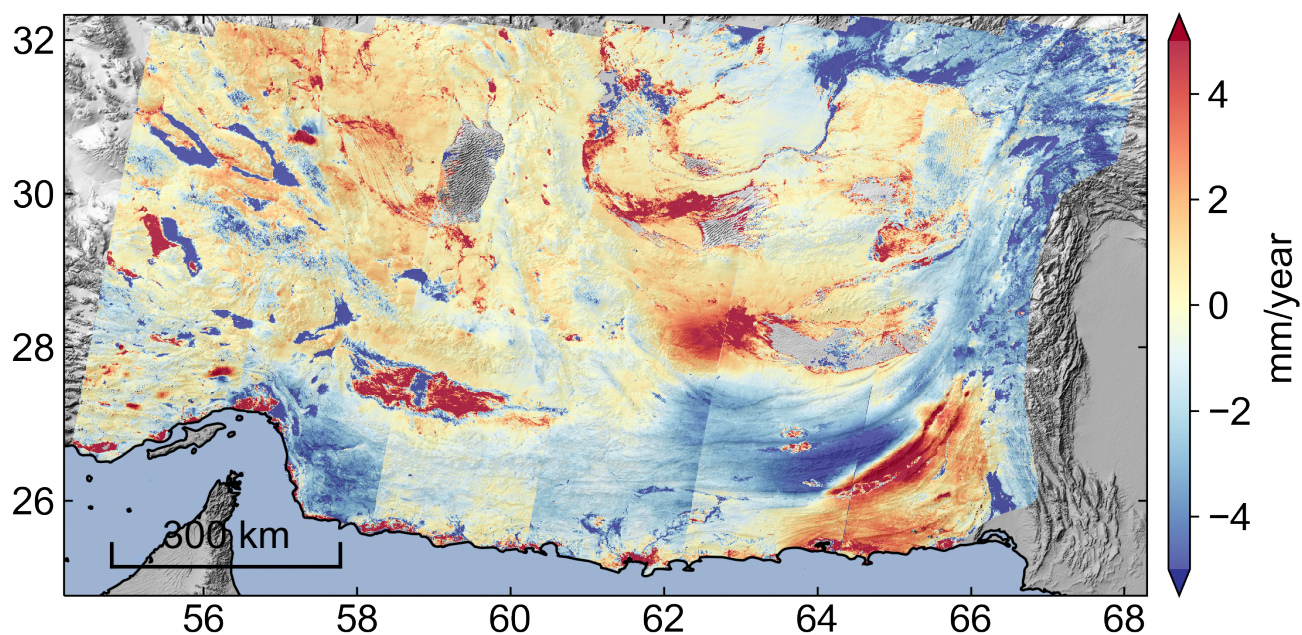


Figure 2.13: LOS velocity for descending tracks after removing the effect of plate motion using a plate motion model.

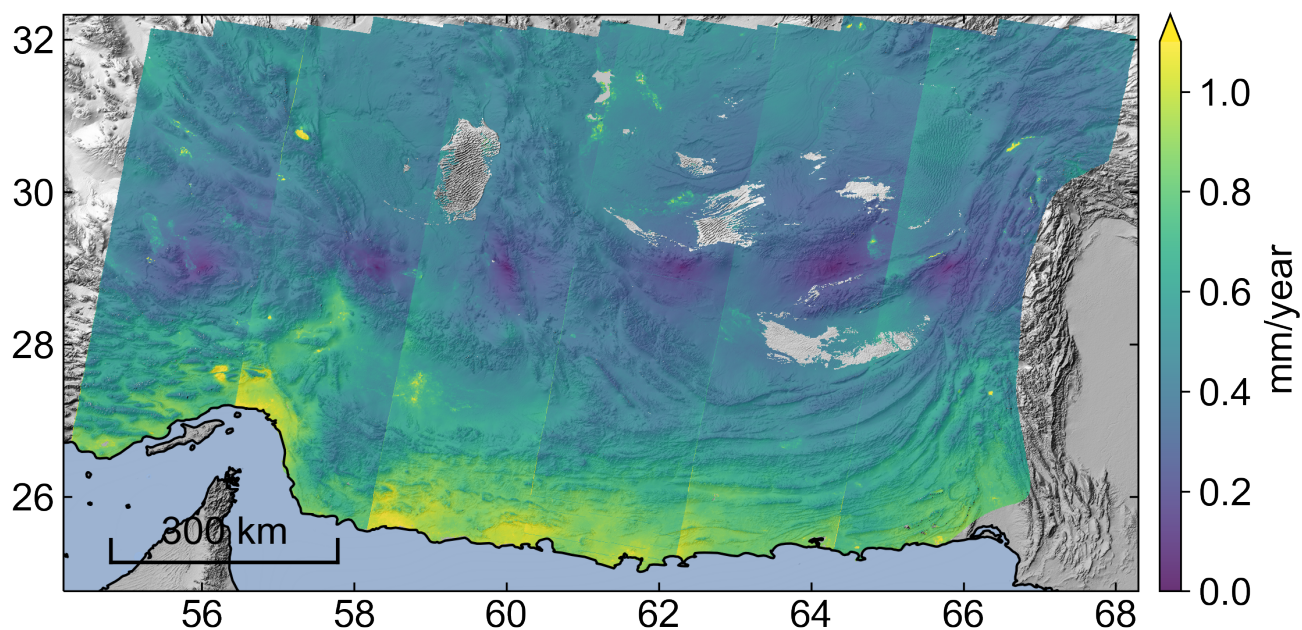


Figure 2.14: Uncertainties for the descending tracks. Caption is otherwise the same as Figure 2.12.

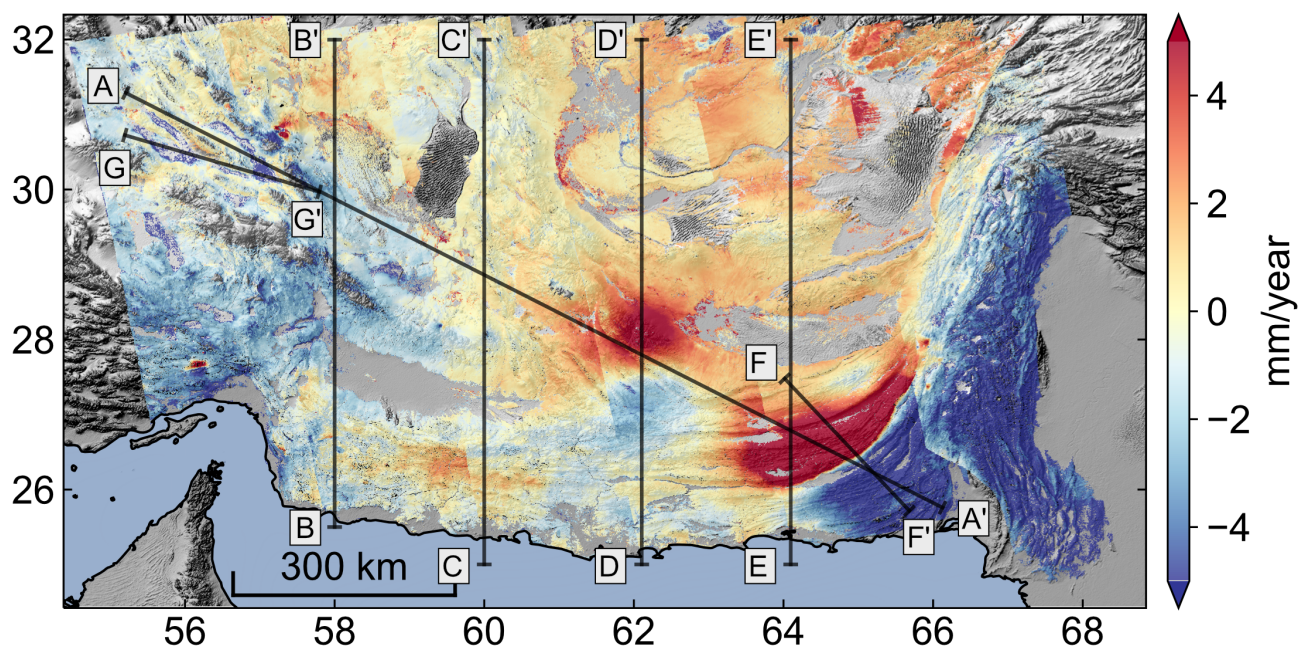


Figure 2.15: Location of the velocity profiles plotted in Figure 2.16, shown over the merged ascending track LOS velocities. Areas which do not unwrap in every interferogram have been masked, and are excluded from profiles, other than for (g), where the temporal coherence mask is used.

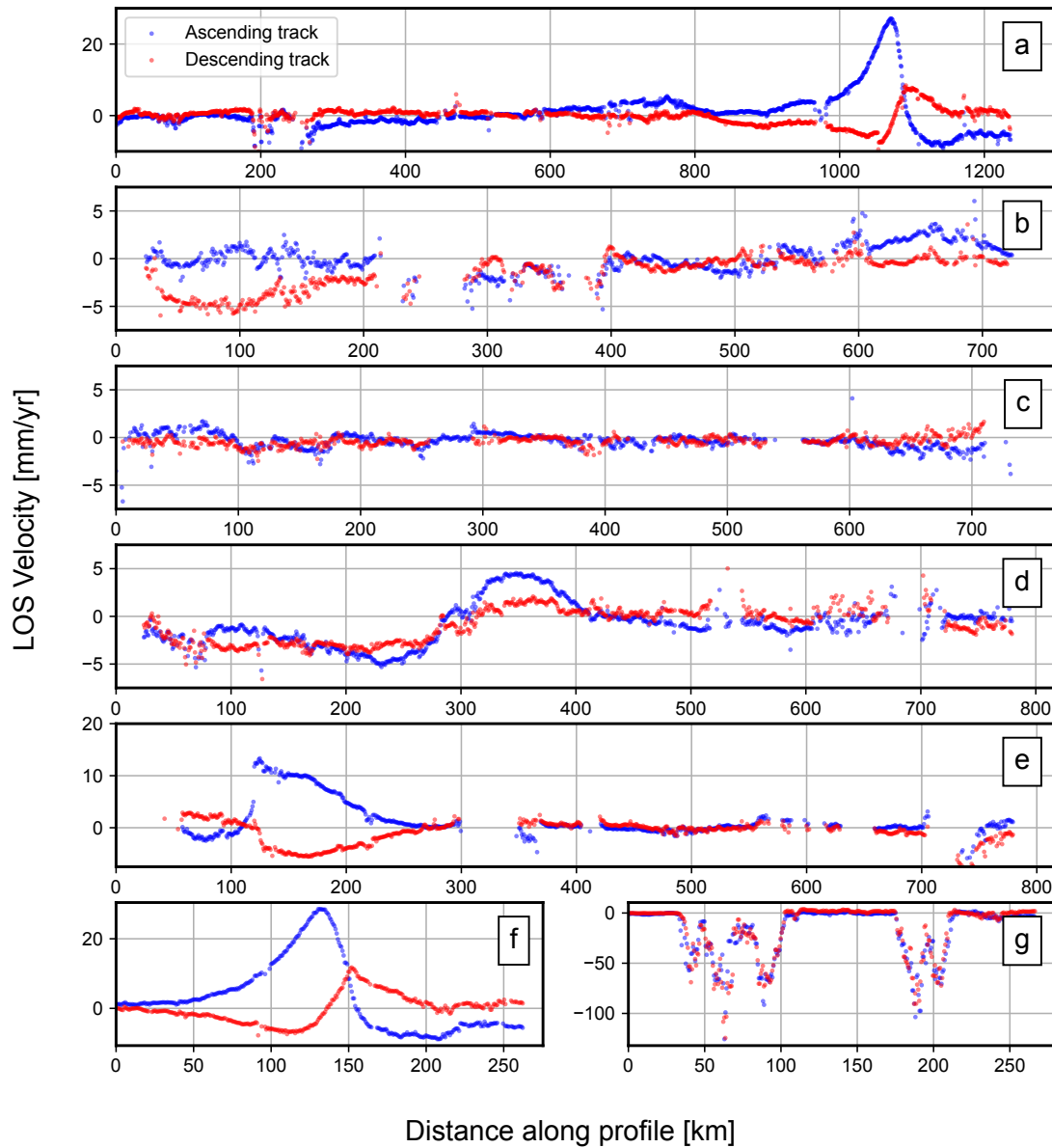


Figure 2.16: Profiles through ascending (blue) and descending (red) track LOS velocity fields. Locations of the profiles are shown in Figure 2.15, and profile distances are calculated from the letter, to the letter primed (e.g., A to A'). **(a)** Region-spanning profile. **(b)** South to north profile at 58°E (25.5°N to 32°N). **(c)** South to north profile at 60°E (25°N to 32°N). **(d)** South to north profile at 62.1°E (25°N to 32°N). The left side of the profile spans post-seismic deformation from the 2013 Khash earthquake. **(e)** South to north profile at 64.1°E (25°N to 32°N). The left side of the profile spans post-seismic deformation from the 2013 Balochistan earthquake. Note that profiles (b)-(e) are chosen to lie within individual tracks to avoid offsets at the track boundaries. **(f)** Profile spanning post-seismic deformation from the 2013 Balochistan earthquake. **(g)** Profile spanning rapid subsidence near the city of Rafsanjan, Iran, likely due to aquifer depletion.

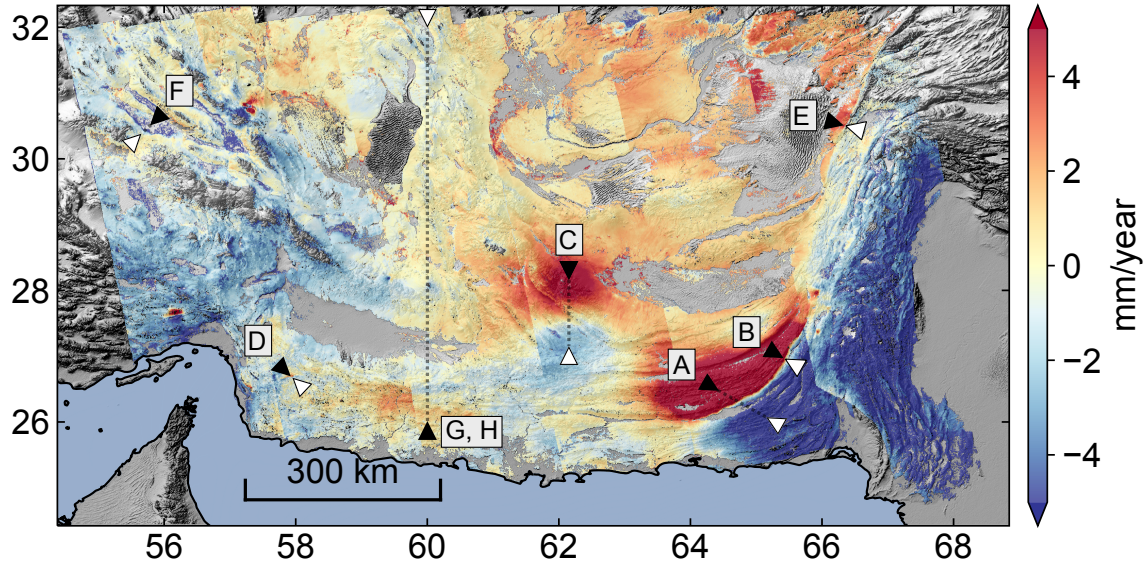


Figure 2.17: Locations of time series (a)-(h) plotted in Figure 2.18, plotted over ascending track velocities. All time series are plotted relative to a reference point. The location of the time series is shown with a black triangle, and the location of the reference point is shown with a white triangle. Note that all time series are relative to a point within the same track as the chosen point, and are plotted before the removal of the plate motion signal (which is only significant for points far from their reference, such as G and H. Tracks and exact locations are given in Table 2.2.

Letter	Name	Track	Location ($^{\circ}$ N, $^{\circ}$ E)	Reference ($^{\circ}$ N, $^{\circ}$ E)	Distance (km)
(a)	Balochistan	115a	26.47, 64.44	26.08, 65.15	83
(b)	Balochistan	115a	26.966, 65.421	26.959, 65.429	1
(c)	Khash	13a	28.16, 62.15	27.15, 62.15	112
(d)	Minab	57a	26.688, 57.928	26.67, 57.95	3
(e)	Chaman	42a	30.511, 66.325	30.50, 66.363	4
(f)	Rafsanjan	166d	30.505, 55.780	30.384, 55.655	18
(g)	Makran	86a	26.0, 60	32.0, 60.0	667
(h)	Makran	20d	26.0, 60	32.0, 60.0	667

Table 2.2: Details of the time series plotted in Figure 2.18. “Track” refers to the Sentinel-1 track, with “a” and “d” referring to ascending and descending tracks respectively. Time series are taken at the point given by “Location” and the InSAR reference point is set at “Reference.” “Distance” refers to the distance between the time series location and the reference point.

which are shown in Figures 2.16(b)-(e). The locations of these profiles are chosen such that they lie entirely within a single ascending and descending track, so we do not introduce additional uncertainties related to track merging. Of these profiles,

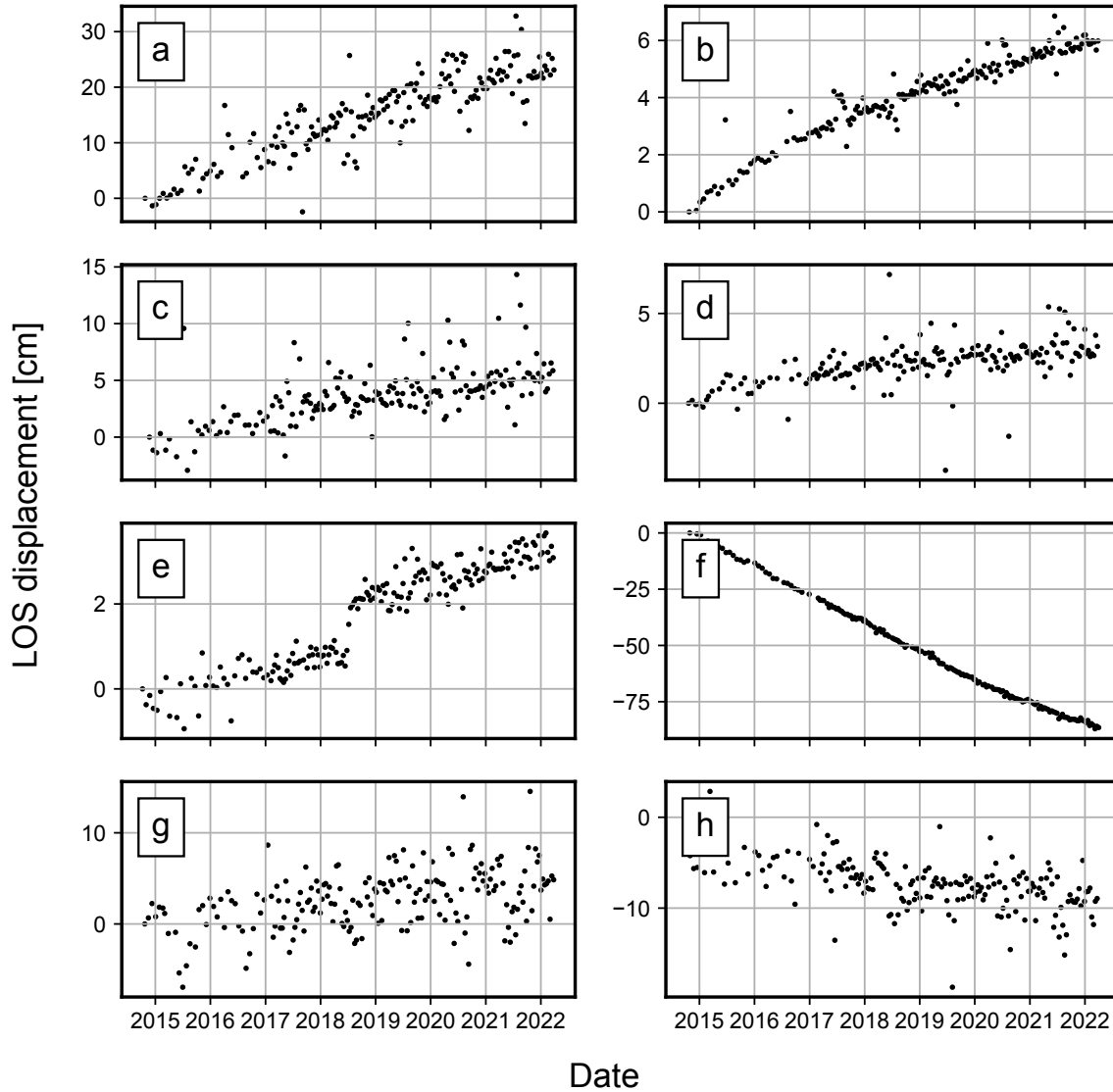


Figure 2.18: InSAR LOS displacement time series illustrating a variety of different temporal behaviors in our region. The locations of the time series and reference points are given in Figure 2.17 and Table 2.2. **(a)** post-seismic deformation from the 2013 Balochistan earthquake near the point of peak deformation. **(b)** Post-seismic deformation from the 2013 Balochistan earthquake across a portion of the fault creeping at the surface. The proximity to the reference point (1 km away) results in much lower tropospheric noise compared to (a). **(c)** post-seismic deformation from the 2013 Khash earthquake. **(d)** post-seismic deformation from the 2013 Minab earthquake. **(e)** Creep on the Chaman fault, and co-seismic offset from an earthquake on 2018-06-27 (Dalaison et al., 2021). **(f)** Rapid subsidence near the city of Rafsanjan, Iran, likely due to aquifer discharge (Motagh et al., 2017). **(g)** Time series near the coast for track 86 after corrections shown in Figure 2.7. A linear trend partly due to plate motion can be seen. **(h)** Same as (g), but for track 20, after corrections shown in Figure 2.8.

2.16(d) and (e) have substantial signals from Khash and Balochistan earthquake post-seismic signals respectively, and (b) lies in the transition zone between the Makran subduction zone and the Zagros mountains collision zone (e.g., Penney et al. (2015) and Regard et al. (2005)). We therefore cannot straightforwardly compare deformation in these profiles to different models of coupling on the megathrust. The profile (c), at 60°E, does not seem to have obvious sources of contamination, so we select this profile for comparison with forward models.

In Figures 2.19 and 2.20, we compare our velocity profiles with predicted LOS deformation rates from the two coupling models presented in Figures 2.4 and 2.5. Note that we plot these profiles in terms of distance from the subduction trench. The high sediment input means there is not a clear trench in the bathymetry (Schlüter et al., 2002), but we can still locate a deformation front, which is at around 24°N for our profile along 60°E. We add to the velocity profiles uncertainty estimates based on the calculated σ_v (Equation 2.1). For a conservative estimate of the error, we set our error range to be $\pm 2\sigma_v$. This choice gives us an error range of around ± 2 mm/yr near the coast, relative to a point around 400 km to the north.

For our simple linear elastic forward models, the transition between coupled and decoupled behavior on the megathrust corresponds to a peak in the surface velocity, with the size of that peak related to the strength of coupling and depth ranges over which that transition occurs. The assumed convergence direction of 10° means that megathrust coupling results in an eastward component of the deformation field. As a result of their differing lines of sight, ascending and descending tracks have different sensitivities to this eastward motion, making the combination of ascending and descending tracks a key tool for constraining potential coupling. This difference in sensitivity can be seen prominently in the high coupling model (Figure 2.20), where the predicted ascending and descending track LOS velocities differ by over 4 mm/yr near the coast.

The low coupling model in Figure 2.19 gives velocity predictions that lie well within the bounds of our observational uncertainty. Therefore, this coupling scenario cannot be ruled out by our data. Lower coupling scenarios, including a completely unlocked megathrust, are also possible, and coupling could be somewhat stronger, particularly off the coast where we have little sensitivity. We illustrated such an intermediate coupling model in Figure 2.21, which is at the limit of what is compatible with our data, given the conservative error bounds. For the high coupling model in Figure 2.20, the predicted velocities are clearly not consistent with the observed ve-

locity fields, lying well outside of the observational uncertainty bounds for much of the 200 km closest to the coast. A visual inspection of the other north-south profiles in Figure 2.16 does not show any clear evidence for the kind of strong coupling north of the coast that we use in Figure 2.20, although, as previously mentioned, there are other signals present in these profiles that could hide coupling signals. These results suggest that, under the simplifying assumptions of our fault geometry, convergence velocity, and forward model, strong coupling on the megathrust is unlikely to be present north of the coastline, at least for the western part of the subduction zone. However, these results do not rule out the potential for a significant earthquake on this part of the subduction zone.

2.4 Discussion

2.4.1 Quality of the InSAR-Derived Velocity Fields

The uncertainties we presented in Figures 2.12 and 2.14 represent the estimate of the error on the velocity, relative to the reference point within the track, and under the assumption of temporally uncorrelated Gaussian errors (Fattahi & Amelung, 2015). It does not capture biases in the velocity, for example related to incomplete removal of plate motion, or errors in the offsets between the tracks used when merging. Within each track the error is likely dominated by residual tropospheric signal (e.g., see Fattahi and Amelung (2015) and Parizzi et al. (2021), and our discussion of uncertainties in Appendix A), with some small contribution from neglecting the burst discontinuities when correcting the ionosphere (Liang et al. (2019), and see Appendix A). Treating the error as uncorrelated in time is an approximation, as there is seasonal variability in the tropospheric signal (Fattahi & Amelung, 2015). Quantifying the uncertainties using the residuals from a purely linear fit is therefore potentially an overestimate of the uncertainty in the velocity.

Our InSAR velocity uncertainties are substantially larger than those found by Lin et al. (2015) (around 0.5 mm/yr, see their Figure 2(b)) for three Envisat tracks in the eastern Makran, in spite of their lack of ionosphere corrections, and much smaller quantities of data. Their use of dusk-acquired descending tracks, which are much less affected by the ionosphere, and were mostly acquired during a relative low in solar activity during the period 2006-2010 (Liang et al., 2019), means the ionosphere likely has a small effect. They also make use of direct observations of the troposphere from the Medium Resolution Imaging Spectrometer on board Envisat, which increases the accuracy of their tropospheric corrections. However, it is still surprising that their estimated errors are so much lower with no more than 23

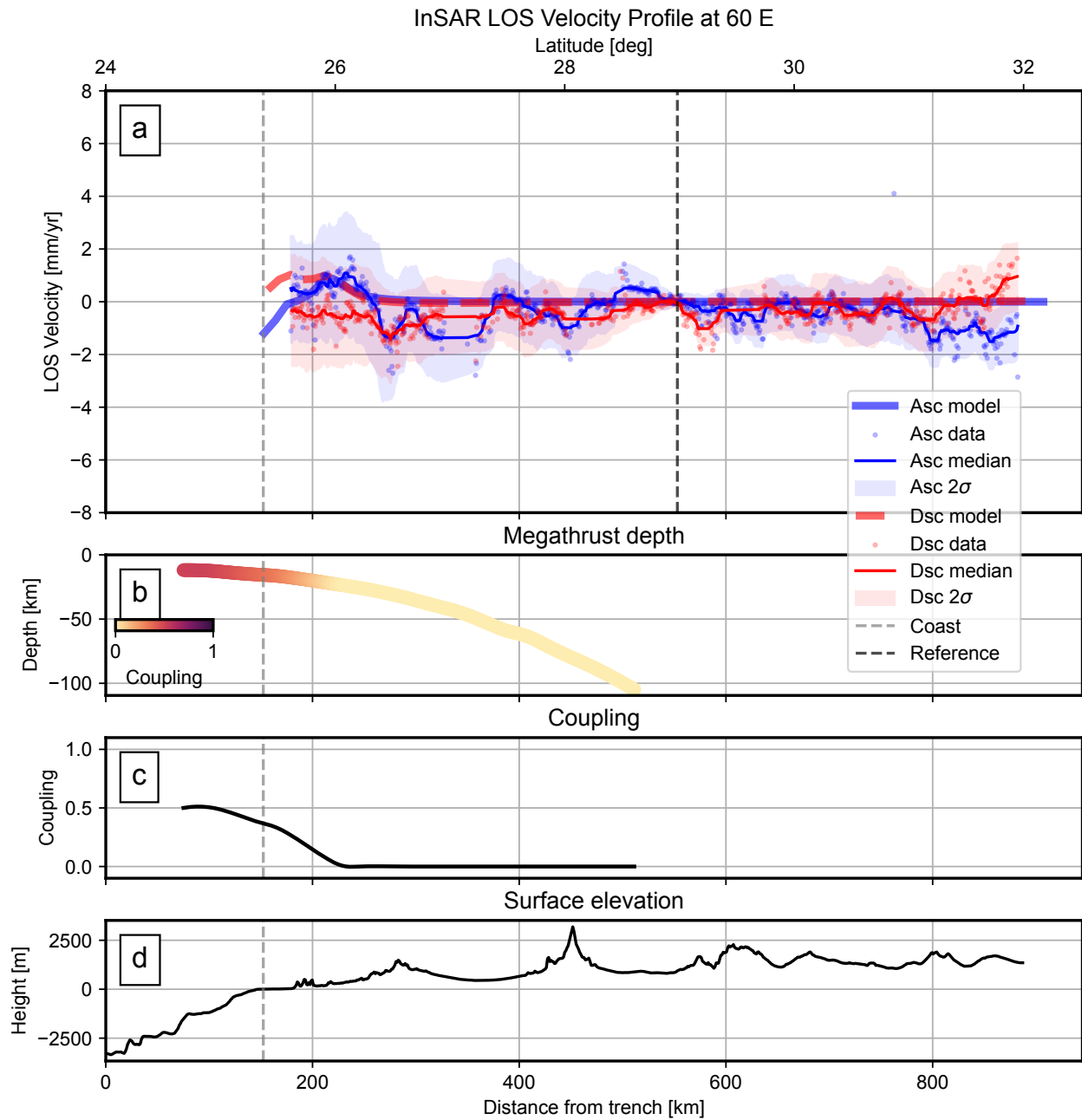


Figure 2.19: Profiles at 60°E for the low coupling model shown in Figure 2.4. **(a)** Profiles through the LOS velocity and estimated uncertainties ($2\sigma_v$) for ascending (asc) track 86 and descending (dsc) track 20. All velocities and uncertainties are relative to the reference point. Thin blue (asc) and red (dsc) lines show the running median of the InSAR data. Thick blue (asc) and red (dsc) lines show the predicted LOS velocity from the forward model. **(b)** Profile through the megathrust depth, with the color showing the coupling. **(c)** Plot of the coupling from (b). **(d)** Surface elevation profile.

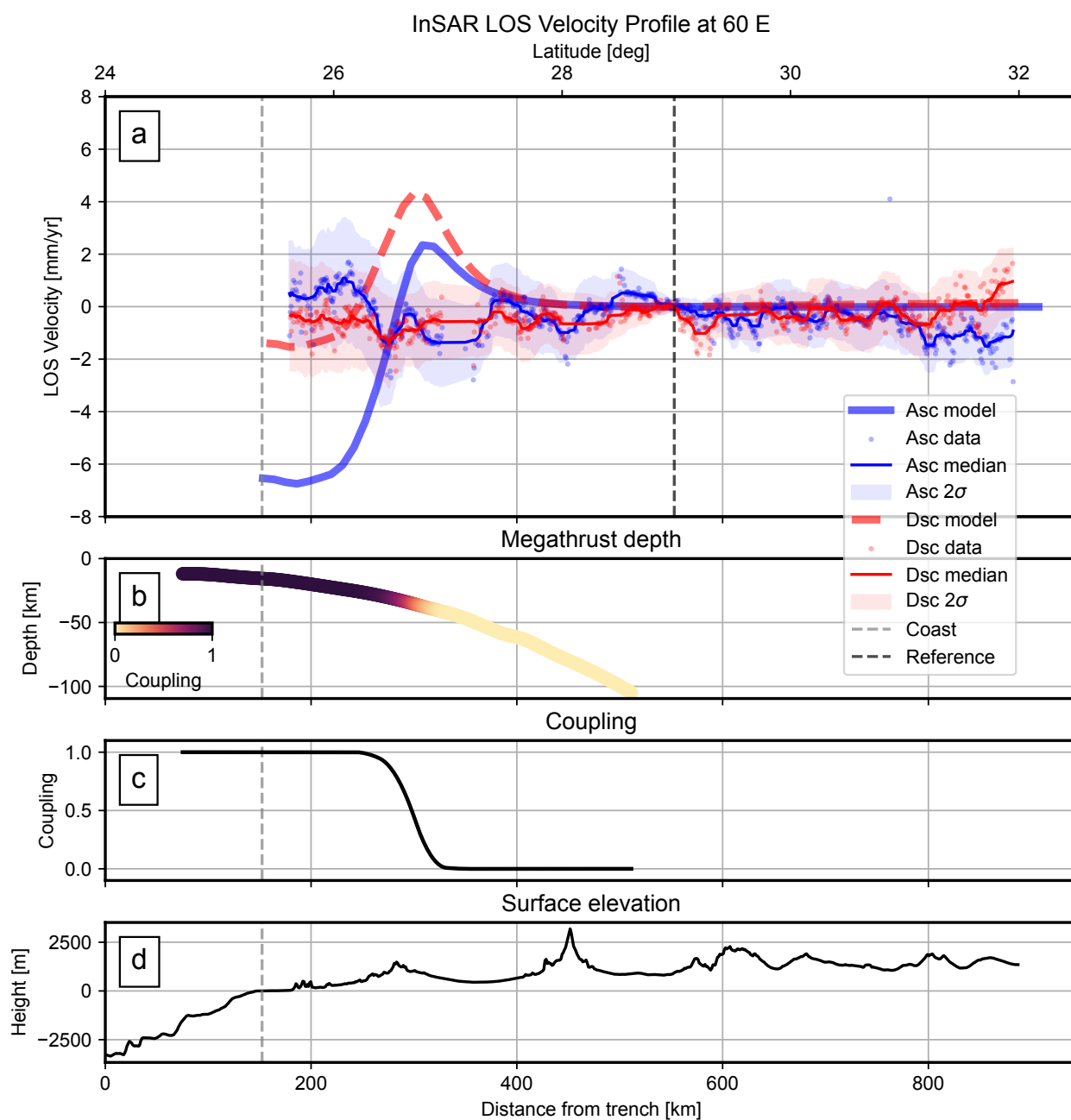


Figure 2.20: Same as Figure 2.19, but with higher coupling on the megathrust (coupling as shown in Figure 2.5). The increased coupling results in larger variations in the predicted InSAR LOS velocity, which are not consistent with our observations.

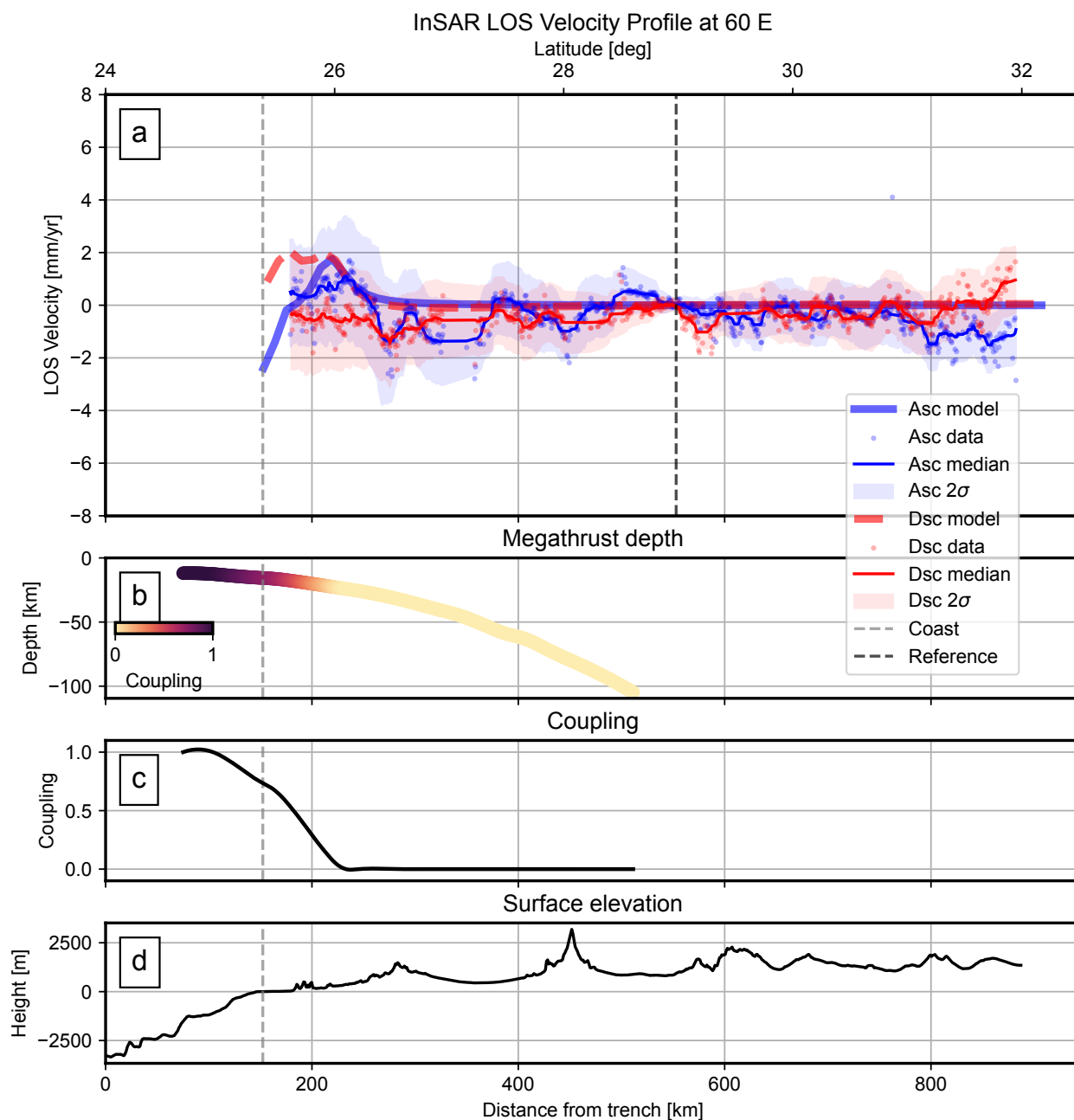


Figure 2.21: Same as Figures 2.19 and 2.20, but with intermediate coupling on the megathrust. This coupling is at the limit of what is compatible with our data, given the conservative errors bounds we have chosen.

acquisitions in a track. This comparison supports the idea that our error estimates are likely overly conservative.

Our approach of rejecting outliers and then performing an unweighted linear fit to the time series is a sub-optimal approach to extracting the velocity and its uncertainty. A more rigorous approach would be to quantify the uncertainty in each interferogram, then propagate that uncertainty through the calculation of the time series and then the velocity. This propagation of uncertainty could be done assuming that the sole contributor to the uncertainty is from residual troposphere (e.g., Parizzi et al. (2021)), or by using a full noise model for the InSAR time series (e.g., Agram and Simons (2015)). Such approaches allow for the use of all available data, with noisier acquisitions down-weighted in the calculation of the velocity, and provide a more rigorous estimate of the error on the final velocity. The Makran region presents specific challenges to such approaches, as the tropospheric signal can show sharp north-south variation. This variation means that the method of calculating a single tropospheric covariance function for the entire track will likely not give a representative uncertainty (e.g., Emardson et al. (2003) and Lohman and Simons (2005)). It may therefore still be necessary to reject acquisitions that show strong variations in tropospheric noise, or better account for this spatial variability in applying corrections and quantifying the uncertainty (e.g., Cao et al. (2021) and Murray et al. (2021)). A rigorous estimate of the data uncertainty is a key element of any probabilistic estimate of megathrust coupling, and is particularly important given the apparently small signal to noise ratio in the Makran region.

The above discussion applies to the uncertainties within individual tracks, but not the merged tracks. When merging tracks, we are attempting to take velocity fields that are expressed relative to a separate reference point within each track, and create a single velocity field that is relative to one reference point, i.e., putting all tracks into the same reference frame. With all tracks showing deformation relative to the same reference frame, we can interpret long wavelength signals across multiple tracks in terms of regional tectonics. The process of putting multiple InSAR tracks into the same reference frame is often done with the aid of GNSS (e.g., Weiss et al., 2020; X. Xu et al., 2021). While algorithms for combining multiple InSAR tracks without GNSS have been developed, they generally rely on the assumption that the LOS displacement does not vary significantly between overlapping tracks (e.g., Shirzaei (2015)).

As mentioned in Section 2.3.3, ground motion with the overlapping region between the tracks can bias the calculated offset due to the projection of that deformation into the different satellite lines of sight. This bias will then propagate as each subsequent track is merged, with the merged tracks before plate motion correction providing a clear example (Figures 2.9 and 2.13). These biases are effectively errors in reference frame estimation. Even after plate motion removal, our merged tracks contain biases in the offset estimation, for example due to the Balochistan post-seismic signal. It is hard to quantify the impact of these errors without comparison to dense GNSS stations; however, it seems reasonable that the bias could be in the range of 1-2 mm/yr from visual inspection of profiles crossing multiple tracks.

In the ideal case, if the reference region for every track is placed in a point that is stationary with respect to the overall plate, and subject to low noise, no offset between the tracks will actually be necessary, as each track will already be expressed relative to the plate and so can be directly compared. Given that we are unlikely to pick perfectly stationary and noise free reference regions, we could attempt to merge tracks by using only the overlap regions which are least affected by noise and ground deformation, which would likely reduce the biases that we see in our current merged velocity fields. Beyond approaches to merge InSAR tracks that are each relative to a separate reference, work on absolute phase change recovery in InSAR (X. Xu & Sandwell, 2020), and SAR geolocation (Cong et al., 2012; Eineder et al., 2011) may also prove useful for combining multiple SAR/InSAR tracks without the aid of GNSS.

2.4.2 Interpretation of Modeling Results

Based on comparisons between our linear elastic forward models and InSAR velocity profiles, we claim that scenarios where the fault is strongly coupled north of the coast are unlikely. The assumed convergence direction, combined with the two lines of sight from ascending and descending tracks, provide a useful tool for resolving strong coupling scenarios, as we would expect them to create resolvable differences in the ascending and descending track profiles. Given the conservative error bars we place on our velocity profiles, the coupling scenario Lin et al. (2015) presented for the eastern end of the subduction zone is possible at the western end, highlighting the importance of precise errors when measuring coupling.

There are several limitations to our model. Low seismicity and logistical challenges of working in the region have resulted in the megathrust geometry being relatively

poorly constrained, and the Slab2 model does not include data from recent studies (e.g., see Priestley et al. (2022), and references therein). While the slab geometry could be wrong, if the convergence direction and assumption of linear elastic deformation are reasonable approximations, we would still expect to be able to resolve high coupling north of the coast, as it would create differences between the ascending and descending tracks on the order of several millimeters per year.

The assumption that strain is accommodated purely elastically in the accretionary wedge is clearly an approximation, as evidenced by co-seismic and post-seismic deformation from the 2013 Balochistan and Minab earthquakes. If a substantial fraction of the convergence was accommodated by permanent deformation, that would reduce the amount of stored elastic energy that could be released in an earthquake. Haghipour et al. (2012) find that 3% of the Arabian-Eurasian convergence is accommodated by internal deformation of the accretionary wedge in the Iranian Makran, suggesting that the assumption of elasticity is not unreasonable. Our model could likely be improved by the use of a layered elastic half space, as was used by Lin et al. (2015).

We use the constant convergence velocity of Lin et al. (2015) (30 mm/yr, at 10°), based on results from DeMets et al. (2010) and Argus et al. (2010). However, this value relies on sparse GNSS stations, we would expect it to vary somewhat along the megathrust (Khorrami et al., 2019). Lin et al. (2015) explored the impact of varying the convergence direction on the fit to their data, finding that a westward component to the convergence gave a worse fit to the data at the eastern end of the subduction zone (see their supplementary Figure 8). Looking at the velocities over the coastal region of the western Makran in Figure 2.11 (around $(26^\circ\text{N}, 59^\circ\text{E})$), and the profiles in Figure 2.16, we can see areas near the coast where the ascending track velocity is larger than the descending track velocity (at 100 km in profile 2.16(b), 50 km in profile 2.16(c), and 100 km in profile 2.16(d)). Such a pattern could potentially indicate westward motion (i.e., moving away from the descending satellite, towards the ascending satellite), rather than the eastward motion predicted by our chosen convergence direction. These velocity differences are within our uncertainties, so we cannot interpret them in terms of deformation, but reducing the noise level with more data and better corrections may allow such differences between the ascending and descending tracks to be interpreted as variations in the convergence direction along strike. The overlap region between adjacent tracks of the same direction

could also be used to provide an additional look angle (assuming the tracks have been merged using a non-deforming part of the overlap region).

While our work suggests that a fully coupled megathrust in the western Makran is unlikely, this does not rule out a substantial rupture. Previous computational work (Noda & Lapusta, 2013), and simulations that we present in Chapter 4, illustrate how dynamic weakening can allow ruptures to propagate through areas of faults with low coupling. We therefore caution that constraining the coupling distribution alone is not sufficient to fully rule out larger magnitude events.

2.5 Conclusions

We test the ability of InSAR to constrain deformation over wide areas after corrections for the ionosphere, troposphere, solid Earth tides and plate motion. We study the Makran subduction zone, on the Iran-Pakistan border, revealing deformation from fault creep, co-seismic and post-seismic offsets, and aquifer depletion. Our results show the importance of plate motion corrections when combining multiple tracks. One notable observation is that of ongoing post-seismic deformation from the 2013 M_w 7.7 Khash earthquake, which ruptured the subducting slab at a depth of 80 km.

Tropospheric noise and post-seismic deformation results in us being unable to resolve deformation due to coupling on the megathrust. However, by comparing our velocity measurements with simple forward models we can place constraints on the likely degree of coupling. The combination of ascending and descending tracks is particularly useful for resolving coupling signals due to the differing sensitivities to east-west motion between the tracks and the eastward component of the convergence direction between the plates. Comparisons between our models and observations suggest that the western Makran is unlikely to be strongly coupled north of the coast. However, significant uncertainties remain, and we cannot rule out the possibility of a significant earthquake in the western Makran.

Future work could focus on rigorous handling of data errors, allowing more data to be used and a reliable uncertainty on the final velocity to be derived. This uncertainty could then be used for a probabilistic inversion of coupling. More data and improved tropospheric correction methods will further help to reduce the noise. Further work is also needed on the ionospheric correction method to automatically deal with unwrapping errors when processing large volumes of data.

DEEP LEARNING-BASED DAMAGE MAPPING WITH INSAR COHERENCE TIME SERIES

3.1 Introduction

In the wake of major natural disasters, emergency services need a rapid and accurate assessment of the damage over a wide area in order to quickly direct their response and estimate losses. However, damage to infrastructure and communications networks often makes prompt on-the-ground damage assessment difficult or impossible. Under these circumstances, remote sensing can either complement, or provide a useful alternative to, ground-based assessments (Voigt et al., 2016).

Assessments of damage due to a natural disaster can be obtained by comparing satellite observations from before and after the event. One approach is the use of change detection on very high resolution optical data ($\approx 50 \text{ cm} \times 50 \text{ cm}$ pixels) (Dalla Mura et al., 2008; Pesaresi et al., 2007). However, the utility of optical images for disaster response can be hampered by the need for timely data, requiring cloud free conditions and sufficient solar illumination (Brunner et al., 2010).

Satellite-based synthetic aperture radar (SAR) is an imaging technique that offers advantages over optical data by providing images in all weather conditions, day or night (e.g., Rosen et al. (2000) and Ulaby and Long (2015)). SAR relies on active imaging using microwave (centimeter-scale) wavelengths emitted by the satellite, with the sensor recording the amplitude and phase of the reflected radar pulse to produce images at meter-scale resolution. The availability of SAR images depends only on the orbital parameters of the satellite.

Damage detection using SAR relies on separating normal changes in the radar backscatter properties of the ground (e.g., due to agricultural activities, vegetation growth, snow, rainfall, and even vehicle motion in a car park) from anomalous changes attributed to disaster-induced damage. The changes can be quantified using the coherence of the radar echo between subsequent acquisitions (Zebker and Villasenor (1992), and see Eq. 3.1). One current method of mapping damage uses a pair of SAR acquisitions just before the event and a pair of acquisitions that span the event, allowing for a comparison between the amount of ground surface change without any damage to the amount of ground surface change that occurs during

the event (Yun et al., 2015). This method relies on human judgment in setting an appropriate threshold for classifying areas as damaged, usually assumed to be constant for all locations in the SAR scene, as well as for selecting suitable pre-event acquisitions (Yun et al., 2015).

There are now satellite SAR missions with revisits on a time-scale of days, and many parts of the Earth have repeat observations by the same satellite constellation going back several years. These developments allow for the possibility of using pre-event multi-year time series to separate out regularly occurring anthropogenic and natural surface changes from changes caused by a given natural disaster. These data have only recently begun to be exploited by researchers for damage mapping purposes (Jung et al., 2018; Karimzadeh et al., 2018; Washaya et al., 2018). The desire and opportunity to perform damage classification on large and complex SAR data sets motivates us to explore the use of deep learning techniques.

Deep learning relies on feeding input data through multiple layers of non-linear parameterized functions, also known as a deep learning architecture, to transform input data into desired outputs which can be used for regression or classification (3Blue1Brown, 2017; Goodfellow et al., 2016; LeCun et al., 2015). In supervised deep learning, the function parameters are optimised, during a process known as training, to minimise the misfit between the functions' output and known training data (*ground truth*). For example, in image classification, the function input is an image with a known classification (e.g., “dog,” “cat,” “tree,” etc.), and the function outputs are the probabilities of the image having each classification, with the set of possible classifications finite and fixed. The functions' parameters are optimized to maximize the probabilities assigned to correct classifications for images in a data set, and the final resulting function can then be used to classify previously unseen images. Generally, in supervised learning, the functions are trained using data from the *training set* and then evaluated on a separate data set, the *validation set*, which is unseen during training to ensure the learned functions represent generalizable rules, rather than just a memorization of the training set.

Deep learning has proven to be an effective way to extract insights from large data sets with little or no assumptions about the underlying data, and minimal human intervention (LeCun et al., 2015). The growing body of available satellite data has prompted recent work to combine satellite data with deep learning techniques to study volcanoes (Anantrasirichai et al., 2019), fires (Kong et al., 2018) and flooding (Y. Li et al., 2019) among other examples.

Recurrent neural networks (RNNs) are a type of deep learning architecture particularly well suited to dealing with sequential (e.g., time-ordered) data (Lipton et al., 2015; Olah, 2015). RNNs have been applied to a wide range of tasks, from predicting the next character in a word (Sutskever et al., 2014) to precipitation forecasting (Shi et al., 2015) and seismic phase association (Ross, Yue, et al., 2019). By training an RNN on a large number of previously observed time series, the network can be used to classify new time series observations and to forecast future time steps. When using an RNN for time series forecasting, the deviation between forecast values and observations can be used for anomaly detection (Malhotra et al., 2015).

The ability of RNNs to learn generalized rules from large time series data sets makes them a good candidate for application to large satellite time series observations, and RNNs have recently begun to be used on satellite data for tasks such as forecasting (Shi et al., 2015), classification (Ndikumana et al., 2018) and anomaly detection (Kong et al., 2018).

In this study, we frame the damage detection problem as one of detecting anomalies in sequential InSAR coherence time series. We train an RNN on a time series of sequential InSAR coherence data taken before a damage event, then use the trained RNN to make a probabilistic forecast for the co-event InSAR coherence (i.e., the coherence of the radar echo between pre- and post-event acquisitions). The probabilistic nature of the forecast allows us to capture the distribution of the coherence values we expect for each location in the absence of any damage. We then calculate the number of standard deviations of the forecast distribution between the forecast mean and the observed co-event coherence value for each point in the region of study. The number of standard deviations between the forecast mean and observed values is used to quantify how anomalous each coherence value is, with anomalously low coherence values attributed to damage. The use of a probabilistic forecast for each pixel allows us to create a location-dependent threshold for damage which depends on the specific time series characteristics of that location.

In what follows, we summarize the underlying SAR methodology and give a brief overview of previous work on using SAR for damage mapping (Section 3.2). We then discuss how damage mapping can be formulated in terms of a machine learning problem, and present our method for deploying recurrent neural networks for damage detection (Section 3.3). We apply our method to three earthquakes, which had either substantial building damage or surface rupture (Section 3.4):

- The August 24, 2016 M_w 6.2 central Italy earthquake
- The November 12, 2017 M_w 7.3 Iran-Iraq earthquake
- The July 2019 M_w 6.4 and M_w 7.1 Ridgecrest, California, USA earthquakes.

Through these examples, we illustrate how combining a long pre-event SAR time series with RNN-based anomaly detection can improve results compared to an existing SAR damage mapping method (Section 3.5). We discuss the strengths and limitations of our proposed method (Section 3.6) then present conclusions and outline potential further work (Section 3.7). Further details of our deep learning architecture as well as the satellite data, damage assessments, and example code used in this study are presented in the supplementary materials.

3.2 Background and Previous Work

3.2.1 Synthetic Aperture Radar

Synthetic Aperture Radar (SAR) is a coherent active imaging method operating at microwave wavelengths used for mapping the Earth’s surface (Ulaby & Long, 2015). The method relies on satellite-based illumination of the ground with 1–30 cm wavelength microwaves, then recording the amplitude and phase of the reflected wave. In our work, we begin with processed full-resolution data known as *single look complex* (SLC) images. Each SLC pixel in the image corresponds to a region on the Earth’s surface and records the amplitude and phase of the radar echo from that region. The reflected wave depends on the properties of the Earth’s surface, with the echo being a combination of the coherent sum of the backscatter from all of the reflectors within an SLC pixel, or *resolution element* (e.g., Section 3.12.2 of Simons and Rosen (2015)), as well as delays accrued during propagation through the atmosphere (e.g., Section 3.12.4.2 of Simons and Rosen (2015)).

3.2.2 Change Detection using Synthetic Aperture Radar

Changes in the imaging or viewing geometry, surface roughness, and dielectric properties of the ground within a resolution element will affect the measured radar return (Jordan et al., 2020; Zebker & Villasenor, 1992). For example, the collapse of buildings changes the path length travelled by the radar wave and randomly rearranges the radar reflectors within a given SLC pixel, leading to a random change in each SLC pixel’s phase.

Comparing SAR images of the same point on Earth from the same satellite taken at different times provides proxies for changes in the Earth's surface. These measurements can be classified as coherent or incoherent, depending on whether or not the SAR phase is used (Jung & Yun, 2020). In this study, we focus on coherent change detection where the change between two SAR acquisitions can be quantified by calculating the magnitude of the complex correlation coefficient, also known as the *interferometric coherence*, or simply *coherence*, between the two complex SAR signals. For a given SLC pixel, coherence is defined as:

$$\gamma_{i,j} = \left\| \frac{|\langle \Gamma_i \Gamma_j^* \rangle|}{\sqrt{\langle |\Gamma_i|^2 \rangle \langle |\Gamma_j|^2 \rangle}} \right\|, \quad (3.1)$$

where Γ_i is the complex amplitude and phase for SAR acquisition at time step i , $*$ represents complex conjugation, and $\langle \rangle$ denotes an ensemble average, generally approximated as a local spatial average (e.g., see Section 3.12.2.5 of Simons and Rosen (2015)). $\gamma_{i,j}$ is known as the coherence of the signal between SAR acquisitions at time steps i and j . This measure incorporates information about changes in both the amplitude and phase of the SAR signal. The coherent nature of SAR means that it is possible to sense changes on the scale of the radar wavelength (1-30 cm) when using phase information, allowing for very sensitive change detection compared to most optical data.

The use of a local spatial average in the coherence calculation means that the resolution of the coherence image is necessarily lower than the original SAR SLC image, as multiple pixels in the SLC image (SLC pixels) are used to calculate a single pixel in the coherence image (coherence pixel). Unless stated otherwise, the term *pixel* refers to coherence pixels for the rest of this paper.

For completely coherent echos $\gamma_{i,j} = 1$, whereas $\gamma_{i,j} = 0$ implies that the two echos are completely uncorrelated (a low or zero coherence value is also known as decorrelation). Stable, concrete structures, for example, will reflect radiation in the same manner through time, and thus exhibit high coherence, whereas bodies of water, which change their radar scattering properties on a time-scale of less than a second, will completely decorrelate (Bamler & Hartl, 1998).

Coherence for a given pixel will tend to decrease with the time between SAR acquisitions due to natural changes in the Earth's surface properties, with the rate of decrease depending on the rate of change of the Earth's backscattering properties at length scales similar to the radar wavelength (Zebker & Villasenor, 1992). However,

the presence of seasonal effects such as snow can also lead to seasonal coherence variations as the ground surface is covered and uncovered, and rainfall can lead to sudden drops in coherence (Jordan et al., 2020). The time between the two acquisitions, known as the temporal baseline, is therefore an important indicator of how much coherence loss to expect. Increasing the spatial separation, known as the spatial baseline, between the SAR sensor's image acquisition position for repeat images will also lead to a decrease in coherence (Zebker & Villasenor, 1992). Currently orbiting satellites have tight orbital control, such that spatial baseline decorrelation is a less significant problem than it was for previous generations of sensors.

A spatial image of coherences calculated from two SAR acquisitions, acquired at different times, allows for mapping of changes in the Earth's surface properties, on the scale of the radar wavelength. For example, Simons et al. (2002) used the spatial pattern of low coherence to map the location of fault surface rupture due to the 1999 M_w 7.1 Hector Mine, California earthquake. However, decorrelation effects from regularly occurring natural processes often occur together with those induced by damage events (Jung et al., 2018). Within a coherence image spanning an earthquake (co-event coherence), we may detect decorrelation due to collapsed buildings as well as, for example, agricultural activity, vegetation growth and the changing position of vehicles in a car park, making the isolation of damage effects challenging.

The need to separate changes in surface properties due to damage from other changes motivates the framing of this problem as one of anomaly detection. If we are able to identify the nominal distribution of coherence (at a given temporal baseline) for each pixel before any damage has occurred, we can then identify which pixels have an anomalously low co-event coherence with respect to their pre-event distribution and use the presence of anomalous coherence as a proxy for damage. This nominal distribution may be a complicated function of underlying physical properties, and may not be stationary in time.

One way to characterize the pre-event coherence is to calculate the coherence between two SAR SLCs acquired as close as possible before the event. The co-event coherence can then be compared to the pre-event coherence and the magnitude of the relative coherence loss can be used to identify areas where the coherence has dropped anomalously. Generally, a threshold for the amount of coherence loss required for a pixel to be marked as damaged is chosen manually, by including areas where it

is known that no damage occurred and setting the threshold so that these undamaged areas are correctly classified. This method is sometimes known as Coherence Change Detection (CCD) (e.g., see Bouaraba et al. (2012), Fielding et al. (2005), Geudtner et al. (1996), and Yun et al. (2015) and Fig. 3.1), and is based on the assumption that the calculated pre-event coherence image is a good representation of the normal pre-event coherence. In cases where coherence between sequential SAR acquisitions has a high variance (i.e., there is a lot of variation in the amount of surface change for a given temporal baseline), a single pre-event coherence image will not be a good characterization of the pre-event coherence distribution for the given temporal baseline, and the CCD damage map is likely to be noisy.

To better characterize the pre-event coherence, researchers have begun using the long time series of regular SAR acquisitions that are increasingly available (Jung et al., 2018; Karimzadeh et al., 2018). By calculating the coherence between sequential SAR acquisitions, the mean and standard deviation of the sequential pre-event coherence can be calculated for each pixel. The number of standard deviations between the mean pre-event coherence and the co-event coherence can then be used to detect anomalous co-event decreases in coherence (Olen & Bookhagen, 2018; Washaya et al., 2018). These methods rely on characterising the pre-event coherence with a single distribution through time for each pixel, which can cause problems when the coherence distribution varies substantially through time, for example due to changing precipitation with the seasons.

Additional information can be gained by calculating the coherence between all possible SAR pairs, leading to coherence images with a wide range of temporal baselines (Jordan et al., 2020; Monti-Guarnieri et al., 2018). These coherence values can be used to estimate the parameters, for each pixel, of a model for the various contributors to temporal decorrelation (Jung et al., 2016; Jung et al., 2018). This model can then be used to detect anomalies in coherence images which span the event. Similar to the mean and standard deviation method, these methods generally rely on inferring a single set of physical parameters for each pixel, without taking into account the possible variation of these parameters through time.

Supervised machine learning has also been used for damage mapping with SAR data, using comprehensive damage assessments, often available several months after major events, as ground truth to train damage classifiers (Endo et al., 2018; Y. Li et al., 2019; Wieland et al., 2016). While supervised machine learning approaches avoid the problem of manually selecting a uniform damage threshold in CCD, they

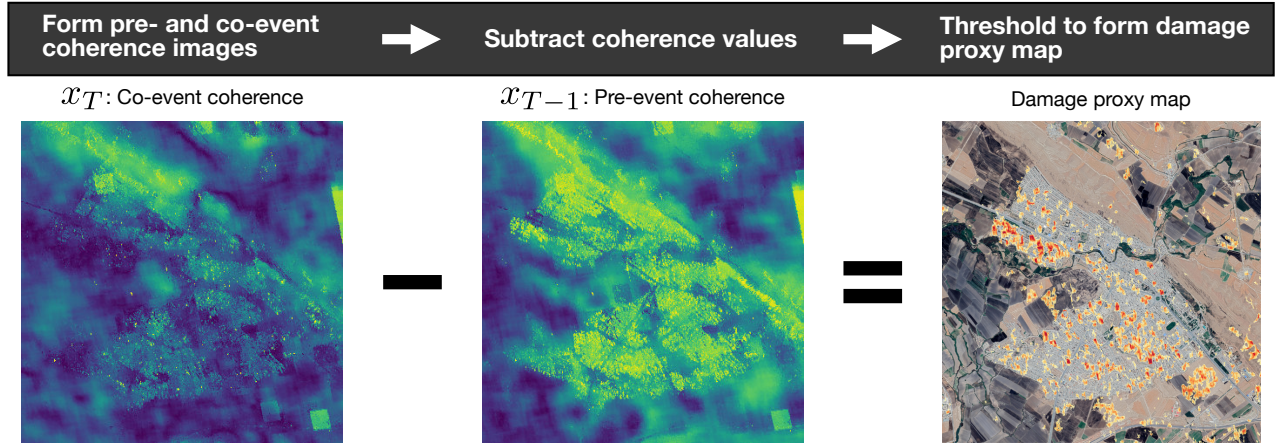


Figure 3.1: Schematic of the existing Coherence Change Detection (CCD) method for damage mapping (Yun et al., 2015), presented for the town of Sarpol-e-Zahab, damaged during the November 2017 Iran-Iraq earthquake. A pre-event coherence image (x_{T-1}) is subtracted from the co-event coherence image (x_T) in order to calculate the coherence loss. The coherence loss is thresholded and plotted to produce a damage proxy map. Optical data from Google, CNES/Airbus, taken July 27th 2020.

rely on extensive ground truth damage assessment data for training. Additionally, if the damage classifiers are to be useful for future events, the trained classifiers must be applied to new areas and it is unclear to what extent this training readily transfers to totally different regions of the Earth’s surface.

In our work, we seek to make use of all available SAR data before an event in order to make a deep learning-based, time-dependent forecast of a co-event coherence distribution that we would expect without any damage event. This approach allows us to detect anomalous changes in coherence. As we only use ground truth damage data to quantify our damage detection algorithm, and not for training, our method does not depend on ground truth damage data.

3.3 Proposed Approach

3.3.1 Notation

We have a total of $T + 2$ SAR acquisitions, ordered in time and indexed from 0 to $T + 1$; the last acquisition, $T + 1$, is post-event, while all others are pre-event. Between all pairs of consecutive acquisitions we compute the coherence values which we map to an unbounded space using a logit transform on the squared coherence (discussed below, see Eq. 3.9). We write the coherence between time steps $t - 1$ and t as $\gamma_{t,t-1}$ and the transformed coherence as x_t . Throughout the rest

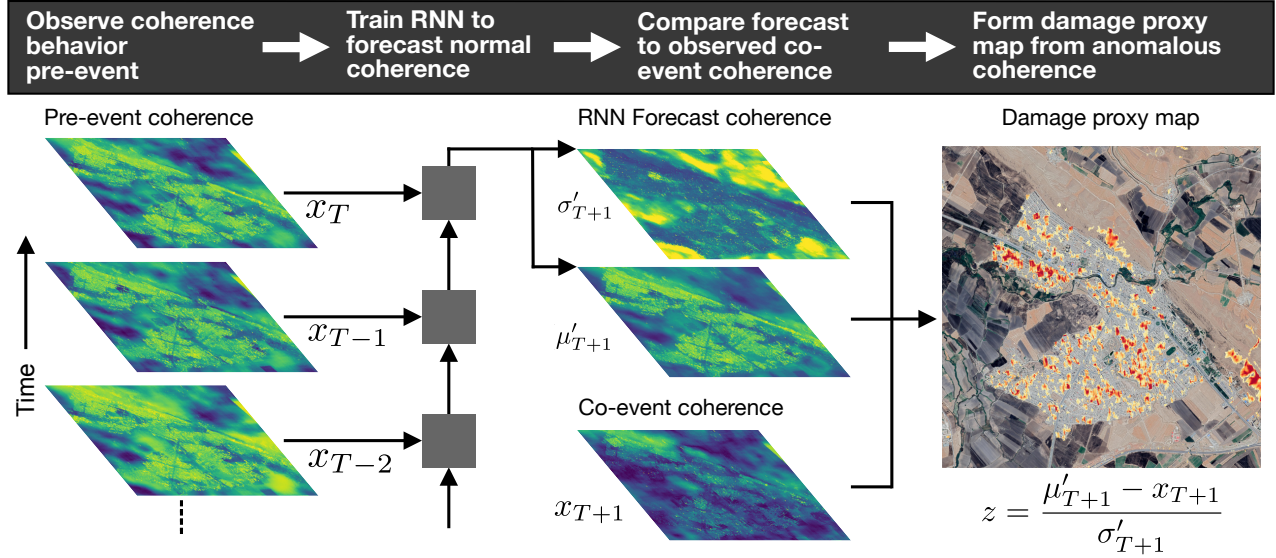


Figure 3.2: Schematic of our proposed recurrent neural network (RNN) method presented for the town of Sarpol-e-Zahab, damaged during the November 2017 Iran-Iraq earthquake. The transformed coherence values (x) are used to train a recurrent neural network to make a Gaussian forecast of the co-event coherence with mean μ'_{T+1} and standard deviation σ'_{T+1} . The forecast is compared with the observed co-event coherence, x_{T+1} , to calculate the z-score, z (see Eq. 3.10). The z-score is thresholded and plotted to produce a damage proxy map. A more detailed illustration of the neural network architecture can be found in Fig. B.1. Optical data from Google, CNES/Airbus, taken July 27th 2020.

of the paper, *coherence* refers to the transformed coherence unless otherwise stated. Let $x_{\leq T} = \{x_1, \dots, x_T\}$ denote the sequence of T pre-event sequential coherence values for a given coherence pixel location, with x_{T+1} the co-event coherence. Additionally, let $\mathcal{D}_t$ denote the collection of M coherence sequences that we have available for training (i.e., coherence sequences from M different coherence pixels), each containing T pre-event coherence values. We also have $\mathcal{D}_f$, the set of coherence sequences on which we wish to perform forecasting and anomaly detection for x_{T+1} , which contains T pre-event ($x_{\leq T}$) and one co-event (x_{T+1}) coherence values for each pixel.

Our goal is to train a model that can capture the range of possible behaviors across time for the sequences in $\mathcal{D}_t$. We can then use the model to make a forecast, x'_{T+1} (we use $'$ to denote a forecast value), for each co-event coherence pixel in $\mathcal{D}_f$, based on the coherence time series of that pixel. We can detect anomalies by comparing the forecast with the ground truth co-event coherence value, x_{T+1} , at each pixel, mapping anomalous changes in ground surface properties that have occurred between SAR

acquisitions T and $T + 1$. Note that the model does not see x_{T+1} , or any damage data, during training.

3.3.2 Recurrent Neural Networks

Motivated by the sequential nature of our data, we use a recurrent neural network (RNN) as our model for forecasting the coherence time series. RNNs are a class of models frequently used on sequential data for machine learning tasks such as speech recognition, machine translation, motion tracking and time series classification and forecasting (Lipton et al., 2015; Olah, 2015). RNNs maintain a fixed-length hidden state vector, h_t , that summarizes a sequence up to time t , and is updated at every time step with observations:

$$h_t = f_\phi(h_{t-1}, x_t), \quad (3.2)$$

where f is a deterministic function, learned during training and parameterized by ϕ , and x_t is the transformed coherence (Eq. 3.9), calculated from SAR data, at time step t for a given pixel. Forecasting future values involves another function g parameterized by ψ :

$$x'_{t+1} = g_\psi(h_t). \quad (3.3)$$

In general one can optimize for parameters ϕ and ψ to minimize some loss, or cost function, between the model forecast and the coherence ground truth (here the transformed coherence values, see Eq. 3.9), for example the mean-squared error:

$$\phi^*, \psi^* = \arg \min_{\phi, \psi} \sum_{x \leq T \in \mathcal{D}_t} \sum_{t=1}^T (x_t - x'_t)^2. \quad (3.4)$$

RNNs use neural networks as function approximators for f and g , and this optimization can be solved with some form of gradient descent (e.g., Kingma and Ba (2014)). See Section B.2 for more details.

3.3.3 Probabilistic Formulation

We aim to forecast the probability distribution over all possible values given the pre-event coherence values. This probabilistic forecast lets us evaluate the probabilities that our model assigns to the coherence values that are actually observed. We locate anomalies by identifying coherence values that have a low probability of occurring given the previous observations. To make a probabilistic forecast, we modify g in Eq. 3.3 to output the parameters of a probability distribution instead of a single

value. In this work we use a Gaussian output probability, so we have:

$$[\mu'_{t+1}, \sigma'_{t+1}] = g_\psi(h_t), \quad (3.5)$$

where μ'_{t+1} and σ'_{t+1} are the forecast mean and standard deviation, respectively. The probability our model assigns to the ground truth co-event coherence, x_{t+1} , is then:

$$p(x_{t+1}; \mu'_{t+1}, \sigma'_{t+1}) = (2\pi\sigma'^2_{t+1})^{-\frac{1}{2}} \times \exp\left(-\frac{(x_{t+1} - \mu'_{t+1})^2}{2\sigma'^2_{t+1}}\right). \quad (3.6)$$

Instead of minimizing the mean-squared error, as shown in Eq. 3.4, the probabilistic forecast allows us to maximise the probability that our model assigns to ground truth sequences in $\mathcal{D}_t$. Computationally, probability maximization is best achieved by minimizing the negative log-likelihood that the model assigns to ground truth sequences in $\mathcal{D}_t$:

$$\begin{aligned} \phi^*, \psi^* &= \arg \min_{\phi, \psi} \sum_{x_{\leq T} \in \mathcal{D}_t} -\log_e p(x_{\leq T}) \\ &= \arg \min_{\phi, \psi} \sum_{x_{\leq T} \in \mathcal{D}_t} \sum_{t=1}^T -\log_e p(x_t | x_{< t}), \end{aligned} \quad (3.7)$$

where in the second step we factorize the conditional probabilities using the relationship $p(x_{\leq T}) = \prod_{t=1}^T p(x_t | x_{< t})$. In our case, all of the information from previous elements in the time series is summarised by the μ' and σ' terms given by the forecast in Eq. 3.5, so we have that $p(x_t | x_{< t}) = p(x_t; \mu'_t, \sigma'_t)$. Therefore, combining Eq. 3.6 and Eq. 3.7 we have:

$$\begin{aligned} \phi^*, \psi^* &= \arg \min_{\phi, \psi} \\ &\sum_{x_{\leq T} \in \mathcal{D}_t} \sum_{t=1}^T \left(\frac{1}{2} \log_e (2\pi\sigma'^2_t) + \frac{(x_t - \mu'_t)^2}{2\sigma'^2_t} \right). \end{aligned} \quad (3.8)$$

Eq. 3.8 gives us a loss function that takes into account both the mean and standard deviation of the forecast, allowing us to optimize a probabilistic forecast for the coherence. We can optimize for the parameters in Eq. 3.8 using some form of gradient descent (e.g., Kingma and Ba (2014)).

Note that a Gaussian distribution assigns nonzero probability everywhere in $\mathbb{R}$, a distribution that is inconsistent with coherence as defined in Eq. 3.1, which is by definition bounded (i.e., $\gamma_{t-1,t} \in [0, 1]$). We therefore transform the coherence

to an unbounded space before training the RNN. We choose the inverse-sigmoid transform (also known as the logit transform) on the square of the coherence values:

$$x_t = S^{-1}(\gamma_{t-1,t}^2) = \log_e \left(\frac{\gamma_{t-1,t}^2}{1 - \gamma_{t-1,t}^2} \right), \quad (3.9)$$

which maps the domain from $[0, 1]$ to $(-\infty, \infty)$. This choice of transform is motivated by near mathematical equivalence between the logit transform of coherence squared and the logarithm of the variance of the interferometric phase, see Section B.1 for more details. We refer to the new unbounded space as the logit space. Our model will then forecast Gaussian distributions over the unbounded logit space.

3.3.4 Model and Training Details

The RNN model we use in this work (represented by f_ϕ in Eq. 3.2) is called a gated recurrent unit (GRU), chosen for its ability to learn long-term dependencies in time series (Cho et al., 2014). The hidden state output from f_ϕ is fed into a feed-forward neural network, represented by g_ψ (Eq. 3.5) which then outputs the parameters of the forecast distribution. To find the optimum model parameters (Eq. 3.8), we train the model using the Adam optimizer (Kingma & Ba, 2014). See Section B.2 for more details of the model and training, as well as further references. An implementation of our deep learning model can be found on GitHub: <https://github.com/olliestephenson/dpm-rnn-public>.

3.3.5 Anomaly Detection for Co-event Coherence

To construct a proxy for damage we normalize the difference between the forecast co-event mean, μ'_{T+1} , and the observed co-event coherence, x_{T+1} , by the standard deviation of the forecast σ'_{T+1} . This quantity is termed the z-score, which we define as:

$$z = \frac{\mu'_{T+1} - x_{T+1}}{\sigma'_{T+1}}. \quad (3.10)$$

Note that we have switched the order of μ'_{T+1} and x_{T+1} terms compared to usual definition of the z-score. With this definition, a large positive z-score implies that the coherence is many standard deviations below the forecast coherence, i.e., we have an anomalous drop in coherence, possibly due to damage. We use this definition of the z-score as the basis of our proxy for damage.

3.4 Data

3.4.1 Coherence Calculation

We use data from the Copernicus Sentinel-1 satellites, a pair of C-band SAR satellites operated by the European Space Agency. We download freely available Level-1 Single Look Complex (SLC) images acquired in interferometric wideswath mode (European Space Agency, [n.d.](#)). We then create a coregistered stack of SLCs covering the region of interest. To generate coherence values, as defined in Eq. 3.1, we average over a rectangle, or chip, of SLC pixels. In this case we use a chip of 15 SLC pixels in range (across the satellite track) and 5 in azimuth (along the satellite track) corresponding to a region of approximately 50 m by 70 m. Note that the resolution in range (across-track) of Sentinel-1 SLCs is higher than in the azimuth (along-track) direction. As stated above, the use of a chip to calculate coherence means that the coherence map is lower resolution than the SLC image as each coherence pixel contains information from a 50 m by 70 m area.

For each study area, we produce two separate coherence data sets: one for training the network ($\mathcal{D}_t$), and one for forecasting purposes ($\mathcal{D}_f$). We construct the training data set to have a large number of pixels drawn from a wide area surrounding the area of interest, while the forecasting data set focuses just on the area of interest to be mapped. More details on how these data sets are constructed can be found in Section B.3.

3.4.2 Study Areas

In this study, we consider three earthquakes:

August 24, 2016 M_w 6.2 central Italy earthquake

This event destroyed much of the town of Amatrice in central Italy. (United States Geological Survey, [2016](#)). The Copernicus Emergency Management Service produced a damage map assessing the damage level of every building in the town (Copernicus Emergency Management Service, [2016](#)). This comprehensive damage assessment allows us to quantitatively validate the RNN and CCD methods against the known damage levels.

November 12, 2017 M_w 7.3 Iran-Iraq earthquake

This event damaged the city of Sarpol-e-Zahab on the Iran-Iraq border (United States Geological Survey, [2017](#)). The United Nations Institute for Training and

Research (UNITAR) produced a damage map for Sarpol-e-Zahab in the wake of the earthquake (United Nations Institute for Training and Research, 2017), allowing for a qualitative test of our damage proxy map and comparison with the CCD method.

July 2019 M_w 6.4 and M_w 7.1 Ridgecrest, California, USA earthquakes

To explore the ability of our method to capture other forms of anomalous ground disturbances, we also consider the Ridgecrest earthquakes which struck the Mojave desert, California, in early July 2019. The Ridgecrest sequence contained two earthquakes with substantial surface rupture tens of kilometers long: an M_w 6.4 event on July 4th, and, 34 hours later, an M_w 7.1 event (Kendrick et al., 2019; Ponti et al., 2020; Ross, Idini, et al., 2019). The earthquakes also caused liquefaction, small rock falls and minor damage to buildings (Brandenberg et al., 2019; Hough et al., 2020; Zimmaro et al., 2020). The mapping of surface ruptures and location of liquefaction allows us to qualitatively compare the damage map to the location of known ground surface changes.

In Section B.3, we give more detailed information about these three earthquakes and the available data for each event.

3.5 Results

We present damage proxy maps for the coherence change detection (CCD) and our proposed RNN methods, then use available independent damage data to validate the efficacy of each method. For each method, we calculate a numerical damage proxy for every pixel, then threshold that damage proxy to create damage proxy maps. For Sarpol-e-Zahab and Ridgecrest, the limited nature of the ground truth data only allows for a qualitative comparison between the methods. For Amatrice, however, more comprehensive ground truth allows us to carry out a quantitative comparison. For Sarpol-e-Zahab, we also explore the forecasts the RNN makes through time for pixels in different locations. We find that the RNN method yields qualitative and quantitative improvements over the the CCD method.

August 24, 2016 M_w 6.2 earthquake, Amatrice, Italy

Fig. 3.3 shows the RNN method applied to mapping the damage in the town of Amatrice due to the August 24, 2016 central Italy earthquake. We use ground truth damage data from the Copernicus Emergency Management Service (Copernicus Emergency Management Service, 2016) to choose an optimum threshold for damage (discussed below), and mask values below that threshold. Details of the damage

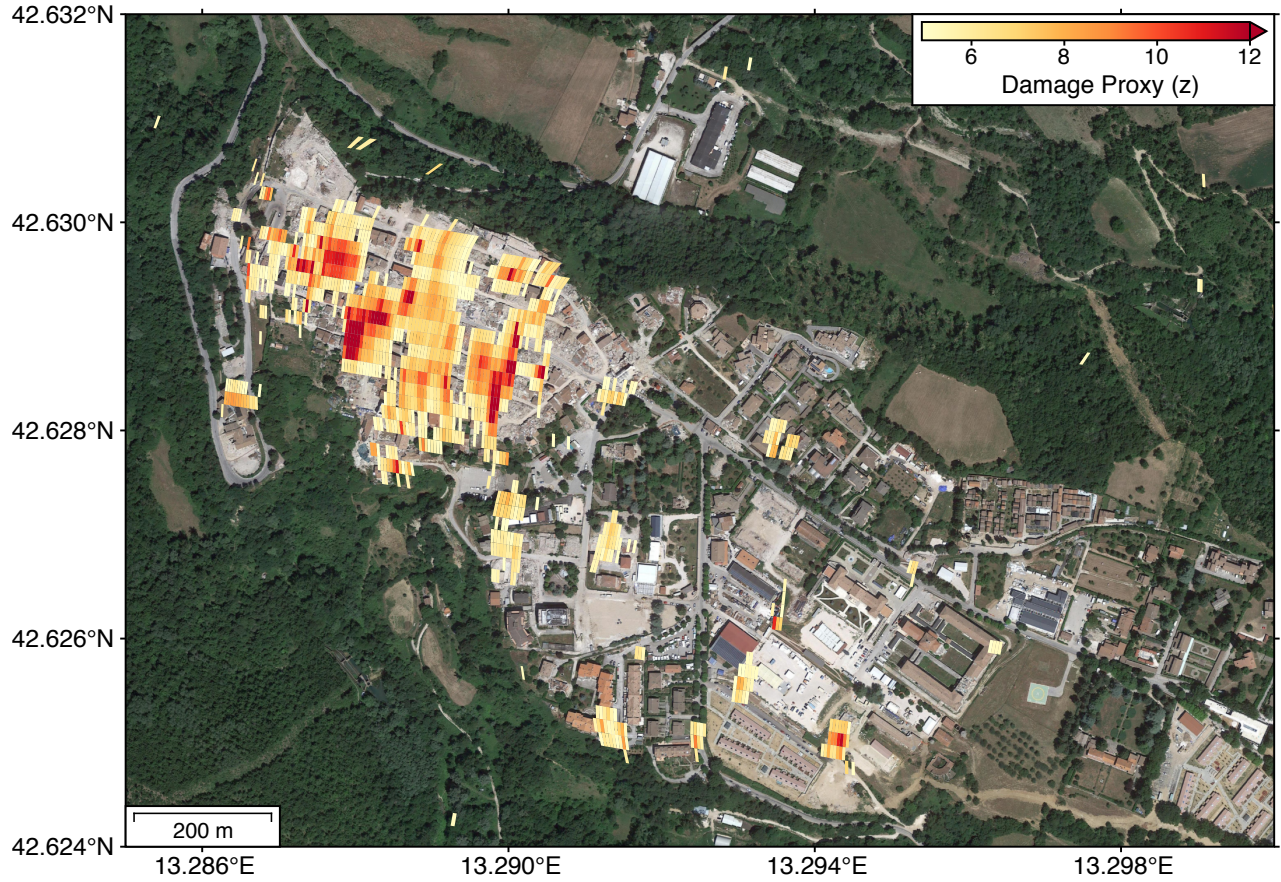


Figure 3.3: RNN DPM for the town of Amatrice, Italy, badly damaged during the 2016 M_w 6.2 central Italy earthquake. The center of the town, which was largely destroyed, is clearly highlighted by elevated damage proxy values towards the top left of the map. Z-score values below 4.93 (chosen from the $F_{0.5}$ score, Eq. 3.11) are masked, values above 12 are set to red as indicated by the color bar. Ground truth damage data are presented in Figure B.3. Optical imagery from Google, taken July 6th 2017.

data are presented in Section B.3. The ground truth damage data also allows for a direct quantitative comparison between the CCD and RNN methods.

We seek to classify each coherence chip as either “damaged” or “undamaged” and compare the classification ability of the CCD and RNN methods. Comparing the classifiers relies on assigning each chip a score (the z-score for the RNN method and the coherence loss for the CCD method), setting a threshold for damage, and then comparing the damaged/undamaged classifications with the ground truth. For every set of classifications, we have four categories: assigned damaged and truly damaged (true positive), assigned damaged but actually undamaged (false positive),

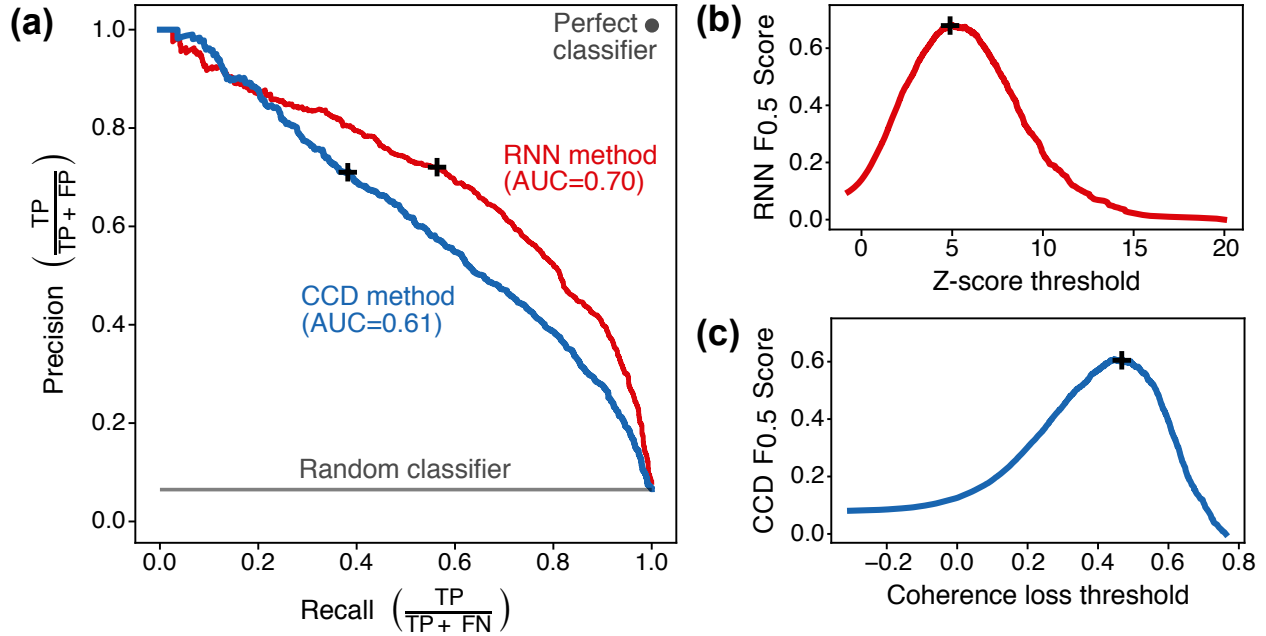


Figure 3.4: Quality metrics for the CCD and RNN damage classification methods. **(a)** Precision-recall curves for Amatrice damage proxy maps using the CCD (blue line) and RNN (red line) methods. For a perfect classifier we can choose a threshold that gives precision and recall equal to one, indicated in the top right corner. A random classifier gives a constant precision, equal to the fraction of the data set that is truly damaged, with recall varying with the threshold, indicated by the grey horizontal line at the bottom of the plot. The larger area under the curve (AUC) for the RNN method indicates improved performance. The black crosses show the position of maximum $F_{0.5}$ score identified in figures **(b)** and **(c)**. **(b)** $F_{0.5}$ score (see Eq. 3.11) for varying z-score damage thresholds using the RNN method. **(c)** $F_{0.5}$ score for varying coherence loss damage thresholds using the CCD method.

assigned undamaged and truly undamaged (true negative) and assigned undamaged but actually damaged (false negative).

In this case the ground truth damage data are building footprints, each with a damage score, which we separate into “damaged” and “undamaged” classes (see Section B.3). To determine the ground truth associated with each coherence chip, we calculate the proportion of each chip’s area that is occupied by damaged buildings. Coherence chips that have at least one third of their area (roughly 1200 m², see the discussion in Section 3.6) occupied by the footprints of damaged buildings, we assign to be truly damaged. Note that as the radar is side-looking, the 3D nature of the buildings means their radar footprint does not exactly match their ground footprint, a fact that we do not take into account in this work.

To compare the two methods quantitatively, we use a standard precision-recall curve (Davis & Goadrich, 2006) (Fig. 3.4). We calculate precision (the fraction of chips classified as damaged that are actually damaged) and recall (the fraction of truly damaged chips that are classified as damaged) for a range of damage proxy thresholds for each method. For a general classifier that is imperfect but better than random, precision and recall will trade off against one another. For example, with a high threshold, only a few points will be classified as damaged, and many of these will be truly damaged, leading to a high precision. However, with a high threshold the recall is low as most truly damaged points are incorrectly classified as undamaged. A low threshold means that many of the damaged points are above the threshold; however, there are also many false positives, leading to low precision and high recall. As our classes are unbalanced (there are many more undamaged points than damaged points) the precision-recall curve is preferred over the receiver operating characteristic (ROC) curve that is also used to assess the quality of classifiers (Davis & Goadrich, 2006).

Different classification methods can be quantitatively compared by calculating the area under the precision-recall curve (known as PR AUC). A perfect classifier will have an area of unity, with better algorithms having PR AUCs closer to this value. The PR AUC for a random classifier will be equal to the fraction of the data set that is truly damaged. Note that the PR AUC is distinct from the ROC AUC which is also used to compare classifiers (Berrar & Flach, 2012). Our PR AUC results presented in Fig. 3.4(a) show a clear quantitative improvement when using the RNN method over the CCD method, with a PR AUC of 0.70 for the RNN method and 0.61 for the CCD method. We achieve this improvement using the RNN method in spite of the relatively poor quality training data (see discussion in Section 3.6).

To compute the optimum threshold for damage for each method, we can use the F_β score, which is the weighted harmonic mean of the precision and the recall, computed as:

$$F_\beta = (1 + \beta^2) \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}}. \quad (3.11)$$

F_β will vary with the threshold, and we can choose a threshold that maximizes the score. This weighting considers recall β times as important as precision. In our case we set $\beta = 0.5$ and thus compute the $F_{0.5}$ score for all possible thresholds for both methods. Our choice of β weights precision as twice as important as recall, based on the assumption that we wish to direct finite emergency response resources to the places most likely to be damaged and thus favor higher precision at the expense

Table 3.1: Optimum threshold and corresponding precision and recall values for both methods using the Amatrice data set, selected using the maximum value of the $F_{0.5}$ score, along with the area under the precision-recall curves (PR AUC).

Method	Optimum threshold	Optimum precision	Optimum recall	PR AUC
RNN	4.93	0.72	0.56	0.70
CCD	0.47	0.71	0.38	0.61

of lower recall. At the maximum $F_{0.5}$ values, both methods have a precision just over 0.7, meaning over 70% of the points classified as damaged are truly damaged; however, the RNN method has recall of 0.56, compared to 0.38 for the CCD method, a clear quantitative improvement. The $F_{0.5}$ scores are presented in Fig. 3.4(b) and (c), and the optimum threshold and corresponding precision and recall for each method are given in Table 3.1.

Note that we perform the quantification on a pixel-by-pixel basis (using coherence pixels) rather than a building-by-building basis. Building areas can vary greatly, and, for the same level of damage, a small building and a large building can have very different effects on the coherence. Therefore a building-by-building quantification would combine metrics with very different sensitivities, whereas in theory each coherence chip should respond in a more similar way when a given fraction of its area is occupied by buildings with the same level of damage. We also note that in actual deployment scenarios, building footprints may not be available, and we may also be interested in investigating other forms of anomalous surface change (for example fault ruptures and landslides).

November 12, 2017 M_w 7.3 earthquake, Sarpol-e-Zahab, Iran-Iraq Border

In Fig. 3.5, we present damage proxy maps for the town of Sarpol-e-Zahab, damaged during the Iran-Iraq earthquake of November 12, 2017, for both the CCD and RNN methods. Using results from Amatrice (Table 3.1) the threshold for damage is set at $z = 4.93$ for the RNN method, and coherence loss = 0.47 for the CCD method, with damage proxies below these thresholds masked out. For display purposes we again choose the upper limit of the RNN color bar to be $z = 12$ and set the limit of the CCD color bar such that the same number of points are above the color bar limit as for the RNN case to provide a fair visual comparison. The damage data from the UN (United Nations Institute for Training and Research, 2017) (Figure B.4) allows us to qualitatively compare the damage maps to the documented damage in the city.

Within the city, both methods highlight neighborhoods where the UN located many collapsed buildings (for example in the northwest of the city) However, they also have elevated damage proxies over areas in the city where the UN did not record damage, possibly due to the sensitivity of InSAR coherence to small changes in surface properties, and possibly due to significant damage that was missed in the UN damage map. This seems to be more significant for the RNN method, which finds a larger amount of damage in the city than the CCD method.

Looking outside the city yields a clearer difference between the methods. In Fig. 3.5, we use white dashed lines to highlight several areas where the CCD method has agricultural fields outside the city with high damage scores that are no longer highlighted in the RNN damage map, indicating that the co-event coherence for these areas was within the bounds of the normal variability for those pixels. This difference between the methods indicates the advantage of taking into account the full temporal behavior of each pixel. In Fig. 3.5, we also highlight one area outside the city where the RNN method has a higher damage proxy than the CCD method. As this area is over a rocky ridge, it is possible that the RNN damage proxy is capturing surface change due to rockfalls caused by the earthquake shaking.

To better understand the damage map produced by the RNN, we select four locations in and around the town that show different styles of coherence time series. We apply this analysis to Sarpol-e-Zahab due to the larger amount of pre-event data compared to Amatrice, and the wider variety of pixel behaviors in a small area compared to Ridgecrest. In Fig. 3.6, we present the coherence time series, as well as the mean and standard deviation of the RNN coherence forecast and resulting z-score through time. For each case, the forecast at each time step is made based on Eq. 3.5, using the hidden state that is output from the trained model with the input being the coherence time series at that pixel up to that time.

For pixel **(a)** of Fig. 3.6, over a rocky ridge, we see a high, stable coherence through time, with no substantial drop in coherence co-seismically, hence a low co-event z-score. Pixel **(b)** is over a river, where surface properties change rapidly between SAR acquisitions, hence the pixel has a low coherence and higher uncertainty, but again has no co-event drop in coherence. Pixel **(c)** is within one of the badly damaged areas of the town. The pre-event coherence is high and relatively stable through time, causing a narrow uncertainty in the forecast. The co-event coherence is around 19 standard deviations below the forecast ($z \approx 19$), implying that the pixel is well out of the bounds of normal behavior. We infer that this is due to building damage. Finally,

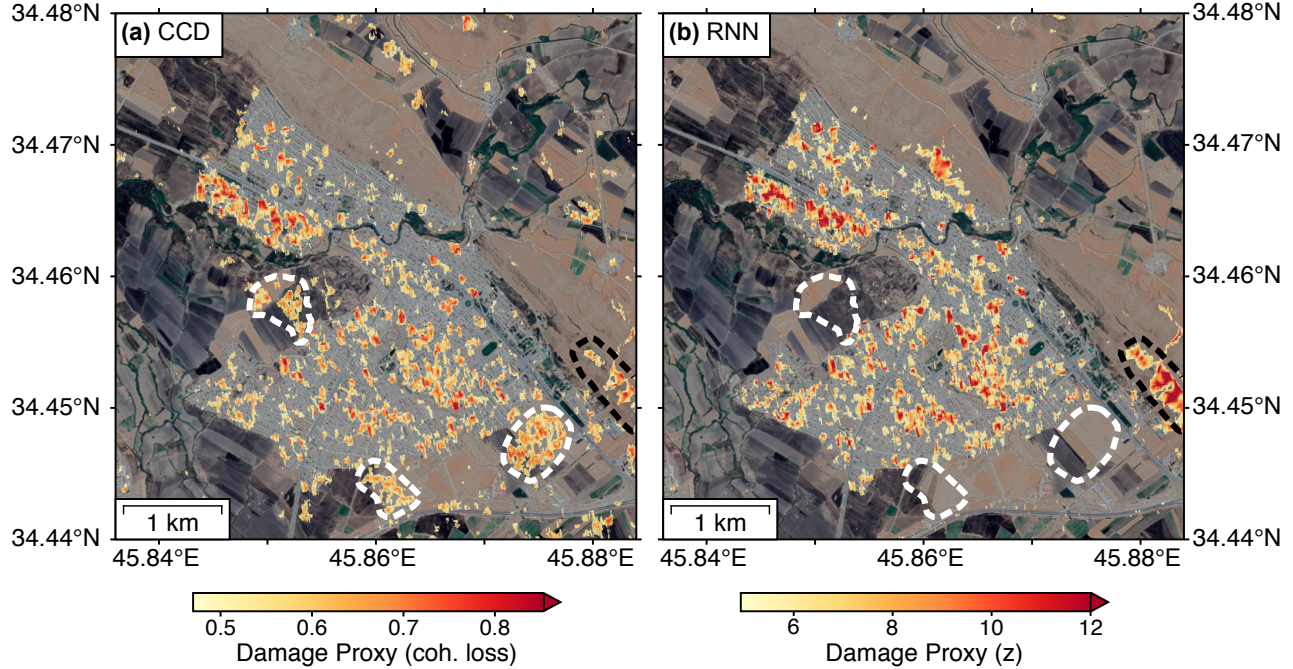


Figure 3.5: Comparison of CCD (a) and RNN (b) DPMs for Sarpol-e-Zahab, damaged in the November 12, 2017 M_w 7.3 Iran-Iraq earthquake. For each plot we mask values below a threshold, with the threshold chosen using the maximum value of the $F_{0.5}$ curve for the Amatrice data set (see Table 3.1). The upper threshold of the color scales are chosen such that both plots have the same number of points above the threshold. The white dashed lines highlight example areas where the CCD method gives false positive damage detection in regions outside of the city that are no longer classified as damaged by the proposed RNN method. The black dashed line shows an area over a rocky ridge where greater damage is shown by the RNN method compared to the CCD method. Ground truth damage data are presented in Figure B.4. Optical imagery from Google, CNES/Airbus, taken July 27th 2020.

pixel (d), covering an agricultural field, has highly variable coherence through time, which causes a large variance in the forecast coherence. The co-event coherence is substantially below the final pre-event coherence, meaning the CCD method shows elevated damage proxy values. However, in the context of the entire time series we see that this coherence is well within the bounds of the forecast coherence variability and thus the z-score is small.

July 2019 Ridgecrest earthquakes, California, USA

The earthquakes that struck near the town of Ridgecrest, California, in July 2019 (Ross, Idini, et al., 2019) provide an opportunity to test our proposed method on other forms of anomalous changes in ground surface properties, including fault

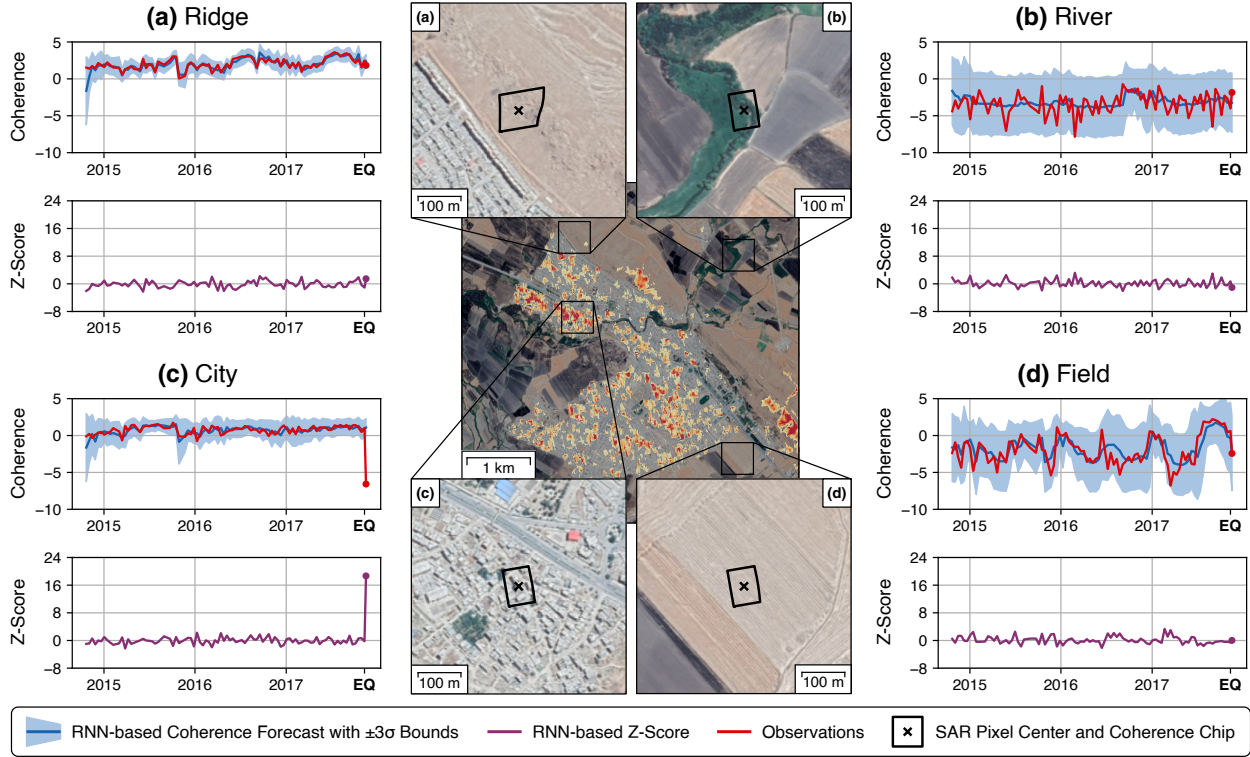


Figure 3.6: Sarpol-e-Zahab RNN damage map along with coherence time series and Gaussian forecasts for four representative locations ((a)-(d)) around the city. The z-score indicates the number of standard deviations between the forecast and the ground truth, and the coherence is plotted in logit space (i.e., it has been transformed into an unbounded space, see Eq. 3.9). Note that the shape of the coherence chip on the ground changes depending on the topography due to the way SAR data is acquired. The “Coherence” plotted on the y-axis is the logit transform of the squared coherence (Eq. 3.9). Optical imagery from Google, CNES/Airbus, taken July 27th 2020.

surface rupture, landslides and liquefaction (Brandenberg et al., 2019; Hough et al., 2020; Kendrick et al., 2019; Ponti et al., 2020; Zimmaro et al., 2020). In Fig. 3.7, we plot the RNN damage proxy map with two different z-score thresholds. The higher threshold allow us to focus on points that have had more anomalous coherence drops compared to their previous behavior through time. Using a threshold of $z = 4.93$ from the Amatrice data above (Fig. 3.7 (a)) we see that the largest anomalies lie on the M_w 7.1 and M_w 6.4 surface ruptures (running NW-SE and NE-SW, respectively) and liquefaction in the Searles Lakebed area, around (35.6°N , 117.3°W).

In Fig. 3.7 (b), we plot all points below $z = 0$ as black, more clearly showing smaller coherence anomalies surrounding the ruptures. Comparison with the mapped rup-

tures shows that some of these anomalies are due to smaller off fault ruptures, and we can also locate a small amount of damage in the town of Ridgecrest (Hough et al., 2020). The correlation with topographic slope of many of the smaller anomalies (for example around (35.80°N, 117.50°W)) suggest that these are due to small rockfalls or landslides induced by the earthquake. Damage maps such as these could be useful for directing mapping of ground failure in the aftermath of earthquakes.

3.6 Discussion

3.6.1 Importance of the training data

The goal of the RNN method is to produce the best possible forecast of the distribution of the co-event coherence value at each location, in the absence of any damage, given that location's pre-event coherence time series and the trained model. The forecast at every location depends on the trained model and thus contains information from every coherence time series used in the training. In this way, every forecast uses information learned from a wide spatial area.

Different parts of a given geographic region will be affected by processes that affect coherence (e.g., rain and snow) in a similar fashion, meaning that some amount of correlation between coherence time series in the region is likely. For example, a storm could cause consistent amounts of change in the surface properties across a wide area, leading to a sudden, correlated drop in coherence for many of the time series in the region.

When training the RNN, we split the overall training set $\mathcal{D}_t$ into training ($\mathcal{D}_{t,t}$) and validation ($\mathcal{D}_{t,v}$) components. $\mathcal{D}_{t,t}$ is used to optimize the RNN parameters, whereas $\mathcal{D}_{t,v}$ is used to evaluate the model performance according to Eq. 3.7 and choose the model with the best loss (note the ultimate task of damage classification is not part of model selection). When the training set $\mathcal{D}_t$ is split into its training and validation components, any correlation between time series from the same geographic area could lead to *data leakage* (Kaufman et al., 2011), whereby information from the validation set can also be found in the training set.

Data leakage can result in the RNN making artificially good forecasts as the RNN memorizes correlated patterns in the data, effectively over-fitting rather than learning generalizable rules. For example, we might get an unexpectedly good forecast of a sudden coherence drop in a time series which the RNN had not previously seen, due to the same pattern also being present in time series used in training. A possible case of this can be seen in time series (a) of Fig. 3.6; in late 2015 there is a sudden drop in

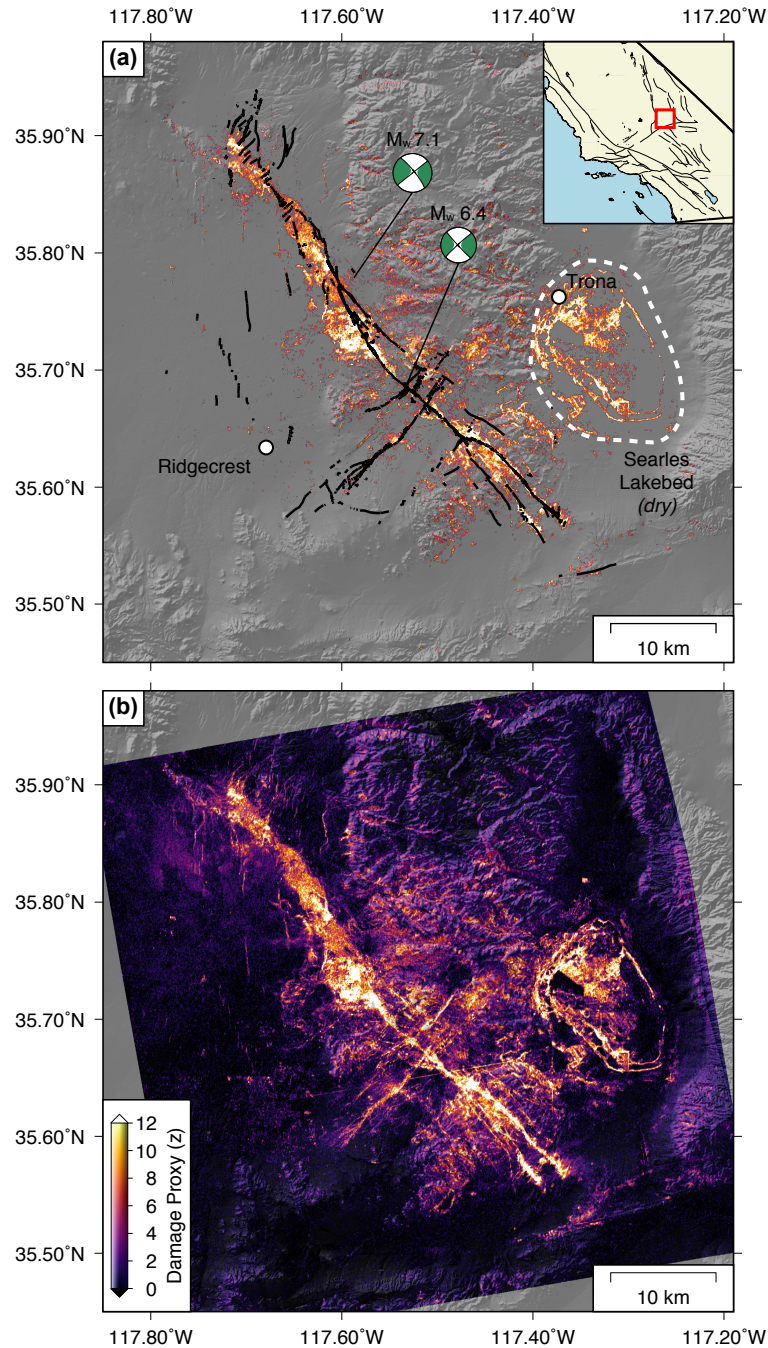


Figure 3.7: Ridgecrest RNN damage proxy map with different thresholds. **(a)** All points with $z < 4.93$ are masked. Black lines indicate mapped surface ruptures from Ponti *et al.* (Ponti *et al.*, 2020). The white dashed line indicates the approximate extent of the dry Searles Lakebed that saw substantial liquefaction (Zimmaro *et al.*, 2020). Global CMT focal mechanisms are plotted for the M_w 6.4 and M_w 7.1 earthquakes (Dziewonski *et al.*, 1981; Ekström *et al.*, 2012). The inset shows a regional map with simplified Quaternary faults. **(b)** Same damage proxy map as **(a)** except with no masking and points with $z < 0$ plotted in black. This threshold allows less intense off-fault anomalous change to be more clearly seen.

coherence (from around 3 to 0 in logit space) which is closely mirrored by the mean and variance of the forecast distribution. It seems unlikely that this coherence drop would be so accurately forecastable unless the network had seen many examples of similar patterns during training. Ideally, the network would instead learn that sudden drops in coherence can occur, and would broaden its uncertainty (i.e., the forecast standard deviation) accordingly.

Since data leakage can lead to artificially high performance on supervised machine learning tasks (Kaufman et al., 2011), we ask if data leakage during RNN training could lead to artificially improved damage classification. In our case, there can be no leakage of the actual ground truth damage data, which is the target for our RNN-based classifier but is not used in training. There can only be leakage in correlated patterns in the pre-event coherence time series. As detailed above, data leakage could cause over-fitting, making the forecast overly confident, i.e., with a standard deviation (σ') that is too small. An overly confident forecast will lead to a z-score that is larger for the same difference between observed coherence (x) and the mean of the forecast (μ'), making the z-score noisier and meaning the likely result of data leakage is a worse performing damage classifier, not one with artificially improved performance.

The quality of the RNN forecast will depend on the training data which we use. In general, training data acquired over a shorter time span and over a smaller spatial area will not sample the full scope of representative coherence behavior. We expect that more limited training data will cause the model to give a less representative forecast coherence distribution, as the network will see fewer examples of how coherence can vary in time and space. In Section B.4, we explore how decreasing the time span of the training data affects the quality of the damage map for the Amatrice example, finding that the results are highly variable when less than a year of data is used. Based on these results, we emphasize the importance of testing the RNN method on a wider variety of disasters, in different geographic regions, to ensure robust performance.

A smaller geographic region is more likely to have strongly correlated coherence time series, leading to more significant data leakage problems and over-fitting during training. Drawing pixels from a very wide geographic area, or potentially many different geographic areas all over the Earth, could ameliorate this problem. A systematic exploration of the relationship between the input SAR data and the quality of the final damage map is beyond the scope of this paper.

Coherence images with longer temporal baselines will generally have lower coherence than images from a similar time period and region with a shorter temporal baseline (Zebker & Villasenor, 1992). Our current RNN training approach ignores the temporal baseline of the coherence images, effectively assuming that all coherence images have the same temporal baseline. However, our coherence time series have some variation in temporal baselines due to the varying acquisition frequency of Sentinel-1 SAR data, with the repeat frequency tending to become more stable and more frequent with time (see Section B.3 for details of the data). A variable temporal baseline adds an extra noise term due to changing amounts of temporal decorrelation. Therefore, the likely effect of a variable temporal baseline is to make the coherence time series noisier, thus decreasing the confidence of the forecast and reducing the sensitivity to damage.

For the Amatrice case, the timing of the event meant that we had less than two years of Sentinel-1 data preceding the earthquake, and the data had a higher variance in the temporal baseline. The lack of data and variable temporal baseline degrades the performance of the RNN, so the precision-recall results presented are likely a lower bound on the possible performance for damage mapping in this area.

3.6.2 Sensitivity to damage in different geographic areas

Random motion of scatterers within a resolution element on the scale of the wavelength of the satellite radar signal (≈ 5.6 cm for this work) will cause decorrelation between two radar echoes (Zebker & Villasenor, 1992). Different regions of any study area will have different background rates of change in their surface properties, and these rates may vary through time. High rates of surface change will lead to low coherence, and variability through time in the rate of change will create a larger standard deviation in the coherence time series.

Robust damage detection relies on separating normal changes from damage induced changes, and the ability to do this separation depends on the rate of surface change and how much this rate varies in time. At a given time step, for a given pixel, the RNN forecasts the average rate of change with the mean of the forecast (μ'), and the variability with its standard deviation (σ'), both of which are used to calculate the z-score (Eq. 3.10). A high background rate of surface change will lead to a low μ' , thus obscuring coherence drops due to damage as $\mu' - x$ will be small. Similarly, large coherence variability will cause larger σ' values leading to smaller

z-scores, making it hard to separate coherence drops that are natural from those that are damage induced.

Differing behavior of pixels means that, for two different pixels, the same z-score does not necessarily imply the same level of damage, but instead the same ratio of coherence change to background coherence variability. Thus, when interpreting the z-score map, it may also be useful to consider the forecast mean and standard deviation to understand the noise level for each pixel, with low mean and high standard deviation indicating noisy pixels and thus lower sensitivity to damage. We note that stable, human-made structures typically have higher and less variable coherence, whereas areas with vegetation, water and snow typically have lower and more variable coherence.

For the Ridgecrest damage proxy map (Fig. 3.7), it is noticeable that the largest z-scores appear to correspond to the most significant ground disruption, specifically the surface ruptures from the M_w 6.4 and M_w 7.1 earthquakes and liquefaction near the town of Trona (Brandenberg et al., 2019; Hough et al., 2020; Zimmaro et al., 2020). The apparent link between z-score magnitude and intensity of ground surface change may be due to the desertic conditions in the area. The dry, stable conditions mean that most pixels have similar behavior through time, i.e., they have a similar noise level, meaning z-scores are more directly comparable between pixels. More generally, we should be able to use z-scores as proxies for levels of damage within groups of pixels that have similar forecast standard deviations. However, the z-score is less comparable between groups of pixels with very different forecast means and standard deviations.

The desertic conditions in the Ridgecrest area mean that coherence is comparatively high and stable through time. Because of this stability, the single pre-event image used in the CCD method is a better proxy for the pre-event coherence than the single images used in the other regions, causing the RNN and CCD methods to be more similar than for the other two case studies (we do not present the CCD results here).

Damaged buildings that are smaller or on poorly orientated slopes with respect to the satellite line of sight will occupy a smaller fraction of the coherence chip. These buildings will therefore have a smaller effect on the coherence, making them harder to identify using coherence based methods. The choice of the fraction of the chip area which has to be occupied by damaged buildings in order for that chip to be in the “truly damaged” class can therefore have a substantial effect on the final precision-recall area under the curve (PR AUC) values for the Amatrice data set.

We choose one third of the chip area (roughly 1200 m^2) as it gives approximately the largest PR AUC values, although the change in PR AUC values between fractions of 20% and 50% are small (RNN PR AUC in the range 0.68-0.70). The decrease of the RNN PR AUC below 20% building fraction suggests that our method has difficulty identifying damage where the damaged area occupies less than around 700 m^2 of the coherence chip, meaning, for example, our method could have difficulties correctly classifying an isolated damaged building that is smaller than $25 \text{ m} \times 25 \text{ m}$. With the present resolution, this method is likely most useful for detecting large damaged buildings and damaged blocks, rather than damage to individual smaller buildings.

3.6.3 The choice of change metric and forecast distribution

While coherence has proven useful for surface change detection, metrics such as the SAR amplitude correlation can also be used (Jung & Yun, 2020). Using the amplitude can be particularly useful when the InSAR coherence is low. We suspect that a similar RNN-based anomaly detection approach would also work on time series of other metrics used for SAR change detection, and could also be combined with the coherence metric; however, we do not pursue this here.

When calculating coherence, we approximate an ensemble average for a given SLC pixel with a spatial average around that SLC pixel when evaluating Eq. 3.1 (e.g., see Section 3.12.2.5 of Simons and Rosen (2015)). In our case, we use a local chip (15 by 5 SLC pixels in size) and assume that these SLC pixels are drawn from the same statistical distribution of amplitude and phase. There is a trade off in the choice of chip size—smaller chips are more likely to combine pixels that demonstrate the same statistics through time (as a result of being over the same type of land surface). However, they also contain fewer samples with which to estimate the coherence, leading to a larger variance in the estimate and a noisier coherence time series. Larger coherence chips contain more samples, but are more likely to include SLC pixels with greater differences in statistical behavior through time, less consistent with the assumption that we used to approximate the ensemble average. Larger chips also provide lower spatial resolution. Limited testing with the data presented in this study indicates that the results are made worse by reducing the chip size to 9 by 3 SLC pixels; however, we have not systematically explored how the results vary with chip size.

The choice of forecasting a Gaussian distribution on the logit transform of coherence squared is motivated by the desire to use an unbounded distribution on an unbounded

space, and the mathematical relationship to the logarithm of the phase variance (Section B.1). However, the specific transform and distribution are ad-hoc and in general they may not produce the optimal forecast. We leave the exploration of the best transform and distribution to future work.

Rather than using a local chip to estimate coherence, a stack of SAR images can be used to identify a non-local neighborhood of SLC pixels, within some distance of a central SLC pixel, that behave in a statistically similar pattern through time. The coherence can then be evaluated over those SLC pixels, which has been shown to give a better coherence estimate (Ferretti et al., 2011). However, using a non-local coherence evaluation introduces problems when doing damage mapping. Damage will change the statistical characteristics of an SLC pixel, meaning that co-event SLC pixels may no longer be in their pre-event statistical groupings. For example, some of the SLC pixels may be over collapsed buildings, with other SLC pixels over buildings that remained undamaged. On the other hand, SLC pixels in a local chip are more likely to have been affected by the same process (e.g., building collapse). Thus, while a non-local coherence calculation may give improved results for pre-event coherence calculations, our method requires us to use a local chip for the coherence calculation.

3.6.4 Near real-time deployment

When deploying damage mapping for rapid response, delivering timely products, ideally within hours, is exceptionally important. While the wait time for a Sentinel-1 post-event acquisition could be up to six days, in many cases we will have an image before this, allowing the information to feed in to rapid disaster response. Other SAR satellites also acquire data; however, they do not have the same long time series of open access acquisitions that is available from Sentinel-1. Planned SAR missions, such as NISAR (Sharma, 2019) and ALOS-4 (Motohka et al., 2019) should improve the availability of frequently acquired SAR data.

A near real-time deployment could follow this workflow, which would ideally be almost completely automated:

1. Identify the natural disaster
2. Find the SAR satellite which has a sufficient time series of existing observations, and will have an overflight of the affected area in the immediate future

3. Download the existing SAR data, coregister, and calculate a coherence time series
4. Train the RNN and forecast the co-event coherence
5. Obtain the first post-event acquisition, coregister and calculate co-event coherence
6. Compare with the forecast to calculate the damage proxy map
7. Inspect DPM and distribute to first responders via previously established channels of communication, ensuring that they have a clear understanding of the information the damage proxy map is providing.

The most computationally demanding step of this process is the coregistration of the large quantities of pre-event SAR data, with the coherence time series calculation and RNN training also requiring substantial computational resources. Combined, these steps can take several days of computing using our current codes and resources, potentially impacting response times. However, steps 1–4 do not necessarily affect the post-disaster response time, as they can either be pre-computed and regularly updated, or in some cases, be completed before the essential post-event acquisition becomes available. Additionally, the use of cloud computing resources and improved algorithms can greatly decrease processing time. Free and open data, accessible with minimal latency, are vital for the effective deployment of such a disaster monitoring system.

The quality of the damage map could be further enhanced by combination with other damage assessments, e.g., maps of shaking intensity and on-the-ground reports, as well as previously identified zones of higher risk for building collapse, fault surface rupture, landslides and liquefaction (Loos et al., [2020](#)).

3.7 Conclusions

In this work, we present a deep learning-based damage mapping algorithm for synthetic aperture radar (SAR) sequential coherence time series. Coherence represents a proxy for ground surface change, and thus by separating out anomalous from expected coherence values after a natural disaster we can find regions that have had anomalous surface change, potentially due to building collapse, surface rupture, landslides or other hazards. We use a recurrent neural network to learn the normal behavior of coherence through time by training on SAR coherence time series

spanning a large area, then forecast the probability distribution of the coherence we expect without any disaster. We then use the deviation between the observed and forecast coherence to locate anomalous coherence changes, which we assume to be due to collapsed buildings. A comparison with on-the-ground building damage assessments shows that this method is quantitatively better than an alternative method of damage mapping based on coherence loss. We discuss the advantages and limitations of our proposed SAR damage mapping method and outline how it could be deployed in disaster response scenarios.

The problem of RNN over-fitting due to the correlated nature of coherence time series in a single region could potentially be ameliorated by simultaneously training on a large number of coherence time series drawn from areas all over the planet displaying very different temporal behaviors. Furthermore, including additional training features such as the spatial and temporal baseline and the amount of precipitation between sequential SAR acquisitions might allow the network to learn the dependence on additional physical parameters relevant to coherence, thus improving its forecast. The ability to learn from many different input features without a physical model is a key advantage of the deep learning approach.

The work presented here has been using C-band (5.6 cm wavelength) SAR data. As coherence is sensitive to surface disruptions on the scale of the radar wavelength, it could be that we are picking up many false positives caused by superficial damage. Investigating the same disasters with 24 cm wavelength L-band data may provide damage maps that are less prone to pick up small surface disturbances. Unfortunately, dense time series of L-band SAR data are not easily available, although the future launches of L-band SAR satellites (e.g., NISAR Sharma (2019) and ALOS-4 Motohka et al. (2019)) will allow for further exploitation of L-band SAR data for damage mapping. NISAR will also image selected regions using S-band radar, allowing the use of SAR data at multiple wavelengths to further increase the ability of the damage map to distinguish different types of surface change.

Chapter 4

EXPLORING THE EFFECT OF THERMAL PRESSURIZATION ON EARTHQUAKE PROPAGATION THROUGH A CREEPING FAULT USING SIMPLIFIED 2-D MODELS

4.1 Introduction

Faults accommodate strain via rapid slip during earthquakes, or slow slip, known as creep (Avouac, 2015; Burgmann, 2000; Simons et al., 2011). Based on long-term geodetic observations, faults are often divided into locked and creeping sections, with their behavior explained by the frictional properties of each section (Harris, 2017). Locked sections are thought to have velocity-weakening friction, which promotes a stick-slip response to loading. Creeping sections, on the other hand, can be modeled as having velocity-strengthening friction, allowing stress to be accommodated via gradual creep, rather than dynamic rupture. Within this framework, creeping sections of faults represent barriers to earthquake rupture. Properly classifying faults as locked or creeping can therefore allow the possible locations and magnitudes of future earthquakes to be estimated (Kaneko et al., 2010).

The idea that creeping faults represent barriers to earthquake rupture has been challenged by evidence that faults can dramatically weaken at high slip rates (Di Toro et al., 2011; Tsutsumi & Shimamoto, 1997; Tullis, 2007), potentially allowing a fault to exhibit both creep due to velocity-strengthening behavior at slow slip rates, and rapid dynamic failure. One proposed dynamic weakening mechanism is thermal pressurization of pore fluids (Lachenbruch & Sass, 1980; Mase & Smith, 1987; Rice, 2006; Sibson, 1973), whereby rapid fault slip at high stresses generates heat, which in turn raises the pressure of fluids infiltrating the fault zone. This increase in pressure reduces the effective normal stress on the fault, resulting in a decrease in fault strength. Evidence from lab experiments (Badt et al., 2020; Faulkner et al., 2011; French et al., 2014), direct observations of fault zones (Kuo et al., 2022; Ujiie et al., 2010), seismological observations (Viesca & Garagash, 2015), computational models (Cubas et al., 2015; Noda & Lapusta, 2013), and theory (Lachenbruch, 1980; Mase & Smith, 1987; Rice, 2006; Sibson, 1973) suggests that thermal pressurization (TP) can play a significant role in weakening natural faults at high slip rates, and can allow dynamic ruptures to propagate into creeping sections. A variety of processes

other than TP have been proposed that would lead to substantial decreases in fault strength during sliding. These include flash heating, gel lubrication, nanoparticle lubrication, decarbonation, dehydration, and fault melting (Di Toro et al., 2011); however, we focus on TP in this work.

The efficiency of TP depends on numerous factors, including the rate of heat generation in the fault zone, the coupling between pore fluid temperature and pressure changes, and the time scales of diffusion of heat and fluid out of the fault zone. A slow rate of heat generation, or rapid diffusion of pore fluid pressure, can result in TP playing a limited role, while the generation and maintenance of high temperatures and pressures can rapidly reduce fault strength (e.g., Lachenbruch (1980) and Rice (2006)). Many of the parameters governing TP are highly uncertain, and can vary in space and time (Aben et al., 2020; Rice, 2006). This uncertainty makes predicting when TP would occur a challenging prospect, leading to significant uncertainties in understanding the behavior of creeping faults when interacting with dynamic ruptures.

The San Andreas Fault (SAF) in California (USA) exhibits both stick-slip and creeping behavior. A long creeping section stretches for around 140 km, from the town of Parkfield in the south to San Juan Baptista in the north (Scott et al., 2020), and is bounded by locked sections to the north and south which have hosted large earthquakes. At the southern end, the transition zone between locked and creeping segments has hosted repeated M_w 6 earthquakes (Bakun et al., 2005; Bakun & Lindh, 1985) and has been the site of intensive study, including direct drilling in to the creeping fault as part of the San Andreas Fault Observatory at Depth (SAFOD) project (Lockner et al., 2011).

The creeping section has not hosted a major earthquake historically, and paleoseismic studies do not show evidence of a major rupture within the past 5000 years (Toké et al., 2011; Toké et al., 2006; Toké & Arrowsmith, 2015). However, geodetic measurements indicate a slip deficit within the creeping section (Jolivet et al., 2015; Maurer & Johnson, 2014) that could potentially be released by seismic slip. French et al. (2014) found that the material from the creeping fault zone at SAFOD could dramatically weaken at seismic slip rates, providing a possible mechanism for ruptures to propagate into the creeping segment. Their results were consistent with TP playing a significant role in the weakening. Recent results by Coffey et al. (2022) present evidence of repeated dynamic ruptures within the creeping section, based on the use of biomarker thermal maturity to calculate previous temperature rises

from SAFOD samples. They calculated temperature rises of between 570-1100 °C within a zone that they estimated to have hosted more than 100 earthquakes.

Motivated by these observations, we seek to understand the role that TP can play in allowing earthquakes to rupture the creeping section of the SAF. We construct a highly simplified 2-D model of the fault, centered around Parkfield, with a creeping section that exhibits stable, velocity-strengthening behavior at slow slip rates but can undergo dramatic co-seismic weakening due to TP. We nucleate events on a nearby velocity-weakening patch, then explore how these dynamic ruptures interact with the creeping section as we vary the efficiency of TP. We use a computational methodology that allows us to efficiently model decades-long periods of creep, followed by rapid dynamic ruptures in the presence of TP (Lapusta et al., 2000; Noda & Lapusta, 2010).

Parameters controlling TP are often determined by laboratory measurements (e.g., Tanikawa and Shimamoto (2009) and Wibberley and Shimamoto (2003)), and are then used in simulations (e.g., Cubas et al. (2015) and Noda and Lapusta (2013)). However, several of the parameters controlling TP are uncertain, and can potentially vary substantially during the rupture process (e.g., Aben et al. (2020), Badt et al. (2020), and Brantut and Mitchell (2018)). In this study, we focus on three of the parameters that are least well constrained: the width of the shearing zone, the hydraulic diffusivity, and the coupling coefficient between temperature rise and pressure rise. Proposed values span several orders of magnitude, and we draw on published results to determine the ranges for each parameter. We use these parameters to vary the strength of TP within the creeping zone.

We begin by modeling events in the style of the 2004 M_w 6 Parkfield earthquake. This rupture rapidly arrested in the creeping section of the SAF (Barbot et al., 2012), and we use this observation to rule out the parameter choices for which TP allows a M_w 6-style rupture to propagate substantially into the creeping section. We then simulate larger events, in the style of a hypothetical future M_w 7 rupture on the Cholame segment of the SAF, and observe how they interact with the creeping section. We find a range of physically reasonable TP parameters for which M_w 6-style events rapidly arrest, but M_w 7-style events propagate partially or totally through our simulated creeping section, allowing the event to grow much larger. We also present some simple analytical models that give further insight into the conditions which allow for the creeping section to be ruptured, and discuss the limitations of our simulations.

4.2 Methods

We apply the computational methods developed by Lapusta et al. (2000) and Noda and Lapusta (2010) to simulate long sequences of inter-seismic slip, followed by earthquake nucleation, propagation, and arrest. In order to efficiently explore a wide parameter space, we limit our study to ruptures propagating along a one-dimensional (1-D) fault embedded in a 2-D linear elastic isotropic medium. To approximate the effects of a finite width of the seismogenic zone on a strike-slip fault, we adopt a crustal plane model (Kaneko & Lapusta, 2008; Lapusta, 2001; Lehner et al., 1981). The only non-zero component of slip is along-strike, averaged over the assumed depth of the fault, H_{seis} . Below H_{seis} , the medium is coupled to a substrate moving at the loading rate, v_{pl} . The parameters we adopt in this study are given in Table 4.1, and further description is given in Section 4.2.4.

4.2.1 Rate-and-State Friction

We model the shear resistance of the fault as the product of the effective normal stress and the friction coefficient given by the standard laboratory-derived rate-and-state law (Dieterich, 1978; Ruina, 1983):

$$\tau = \sigma_{eff} f(v, \theta) \quad (4.1)$$

$$= (\sigma - p) \left[f^* + a \log \left(\frac{v}{V^*} \right) + b \log \left(\frac{\theta v^*}{D_{RS}} \right) \right], \quad (4.2)$$

where τ is the shear strength which is equal to the shear stress, $\sigma_{eff} (= \sigma - p)$ is the Terzaghi effective stress, σ is the normal stress, p is the pore fluid pressure, f is the rate-and-state friction coefficient, v is the sliding velocity, f^* is the reference friction coefficient, defined as the steady state value of f when sliding at the reference velocity V^* , D_{RS} is the characteristic slip for the evolution of the state variable, and a and b are the rate-and-state constitutive parameters.

The evolution of the state variable, θ , is given by the aging law:

$$\dot{\theta} = 1 - \frac{v\theta}{D_{RS}}. \quad (4.3)$$

At steady state, $\dot{\theta} = 0$ and the friction coefficient reduces to:

$$f_{ss}(v) = f^* + (a - b) \log \left(\frac{v}{V^*} \right). \quad (4.4)$$

Equation 4.4 shows that the steady-state friction is governed by the sliding velocity v and the parameter $(a - b)$. When $(a - b) > 0$, friction increases with sliding velocity, a situation known as velocity strengthening. Velocity strengthening results in stable sliding when a steady load is applied. Patches with $(a - b) < 0$ are referred to as velocity weakening, and can potentially nucleate seismic ruptures (e.g., Rice and Ruina (1983) and Rubin and Ampuero (2005)).

4.2.2 Enhanced Dynamic Weakening via TP

To simulate the evolution of temperature due to shear heating and the resulting changes in pore fluid pressure, we use the following two coupled partial differential equations (Noda & Lapusta, 2010):

$$\frac{\partial T}{\partial t} = \alpha_{th} \frac{\partial^2 T}{\partial y^2} + \frac{\tau v \exp(-y^2/2w^2)}{\rho c \sqrt{2\pi} w}, \quad (4.5)$$

$$\frac{\partial p}{\partial t} = \alpha_{hy} \frac{\partial^2 p}{\partial y^2} + \Lambda \frac{\partial T}{\partial t}, \quad (4.6)$$

where T is the temperature, α_{th} is the thermal diffusivity, τ is the shear stress on the fault, v is the sliding velocity, ρc is the specific heat, p is the pore fluid pressure, α_{hy} is the hydraulic diffusivity, and Λ is the coupling coefficient that gives the change in pore pressure for a given change in temperature under undrained conditions. We assume a Gaussian distribution of the shearing rate, with w being the root-mean-square half-width of the shear-rate distribution and y being the perpendicular distance from the fault.

4.2.3 Model Setup

Our model is designed as a highly simplified representation of the southern portion of the creeping section of the SAF, and the locked Cholame segment, centered around the town of Parkfield, California. At Parkfield, the fault transitions from locked in the south, to the creeping segment in the north. This transition zone has hosted repeated M_w 6 earthquakes over the last century (Bakun et al., 2005; Bakun & Lindh, 1985).

We draw on the work of Barbot et al. (2012) in constructing a 2-D model consisting of a velocity-weakening patch with velocity-strengthening segments on either side. When loaded, these velocity-strengthening sections creep at the plate rate, set at 23 mm/yr based on observed creep rates (Titus, 2006). This creep loads the velocity-weakening patch, causing dynamic ruptures to nucleate. We put a small initial

stress perturbation near the right-hand edge of the velocity-weakening patch, which results in the events nucleating from the right and propagating to the left, where they hit the creeping section (Figure 4.1). Our model setup gives a total length of the creeping section of around 80 km. While this is shorter than the creeping section of the SAF, in practice this makes little difference to our results and reduces the computational burden. The velocity-strengthening patch on the right-hand side of the model is similar to the modeling approach used by Barbot et al. (2012). It allows the velocity-weakening patch to be loaded from the right-hand side in order to create ruptures that propagate to the left. However, the patch does not directly relate to observed creep rates on the SAF, which decrease rapidly to the south of Parkfield. We note that our goal is to focus on the impact of variable TP efficiency in the creeping section, and not to realistically model the potentially complex transition zone around Parkfield (e.g., Simpson et al. (2006)).

We seek to understand the response of the creeping section to the incoming ruptures of different sizes as the efficiency of TP varies. TP parameters are constant throughout the left-hand side creeping section, with no TP in other regions ($\Lambda = 0$ MPa/K). In our simulations, we create incoming ruptures of two different styles. The first type of rupture is designed to be similar to the 2004 M_w 6 Parkfield event. We use the observation that this event rapidly arrested in the creeping section to place constraints on the efficiency of TP. We also simulate a larger event, with a magnitude of around M_w 7 in the absence of TP. To control the event size, we vary two parameters: the length of the velocity-weakening patch, λ_{vw} , and the crustal plane width, H_{seis} . We determine the approximate event sizes by examining the rupture behaviour in the absence of TP, where the rupture is constrained to the velocity-weakening patch. In Figure 4.1, we illustrate our model setup and slip profiles for the two styles of events, and describe them in more detail below.

The crustal plane model is used to approximate the depth-averaged slip over the seismogenic fault depth H_{seis} , with the fault loaded at v_{pl} at the bottom, and a free surface at the top (Lapusta, 2001; Lehner et al., 1981). H_{seis} can also be regarded at the half-width of a fault that is constrained to move at v_{pl} from the bottom and the top, with no free surface. This situation is more akin to the 2004 Parkfield earthquake, which was predominantly confined to below 5-km depth and did not rupture to the surface (e.g., Barbot et al. (2012)).

For the M_w 6-style events, we base our geometry on the inferred rupture parameters for the M_w 6 2004 Parkfield earthquake. This event propagated to the north,

rupturing a patch roughly 25 km long by 4 km wide, and rapidly arrested in the creeping section (e.g., Barbot et al. (2012)). We create a velocity-weakening patch 25 km long, and set the crustal plane thickness, H_{seis} , to 2 km, or half of the inferred rupture patch width. This model setup effectively constrains ruptures to a 4-km wide patch along the entire length of the fault. In reality, in situations where TP allowed ruptures to propagate substantially into the creeping section, the rupture would likely expand over more of the fault depth. This artificial constraint on the depth extent of the ruptures acts to inhibit propagation into the creeping section for the M_w 6-style events.

For the M_w 6-style events, our choice of frictional parameters results in slip of around 66 cm along most of the velocity-weakening patch (Figure 4.1(a)). This number is consistent with the maximum slip, found at around 7 km depth, in the kinematic inversions and simulations of Barbot et al. (2012), although is larger than that found by some other authors (e.g., Johanson et al. (2006) and Langbein et al. (2006)). The use of the crustal plane model means that the slip in our model is the average over the depth of the patch, so the magnitude of our events is somewhat larger than the observed M_w 6; however, this difference should not significantly affect our conclusions.

For the M_w 7-style events, we extend the velocity-weakening patch to be 80 km long, and set the seismogenic depth to be larger, $H_{seis} = 8$ km, resulting in ruptures with slip of around 2.2 m along most of the velocity-weakening patch. H_{seis} is here used to represent the full depth of the rupture, which is assumed to span from the surface to a depth of 8 km. These parameters are not directly based on observational constraints, but are designed to produce an intermediate size of event, larger than the M_w 6-style, but still much smaller than the M 7.7 1857 Fort Tejon earthquake (e.g., Zielke et al. (2012)), the most recent major earthquake to rupture the SAF near Parkfield. Note that the Fort Tejon event likely nucleated around the Parkfield transition zone, then propagated from the north to the south (Sieh, 1978a), unlike the ruptures in our simulations.

4.2.4 Parameters Controlling TP

Equations 4.2, 4.5, and 4.6 describe a coupled system whereby the shear stress controls the rate of heat production, which in turn governs the evolution of pore pressure, which feeds back into the shear stress via the effective normal stress. In order for TP of pore fluids to occur, there must be sufficient heat generated to increase

the pore fluid pressure, and this heat and pressure must remain sufficiently confined to the shearing layer over the time scales of seismic slip. The relative weight of these processes is controlled by the thermal and hydraulic diffusivities α_{th} and α_{hy} , specific heat ρc , fault zone half-width w , and coupling coefficient Λ . Of these parameters, the values of $\alpha_{th} = 10^{-6}$ m²/s and $\rho c = 2.7$ MPa/K are relatively well constrained from laboratory results and have been used in previous computational work (e.g., Noda and Lapusta (2010, 2013), Rice (2006), and Wibberley and Shimamoto (2005)).

We therefore focus our attention on Λ , α_{hy} , and w . Given these parameters are uncertain, we search the literature for the range of values proposed for each of these quantities, and explore that range in our simulations. We assume that each parameter is fixed during each simulation, which may not be the case in reality (see Section 4.5.2). Exploring the effects of potential evolution of these quantities during rupture should be the subject of future work.

The coupling coefficient, Λ , depends on the degree of damage to the fault gouge material, and also evolves with the pressure-temperature conditions. Rice (2006) draws on experimental results (including from gouge drawn from the Median Tectonic Line, Japan, presented by Wibberley and Shimamoto (2003)) to suggest values between 0.31 and 0.98 MPa/K, with higher damage resulting in lower values and thus less efficient TP. Noda and Lapusta (2013) calculate values of 0.036 and 0.069 MPa/K based on laboratory measurements from Tanikawa and Shimamoto (2009) of material taken from different points in the Chelungpu fault zone, Taiwan. Aben et al. (2020) give values in the range 0-2 MPa/K depending on the depth and degree of damage. In our work, we vary Λ over the range from 0.01 MPa/K to 1 MPa/K. Most of these values would be appropriate for the fault with off-fault damage due to the rupture, with the simulations essentially ignoring the transition from undamaged to damaged conditions and considering the values for the damaged fault.

The hydraulic diffusivity, α_{hy} , also depends on the degree of damage and the pressure-temperature conditions. Rice (2006) gives values between $0.86 - 6.04 \times 10^{-6}$ m²/s, with higher damage resulting in larger hydraulic diffusivities. Noda and Lapusta (2013) calculate values of 3.5×10^{-2} and 7.0×10^{-5} m²/s, again based on the results of two different samples presented by Tanikawa and Shimamoto (2009). Aben et al. (2020) gives a wide range of parameters, from $\sim 10^{-8}$ to 10^{-2} m²/s, depending on the depth and degree of damage. For our work, we vary the hydraulic diffusivity over the range 10^{-5} to 10^{-2} m²/s.

Theoretical, observational, and experimental results can be used to place bounds on the half-width w of the shearing zone during seismic slip. Field observations of faults often find broad zones of damage, on the scale of meters, surrounding narrow zones of ultracataclasite and fault gouge, centimeters to millimeters wide, which host earthquake slip (e.g., Mitchell and Faulkner (2009) and Sibson (2003)). Microstructural analysis indicates that, within these narrow zones, slip can be accommodated over narrower “principle slip zones,” on the order of 0.1 to 1 mm thick (e.g., Chester and Chester (1998), De Paola et al. (2008), and Heermance et al. (2003)). Rice et al. (2014) and Platt et al. (2014) argued that fault gouge at seismogenic depths would tend to undergo extreme shear localization, resulting in shear being accommodated over zones 5-40 μm wide. Coffey et al. (2022) examined material drawn directly from the SAFOD bore hole in the creeping section of the SAF, finding multiple slip layers with thicknesses between 100 μm and 1.8 cm, with an average of 2 mm. To encapsulate the broad range of the proposed shear-zone widths, we vary the half-width, w , from 0.01 mm (10 μm) to 100 mm. These values remain constant during slip; however, we acknowledge that the wider shearing layers may be unstable, and localize, over extended slip (Platt et al., 2014; Rice et al., 2014).

We set the initial effective normal stress, σ_{eff} , at 120 MPa, the approximate value at 7 km depth with hydrostatic pore pressure. Based on experimental results from material obtained from SAFOD we choose $f^* = 0.15$ within the creeping section (Lockner et al., 2011). The steady-state velocity dependence of the rate-and-state friction is governed by the quantity $\sigma_{eff}(a - b)$ (Equation 4.4), which we set to 0.48 MPa within the creeping section, consistent with values determined from inversions of post-seismic slip (Barbot et al., 2009; Barbot et al., 2012; Chang et al., 2013) and experimental measurements of fault properties (Lockner et al., 2011). Properties within the velocity-weakening section are chosen to give events of the desired size, as discussed by Barbot et al. (2012). All parameters are given in Table 4.1.

The dramatic weakening caused by TP can lead to very rapid increases in slip rates, resulting in high strain rates. In reality, such high strain rates would lead to inelastic yielding, i.e. damage, off the fault around the rupture front. Such behavior can be approximated by imposing a slip velocity limit that varies with the normal stress (Andrews, 2005; Lambert et al., 2021). We set $v_{lim} = 15$ m/s, to approximate the behavior at roughly 7 km depth.

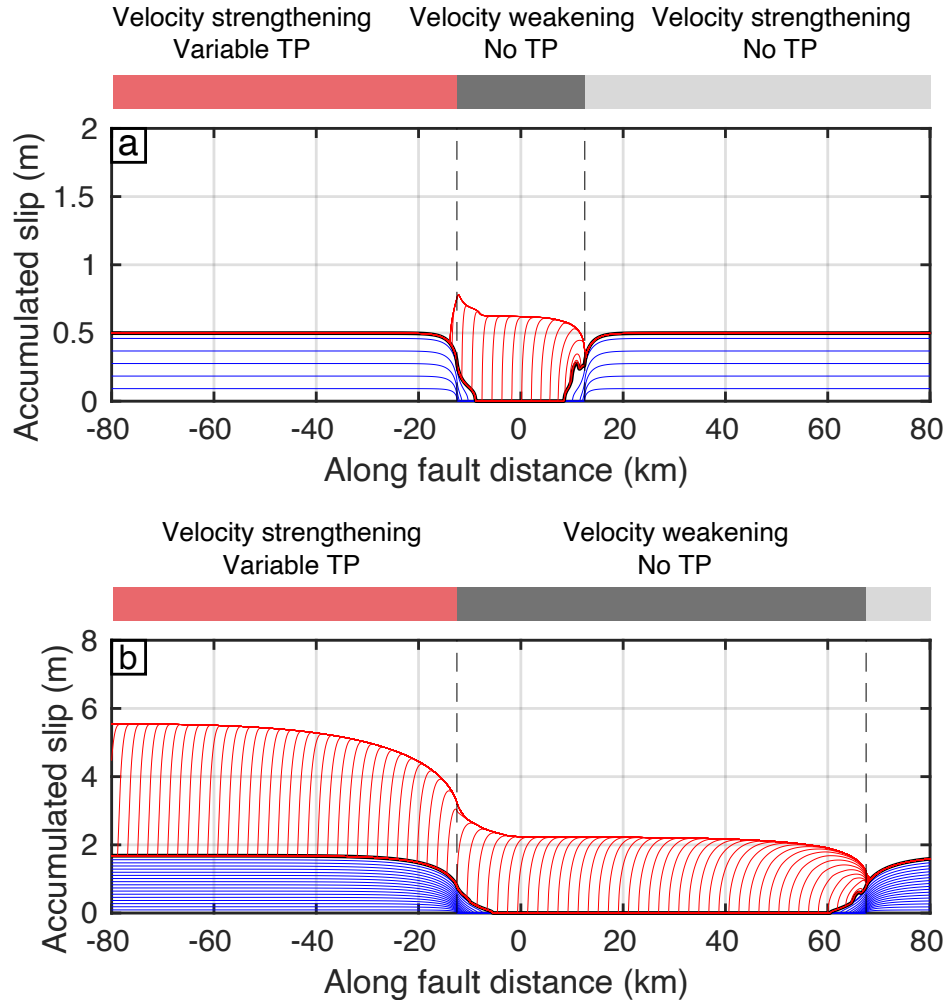


Figure 4.1: Examples of M_w 6 and M_w 7-style events. Blue contours show accumulated inter-seismic slip every four years. Red lines show accumulated co-seismic slip every 0.5 seconds. **(a)** M_w 6-style event. A 25-km-long velocity-weakening patch is placed between two velocity-strengthening patches, the crustal plane thickness is $H_{seis} = 2$ km. The velocity-strengthening patch on the left has TP with $\Lambda = 0.34$ MPa/K, $w = 10$ mm and $\alpha_{hy} = 10^{-4}$ m²/s for this case; the parameters are varied in other simulations (Table 4.1). The TP with these parameters is not efficient enough for the M_w 6-style event to propagate into the creeping section. **(b)** M_w 7-style event. The velocity-weakening patch is expanded to 80 km long, and crustal plane thickness increased to $H_{seis} = 8$ km. TP parameters have the same values as in (a). The larger event now activates sufficient TP to propagate through the creeping section and hit the boundary of our model (not shown here), allowing the event to grow well beyond M_w 7.

4.2.5 Example of Simulated Rupture in the Creeping Section

In Figure 4.1(b), we illustrated a M_w 7-style event propagating through the creeping section due to TP. In Figure 4.2, we also show the evolution of the physical parameters at a point on the edge of the creeping section during a period of creep, followed by a M_w 7-style event that propagates into the creeping section. Figure 4.2(a) shows the evolution of shear stress: steady creep occurs at a background stress set by the rate-and-state parameters, before seismic slip creates a spike to a peak stress and rapid evolution to a higher stress level, again controlled by the rate-and-state parameters (the evolution of the state parameter is shown in Figure 4.2(e)). Without TP, this increase in shear stress would rapidly arrest the ruptures. However, TP causes shear stress to decay exponentially with slip, due to the reduction in effective normal stress (Figure 4.2(b)). The dramatic reduction in stress allows several meters of slip to accumulate with temperature change limited to 370 °C (Figure 4.2(d)).

4.3 Analytical Models of Thermal Pressurization

4.3.1 TP Weakening Length Scales

Previous authors have derived analytical solutions for the evolution of stress due to TP (Lachenbruch, 1980; Mase & Smith, 1987; Rice, 2006). These models allow us to gain some insight into how TP parameters may control the rupture of the creeping section. Two end-member solutions have been derived, one for constant velocity slip on a plane, in which the very existence of the solution is crucially dependent on the off-fault diffusion of heat and fluids, and the other for adiabatic, undrained shear over a zone of finite width, which is only valid if the layer is wide enough for the heat and fluid diffusion to be negligible for the duration of the event.

For slip on a plane (i.e. the case with an infinitely thin shear zone), one has (Mase & Smith, 1987; Rice, 2006):

$$\tau = f(\sigma_n - p_o) \exp\left(\frac{\delta}{L^*}\right) \operatorname{erfc}\left(\sqrt{\frac{\delta}{L^*}}\right), \quad (4.7)$$

where:

$$L^* = 4 \left(\frac{\rho c}{f \Lambda} \right)^2 \frac{(\sqrt{\alpha_{hy}} + \sqrt{\alpha_{th}})^2}{v_{seis}}, \quad (4.8)$$

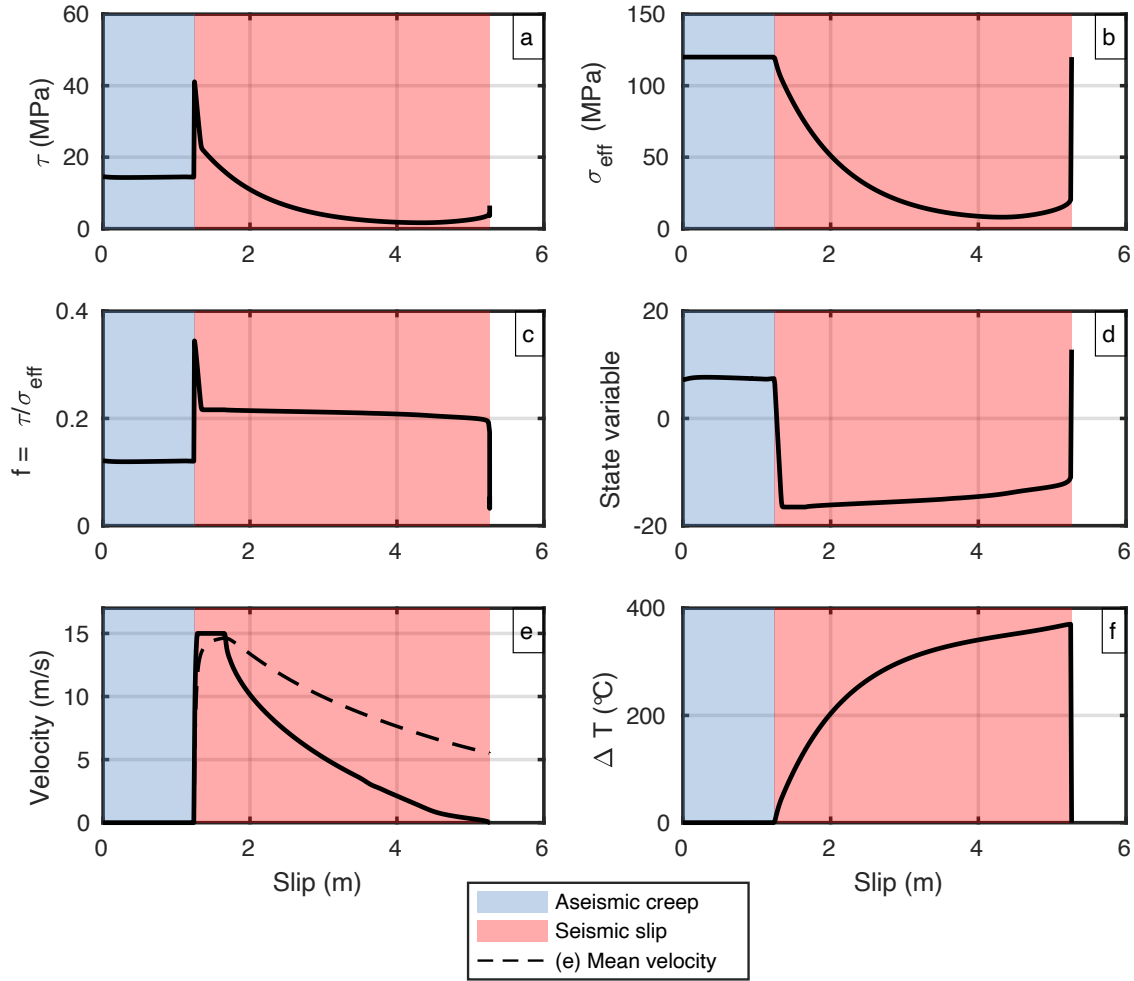


Figure 4.2: Evolution of parameters with slip for a point within the creeping section (position -15 km) during a M_w 7-style event that fully ruptures the creeping section. TP parameters are $\Lambda = 0.34$ MPa/K, $w = 10$ mm, and $\alpha_{hy} = 10^{-5}$ m²/s. **(a)** Shear stress (τ) with slip. **(b)** Effective normal stress (σ_{eff}) with slip. **(c)** Coefficient of friction (τ/σ_{eff}) with slip. **(d)** State variable term ($\ln[V^*\theta/D_{RS}]$) with slip. **(e)** Slip velocity with slip. The dashed line indicates the mean slip velocity since the start of the event. Note the 15 m/s slip velocity limit. **(f)** Temperature change with slip.

and v_{seis} is the representative seismic slip rate, f is the representative dynamic friction coefficient (both assumed constant), δ is the accumulating slip, p_0 is the ambient pore fluid pressure, and the other quantities are as defined above.

The case of adiabatic, undrained shear over the zone of width h gives (Lachenbruch, 1980):

$$\tau = f(\sigma - p_o) \exp\left(-\frac{\delta}{\delta_c}\right), \quad (4.9)$$

where:

$$\delta_c = \frac{\rho ch}{f \Lambda}. \quad (4.10)$$

Note that this solution assumes that shearing is uniformly distributed over the layer of thickness h , compared to our simulations which use a Gaussian distribution of shear rate with root-mean-square half-width w . We assume the approximation $h = w$ when evaluating δ_c for our simulations.

4.3.2 Estimating When TP Allows Propagation Through the Creeping Section

If slip propagates into the creeping section, there would be an initial increase in stress due to the velocity-strengthening properties of the fault, potentially followed by weakening due to TP (Figure 4.3). To estimate the conditions under which TP would allow an event to rupture the creeping section, we estimate the slip required for TP to overcome the velocity-strengthening effect and cause a positive stress drop (Noda & Lapusta, 2013). For a point in the creeping section, undergoing stable creep at the plate rate, v_{pl} , its initial stress τ_{creep} is set by the rate-and-state friction properties of the interface:

$$\tau_{creep} = \sigma_{eff} \left[f^* + (a - b) \ln\left(\frac{v_{pl}}{V^*}\right) \right]. \quad (4.11)$$

The incoming rupture would initially cause the stress to jump to a peak value, the so-called “direct effect” of rate-and-state friction (e.g., Scholz (1998)). Assuming a representative seismic slip velocity, v_{seis} , is supported for a sufficient amount of slip (see Equation 4.21) due to incoming dynamic loading, this stress would then evolve to its steady-state value, τ_{seis} :

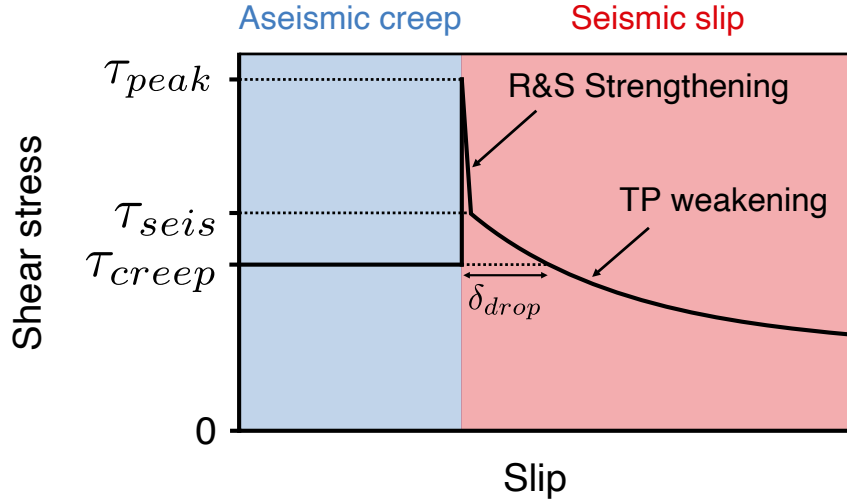


Figure 4.3: Schematic evolution of shear stress, τ , with slip for a point in the creeping section of the fault that experiences TP due to incoming dynamic rupture. The quantities illustrated are used in the theoretical estimates of Section 4.3.1. Steady aseismic creep happens with the slip velocity v_{pl} and stress τ_{creep} . When an event propagates into the creeping section, there is an initial rapid increase of stress to τ_{peak} (the rate-and-state direct effect) and then a decay to τ_{seis} , controlled by the steady-state rate-and-state parameter. TP can then be activated due to the resulting shear heating, causing the stress to continue to decrease. Given sufficient slip before the rupture arrests, the stress could decrease back to τ_{creep} over a slip of δ_{drop} , and then even below this level, allowing for a positive stress drop in the creeping section. A simulation output of this evolution can be seen in Figure 4.2(a).

$$\tau_{seis} = \sigma_{eff} \left[f^* + (a - b) \ln \left(\frac{v_{seis}}{V^*} \right) \right]. \quad (4.12)$$

The length scale of evolution to steady-state is controlled by the rate-and-state critical slip distance, D_{RS} , which is set to 4 mm for our simulations. This value is large compared to laboratory results, but values on this scale are often used in computational work to make the problem tractable (e.g., Barbot et al. (2012), Erickson et al. (2020), and Noda and Lapusta (2013)). Smaller values of D_{RS} would promote rupture, since the slip needed to evolve the shear stress from the larger direct-effect peak to the smaller steady-state seismic resistance would be lower. When applying the slip-on-a-plane and adiabatic, undrained theoretical solutions given above, we assume both a constant seismic slip velocity and a constant friction coefficient over length scales relevant for TP. These assumptions imply that rate-and-state evolution occurs much faster than TP, which we analyse further in Section 4.5.3.

A necessary condition to sustain the rupture in the creeping section is to achieve a positive dynamic stress drop. Given the velocity-strengthening properties of this part of the fault the shear stress is greater when sliding and seismic rates ($\tau_{seis} > \tau_{creep}$), and TP must lower the stress to below τ_{creep} for the stress drop to be positive. The amount of slip required for such stress evolution, δ_{drop} , can be estimated from the analytical solutions presented in Section 4.3.1. We define $\delta_{drop}^{(c)}$ as the estimate of δ_{drop} assuming adiabatic, undrained conditions, and δ_{drop}^* as the estimate assuming slip on a plane. For the adiabatic, undrained solution, the slip for a positive stress drop, δ_{drop} can be derived from Equation 4.9:

$$\delta_{drop}^{(c)} = \delta_c \ln \left(\frac{\tau_{seis}}{\tau_{creep}} \right). \quad (4.13)$$

For the slip-on-a-plane solution, Equation 4.7 gives us an expression for required slip:

$$\exp \left(\frac{\delta_{drop}^*}{L^*} \right) \operatorname{erfc} \left(\sqrt{\frac{\delta_{drop}^*}{L^*}} \right) = \frac{\tau_{creep}}{\tau_{seis}}, \quad (4.14)$$

which can be solved numerically.

To calculate τ_{seis} and δ_{drop}^* , we must choose a seismic slip rate that is representative for a point on the edge of the creeping section over the slip range δ_{drop} . At a point within the velocity-weakening patch, but close to the creeping section, our simulations give an average slip velocity slightly greater than 6 m/s over the entire event for both the M_w 6 and M_w 7-style events. Based on these averages, we choose $v_{seis} = 6$ m/s. We provide further discussion of this value in Section 4.5. For our chosen velocity-strengthening rate-and-state parameters (Table 4.1), and $v_{seis} = 6$ m/s, $\tau_{creep} \approx 15$ MPa, $\tau_{seis} \approx 25$ MPa, and we can calculate $\delta_{drop}^{(c)} \approx 0.53\delta_c$ and $\delta_{drop}^* \approx 0.30L^*$.

4.3.3 The Transition Between Adiabatic, Undrained Solutions and Slip-on-a-Plane Solutions

Our simulated faults have finite widths and non-zero diffusivities, so are never truly in the slip-on-a-plane or adiabatic/undrained regimes. We therefore need to consider the circumstances under which each of these models would be a useful approximation to the full solution. The adiabatic, undrained case would be most applicable when the diffusion length scales are small compared with the width of

the fault for the duration of seismic slip. The slip-on-a-plane solution would be most applicable when the diffusion length scales are large with respect to the fault width (Rempel & Rice, 2006).

Garagash (2012) and Viesca and Garagash (2015) define a lumped hydrothermal diffusivity, $\alpha = (\sqrt{\alpha_{hy}} + \sqrt{\alpha_{th}})^2$, which they use to calculate a length scale for hydrothermal diffusion, $l_d = \sqrt{4\alpha t}$. From this length scale, they derive a hydrothermal diffusion timescale, the time at which the hydrothermal diffusion length scale is equal to the fault zone width: $t_d = h^2/(4\alpha)$.

Assuming that the fault slips at a constant rate v_{seis} , the distance slipped during the diffusion timescale would be $\delta_d = v_{seis}t_d$. Taking the definition of t_d , L^* and δ_c from above, Viesca and Garagash (2015) derived a relationship for the δ_d in terms of the weakening length scales δ_c and L^* :

$$\delta_d = \frac{\delta_c^2}{L^*}, \quad (4.15)$$

which can also be expressed directly in terms of fault zone parameters as:

$$\delta_d = \frac{h^2 v_{seis}}{4\alpha}. \quad (4.16)$$

This consideration implies that, for slip $\delta \ll \delta_c^2/L^*$, the diffusive length scale would be small compared with the fault width, so the shear zone would be in the adiabatic, undrained regime. $\delta \gg \delta_c^2/L^*$ implies that the diffusive length scale is much wider than the fault zone, making the slip-on-a-plane solution the better approximation.

For our parameter combinations, the hydraulic diffusivity varies over several orders of magnitude, and is at least an order of magnitude greater than the thermal diffusivity, resulting in the lumped hydrothermal diffusivity being approximately equal to hydraulic diffusivity. δ_d therefore is best used to determine the slip at which fluid diffusion reaches the scale of the fault width, i.e. the boundary between undrained and drained models. The smaller value of thermal diffusivity means that greater slip than δ_d would be required before thermal diffusion reaches the scale of the fault width and the slip can no longer be approximated as adiabatic. This slip can be estimated from the expression $\delta_{d,th} = (\delta_c^2/L^*)(\alpha/\alpha_{th})$. In this work, we are primarily interested in the variation in fault strength, which is governed directly by the evolution of pore fluid pressure. As δ_d approximately determines the transition from undrained to drained behavior, we use this value to determine which end-member

model best determines the evolution of pore fluid pressure, and thus fault strength, ignoring the fact that this is not the same as the adiabatic boundary. See Section 3.1 of Rice (2006), as well as Rempel and Rice (2006), for more discussion of this issue.

To determine which end-member model of TP best approximates the behavior over the slip required for a positive stress drop, δ_{drop} , we compare the diffusive length scale δ_d (Equation 4.15) to the relevant expression for δ_{drop} , which depends on the chosen model. For the adiabatic, undrained case, we have:

$$\frac{\delta_d}{\delta_{drop}^{(c)}} \approx 1.90 \left(\frac{\delta_c}{L^*} \right) \quad (4.17)$$

$$= 1.90 \left(\frac{f \Lambda h v_{seis}}{4 \rho c \alpha} \right). \quad (4.18)$$

For the slip-on-a-plane case:

$$\frac{\delta_d}{\delta_{drop}^*} \approx 3.28 \left(\frac{\delta_c^2}{L^{*2}} \right) \quad (4.19)$$

$$= 3.28 \left(\frac{f \Lambda h v_{seis}}{4 \rho c \alpha} \right)^2. \quad (4.20)$$

Note that both expressions contain the same physical parameters. The adiabatic, undrained solution would be a good approximation over the slip length required for a positive stress drop when $\delta_d/\delta_{drop}^{(c)} \gg 1$, i.e. $\delta_c \gg 0.53L^*$. The slip-on-a-plane solution would be a good approximation when $\delta_d/\delta_{drop}^* \ll 1$, i.e. $\delta_c \ll 0.55L^*$. Therefore, there is a transition at around $\delta_c \approx L^*/2$ for which solution is a better approximation, but the ratio between δ_d and δ_{drop} scales differently between the two end-member solutions as the physical parameters are varied.

4.4 Rupture Propagation Into the Creeping Section

4.4.1 Simulation results

As we vary the parameters of TP in the velocity-strengthening section, the simulated ruptures span the range from rapidly arresting to propagating through the entire section. In Figures 4.4 and 4.5, we show the fraction of the creeping section that is ruptured for M_w 6 and M_w 7-style events, respectively, as we vary the hydraulic

diffusivity (α_{hy}), fault half-width (w), and TP coupling (Λ) over several orders of magnitude. For $\Lambda = 0.01$ MPa/K, there is negligible influence of TP for both event sizes and all combinations of α_{hy} and w within our chosen ranges, and the rupture rapidly dies out due to the velocity-strengthening rate-and-state properties of the creeping section. TP also has negligible effects when $w = 100$ mm. Some enhanced propagation into the creeping section is possible when $\alpha_{hy} = 10^{-2}$ m²/s, but the creeping section never fully ruptures when the hydraulic diffusivity is this high.

The coupling coefficient of $\Lambda = 0.1$ MPa/K is sufficient to allow M_w 6-style ruptures to propagate several kilometers into the creeping section for the thinnest faults and lowest diffusivities (Figure 4.4(b)). For $\Lambda = 0.34$ and 1 MPa/K, we find numerous parameter combinations for which M_w 6-style events rupture the entire creeping section, resulting in events much larger than M_w 6 (Figure 4.4(c)-(d)). Using the observation that the 2004 M_w 6 Parkfield event arrested in the creeping section, we can rule out these fault models as having overly efficient TP. We mark these parameter combinations with black boxes.

For the M_w 7-style events shown in Figure 4.5, the larger event size results in greater propagation through the creeping section for every parameter combination. For $\Lambda = 0.1$ MPa/K and larger, we find several combinations of parameters for which the M_w 6-style events arrest in the creeping section, but the M_w 7-style events rupture the whole fault. These parameter combinations are highlighted by red boxes. Note we also run simulations for $\Lambda = 0.069$ MPa/K. These results are included in Figures 4.8 and 4.9 but are not plotted in Figures 4.4 or 4.5.

4.4.2 Application of Analytical Models

For each combination of parameters, we use the analytical solutions of Section 4.3.1 to estimate the slip required for TP to give a positive stress drop, δ_{drop} , and compare these values to the fraction of the creeping section ruptured by M_w 7-style events (Figures 4.6(a)-(b) and 4.7(a)-(b)).

We also evaluate δ_d/δ_{drop} for each set of parameters; this ratio allows us to determine whether conditions assumed for the analytical solution are a good approximation to our models (see Section 4.3.3). For the adiabatic, undrained solution to be a good approximation, we need the hydrothermal diffusion length scale to be small in comparison to the fault width over the relevant slip distance, and we set a threshold of $\delta_{drop}^{(c)} < 0.1\delta_d$ (recall that δ_d is the slip required for diffusion scales to be comparable to the fault zone width). For the slip-on-a-plane solution, the approximation should

Mw 6 Event

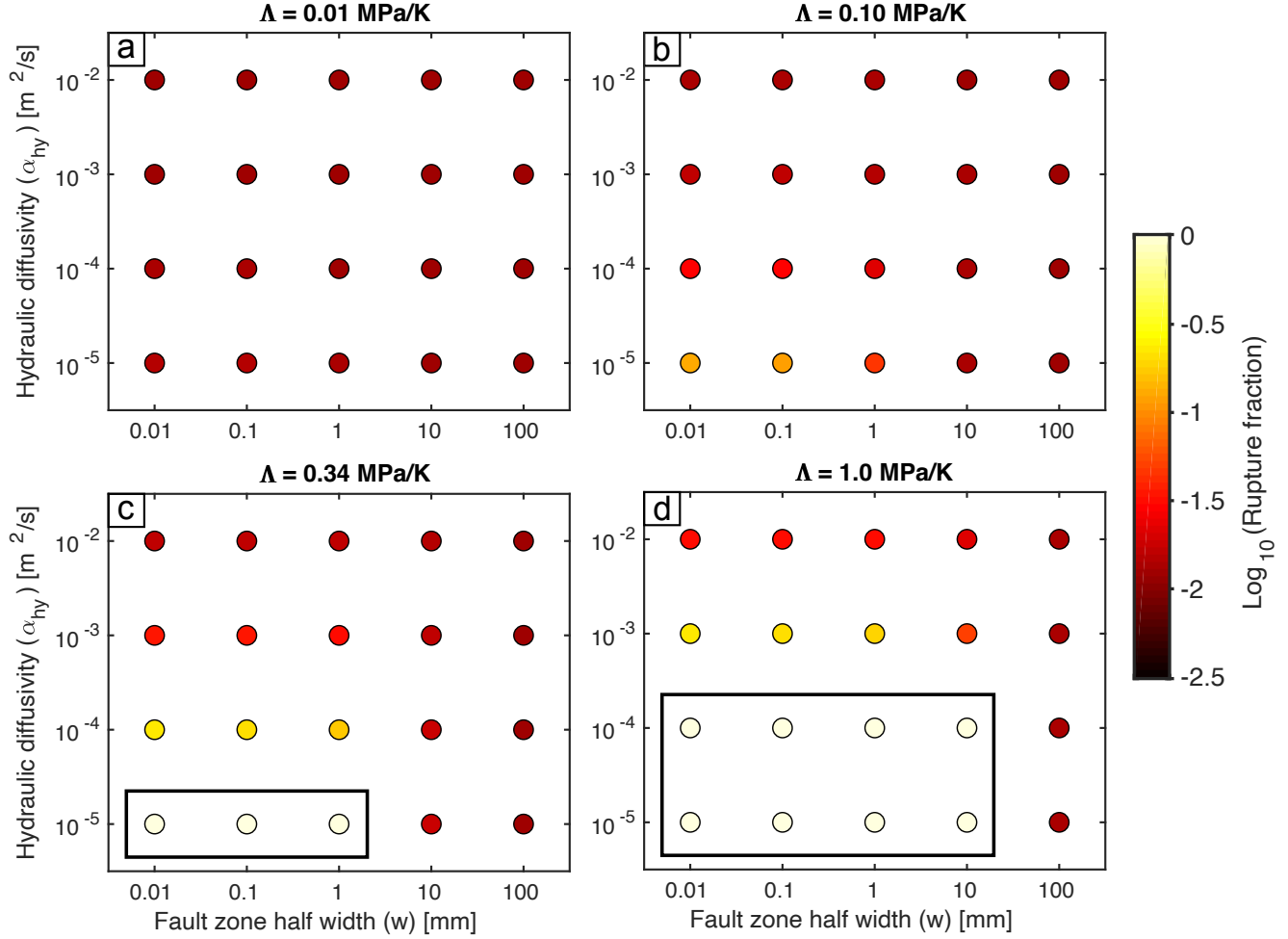


Figure 4.4: Fraction of the creeping section ruptured for a M_w 6-style event and varying TP parameters. The value of the TP coupling coefficient, Λ , increases from (a) to (d), resulting in increasingly efficient TP and greater rupture fractions. Points within black squares are those models for which the event fully ruptures the creeping section, indicating that TP is too efficient to be consistent with the observation of the 2004 M_w 6 event arresting in the creeping section.

Mw 7 Event

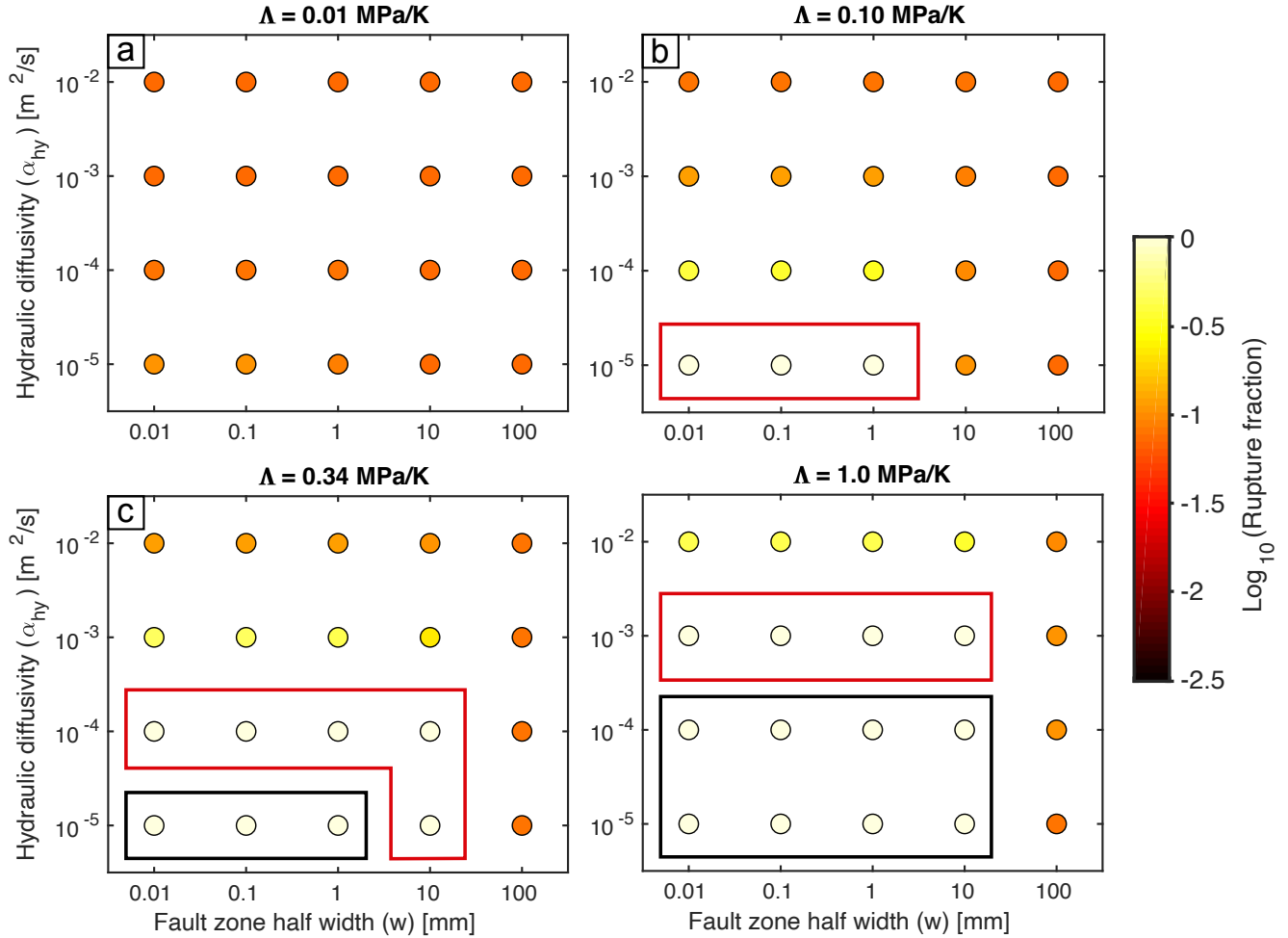


Figure 4.5: Fraction of the creeping section ruptured for a M_w 7-style event and varying TP parameters. The size of the TP coupling coefficient, Λ , increases from (a) to (d), resulting in increasingly efficient TP and greater rupture fractions. Points within black boxes are those that fully ruptured the creeping section for M_w 6-style events (Figure 4.4), indicating that TP is too efficient to be consistent with the observation of the 2004 M_w 6 event arresting in the creeping section. Points within red boxes are those that arrested in the creeping section for M_w 6-style events, but fully rupture the creeping section for M_w 7-style events.

work best for faults that are thin compared to the diffusion lengths, and we require the hydrothermal diffusion length scale to be much larger than the fault width. This condition can be expressed as $\delta_{drop}^* > 10\delta_d$. In Figures 4.6(c)-(d) and 4.7(c)-(d), we use these thresholds of δ_d/δ_{drop} to illustrate the expected validity of the analytical solution, indicating which simulations we expect to be well approximated by the slip-on-a-plane and adiabatic, undrained regimes. For clarity, we shade out areas in Figures 4.6(a)-(b) and 4.7(a)-(b) where the relevant analytical solution is invalid according to our chosen bounds. These plots demonstrate that lower hydraulic diffusivities and larger fault widths move us towards the adiabatic, undrained regime, as they favor the retention of heat and pressure within the fault zone. Increasing the TP coupling coefficient, Λ , shortens the distance δ_{drop} , resulting in less time for hydrothermal diffusion and therefore also shifting us towards the adiabatic, undrained regime (see Equations 4.17 and 4.14).

One would expect that if the dynamic rupture brings a sufficient amount of seismic slip to the velocity-strengthening region, in comparison with δ_{drop} , the associated shear heating could activate TP and allow the rupture to propagate. This is exactly what we observe. Specifically, in Figures 4.6(a)-(b) and 4.7(a)-(b), we highlight the point at which δ_{drop} is equal to the M_w 7 slip (δ_{event}) with a red dashed line. For models that lie within their approximation bounds, this contour marks a transition point for the fraction of the creeping section ruptured. Where $\delta_{event} > \delta_{drop}$, which indicates that the incoming event brings sufficient seismic slip for a positive stress drop from TP, we see a substantial or complete rupture of the creeping section, whereas $\delta_{event} < \delta_{drop}$ results in the events rapidly arresting in the creeping section.

To summarize the outcomes for all combinations of parameters in a single figure, we plot the fraction of the creeping section ruptured by each event against the ratio between the total slip of the incoming rupture pulse without TP, δ_{event} , and the estimated slip required for TP to give a positive stress drop, δ_{drop} . ($\delta_{event} \approx 66$ cm and 2.2 m for the M_w 6 and M_w 7-style events, respectively). We compute δ_{drop} using both the adiabatic, undrained solution ($\delta_{drop}^{(c)}$) and the slip-on-a-plane solution (δ_{drop}^*). We then separate data points with estimates δ_{drop} obtained from analytical models within our chosen bounds, shown in Figure 4.8, and the rest, shown in Figure 4.9.

Figure 4.8 clearly highlights the transition in the fraction of the creeping section ruptured as the total event slip becomes sufficient to create a positive stress drop in the velocity-strengthening region. When $\delta_{event} \ll \delta_{drop}$, TP is inefficient at

overcoming rate-and-state velocity-strengthening properties, and the (small) rupture fraction is determined mainly by the amount of strengthening set by the rate-and-state parameters. As the M_w 7 rupture brings more slip than the M_w 6 event, it can propagate a longer distance into the creeping section without TP, resulting in a larger minimum rupture fraction. Between $\delta_{event} = \delta_{drop}$ and $\delta_{event} = 10\delta_{drop}$, the simulations give at least partial ruptures of the creeping section (most in the range of 10-50%), and for $\delta_{event} > 10\delta_{drop}$ all simulations rupture the entire creeping section.

Figure 4.9 shows that estimates of δ_{drop} based on less relevant analytical models do an overall poor job of estimating when the creeping section would be ruptured. The adiabatic, undrained solution neglects the transport of heat and fluid out of the fault zone, and therefore gives the upper limit on the efficiency of TP for each combination of parameters. The adiabatic, undrained analytical approach applied to conditions where diffusion is significant, i.e. $\delta_{drop}^{(c)} > \delta_d$, often suggests that the creeping section should be fully ruptured (i.e. $\delta_{event} \gg \delta_{drop}^{(c)}$) when this is not observed, as can be seen for the large number of points in the bottom right of the plots in Figure 4.9. Similarly, when the slip-on-a-plane solution is applied to conditions for which $\delta_{drop} < \delta_d$, the finite width of the fault zone results in more distributed heat generation and thus less efficient TP than the solution predicts. The symbols with white fill represent approximations of δ_{drop} that can still be good, just somewhat outside the bounds established earlier, and some of them may belong in Figure 4.8.

To study the validity bounds of each approximation (rather than assuming factors of 10 as done so far), we can compare the analytical estimates of δ_{drop} with its values from our simulations. This is the direction of ongoing work.

4.4.3 Summary of Expressions

Summarizing the expressions defined above, we have the slip-on-a-plane and adiabatic, undrained TP weakening length scales (Equations 4.8 and 4.10):

$$L^* = 4 \left(\frac{\rho c}{f \Lambda} \right)^2 \frac{(\sqrt{\alpha_{hy}} + \sqrt{\alpha_{th}})^2}{v_{seis}},$$

$$\delta_c = \frac{\rho c h}{f \Lambda}.$$

We also have the shear stress values at creep and seismic velocities (Equations 4.11 and 4.12):

$$\begin{aligned}\tau_{creep} &= \sigma_{eff} \left[f^* + (a - b) \ln \left(\frac{v_{pl}}{V^*} \right) \right], \\ \tau_{seis} &= \sigma_{eff} \left[f^* + (a - b) \ln \left(\frac{v_{seis}}{V^*} \right) \right].\end{aligned}$$

We use the TP solutions from Equations 4.7 and 4.9 to define the slip required for a positive stress drop from TP for each end-member solution (Equations 4.13 and 4.14):

$$\begin{aligned}\delta_{drop}^{(c)} &= \delta_c \ln \left(\frac{\tau_{seis}}{\tau_{creep}} \right) \\ \exp \left(\frac{\delta_{drop}^*}{L^*} \right) \operatorname{erfc} \left(\sqrt{\frac{\delta_{drop}^*}{L^*}} \right) &= \frac{\tau_{creep}}{\tau_{seis}}.\end{aligned}$$

Finally, we can use the ratio of the hydrothermal diffusion slip and the relevant δ_{drop} to determine the applicability range of each solution (Equations 4.17 and 4.19):

$$\begin{aligned}\frac{\delta_d}{\delta_{drop}^{(c)}} &\approx 1.90 \left(\frac{\delta_c}{L^*} \right), \\ \frac{\delta_d}{\delta_{drop}^*} &\approx 3.28 \left(\frac{\delta_c^2}{L^{*2}} \right).\end{aligned}$$

We choose $\delta_{drop}^{(c)} < 0.1\delta_d$ for adiabatic, undrained solutions to be applicable, and $\delta_{drop}^* > 10\delta_d$ for slip-on-a-plane solutions to be applicable.

Our results suggest an approximate approach for considering the potential effect of different model parameters, including TP parameters, and different incoming seismic ruptures, without dynamic modeling, by comparing potential δ_{event} and analytically estimated δ_{drop} . Specifically:

- If $\delta_{event} > 10\delta_{drop}$, then full rupture.
- If $\delta_{drop} < \delta_{event} < 10\delta_{drop}$, then partial to full rupture.
- If $\delta_{event} < \delta_{drop}$, then no rupture.

where $\delta_{drop} = \delta_{drop}^*$ for the slip-on-a-plane regime and $\delta_{drop} = \delta_{drop}^{(c)}$ for the adiabatic, undrained regime.

Mw 7-Style Event Slip-on-a-plane Model

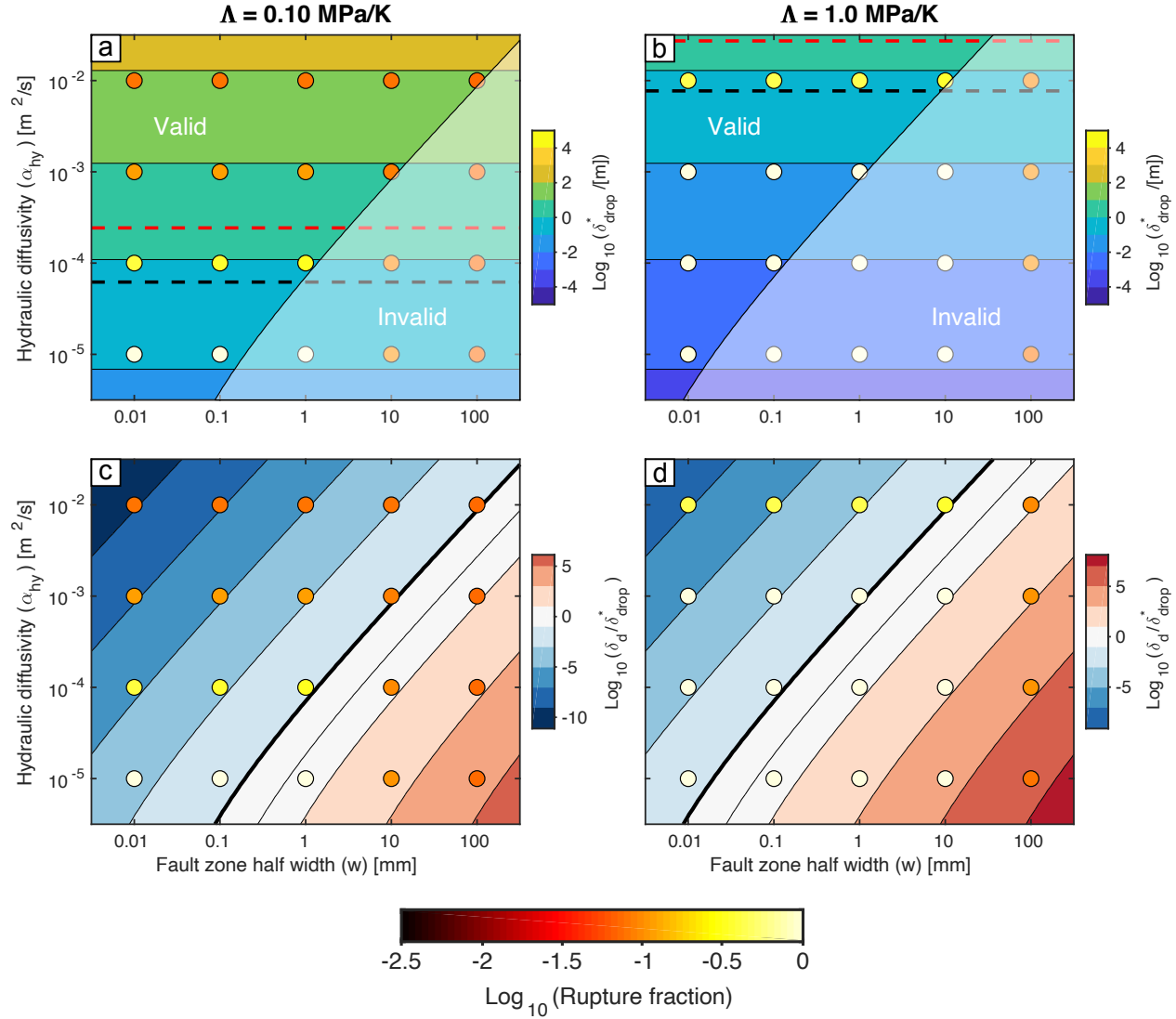


Figure 4.6: Variation in the slip δ_{drop}^* (top row) required for a positive stress drop, assuming the slip-on-a-plane solution, as we change TP parameters w , α_{hy} and Λ , and validity of the approximation (bottom row). The simulated rupture fraction values are also shown, for a Mw 7-style event. **(a)** Colorbar shows variation of δ_{drop}^* with w and α_{hy} , for $\Lambda = 0.1$ MPa/K. Each color change delineates an order of magnitude change in δ_{drop}^* , which only varies with α_{hy} given the assumption of an infinitesimally thin fault in the corresponding analytical solution. The approximation is expected to work for faults that are thin enough compared to diffusion time scales; this regime is highlighted with bright colors (see panel (c)), with the rest of the panel muted. The dashed black and red lines show where $\delta_{drop}^* = \delta_{event}$ for the Mw 6-style and Mw 7-style events, respectively. Note that, within the region of the validity of the analytical approximation (bright colors), the red dashed line predicts the transition from rupture propagation over the creeping section to rupture

arrest. **(b)** Same as (a), except with Λ increased by an order of magnitude to 1 MPa/K. This increased efficiency of TP results in δ_{drop}^* decreasing by two orders of magnitude for each value of α_{hy} . **(c)** Validity for the slip-on-a-plane approximation with $\Lambda = 0.1$ MPa/K. The colorbar shows variation of δ_d/δ_{drop}^* . The thick black contour delineates $\delta_{drop}^* = 10\delta_d$. The blue area to the left and top of this line satisfies $\delta_{drop}^* > 10\delta_d$, the range for which we take the slip-on-a-plane fault model to be a good approximation, as marked in (a). **(d)** Same as (c), except with $\Lambda = 1$ MPa/K. This increased efficiency of TP results in a reduced δ_{drop}^* , as shown in (b). This decrease causes the validity bound for the slip-on-a-plane approximation to shift to the left, reducing the number of our fault models that are well approximated by the slip-on-a-plane fault regime.

4.5 Discussion

4.5.1 Temperature Variation

Coffey et al. (2022) used measurements of biomarker thermal maturity from SAFOD samples to infer that peak temperatures on the fault had reached the range 570-1100 °C. They suggested that a large number of ruptures had broken the fault at this point in the creeping section. Based on assumptions about the evolution of friction during sliding, they calculated a slip range of 0.5-2.9 m associated with these peak temperatures. Our results provide a possible mechanism for these ruptures, consistent both with the arrest of M_w 6 ruptures, and the propagation of larger events with temperatures consistent with the results of Coffey et al. (2022).

For example, with $\Lambda = 0.34$ MPa/K, $w = 10$ mm and $\alpha_{hy} = 10^{-5}$ m²/s, M_w 6-style events rapidly arrest in the creeping section, but M_w 7-style events are able to rupture all the way through. Figure 4.2 illustrates the evolution of the fault parameters with slip for a point on the creeping section at the approximate location of the SAFOD drill site in our model. As shown in Figure 4.2(f), TP allows for 4 meters of slip with a peak temperature change of 370 °C. Assuming an ambient temperature of 200 °C, representative of around 7 km depth, these results give a peak temperature of 570 °C, consistent with the measurements of Coffey et al. (2022).

However, the SAFOD samples are taken from 2.7 km depth, where ambient temperatures are closer to 110 °C (Lockner et al., 2011). We would also expect the stress state and efficiency of TP to vary with depth (Brantut & Platt, 2017), meaning our results cannot be straightforwardly applied to inferred rupture parameters at SAFOD. Our results do indicate how incorporating TP into the evolution of shear strength could allow for slip significantly greater than the 2.9 m upper bound calculated by Coffey et al. (2022), with peak temperatures still in their measured range.

Mw 7-Style Event Adiabatic, Undrained Model

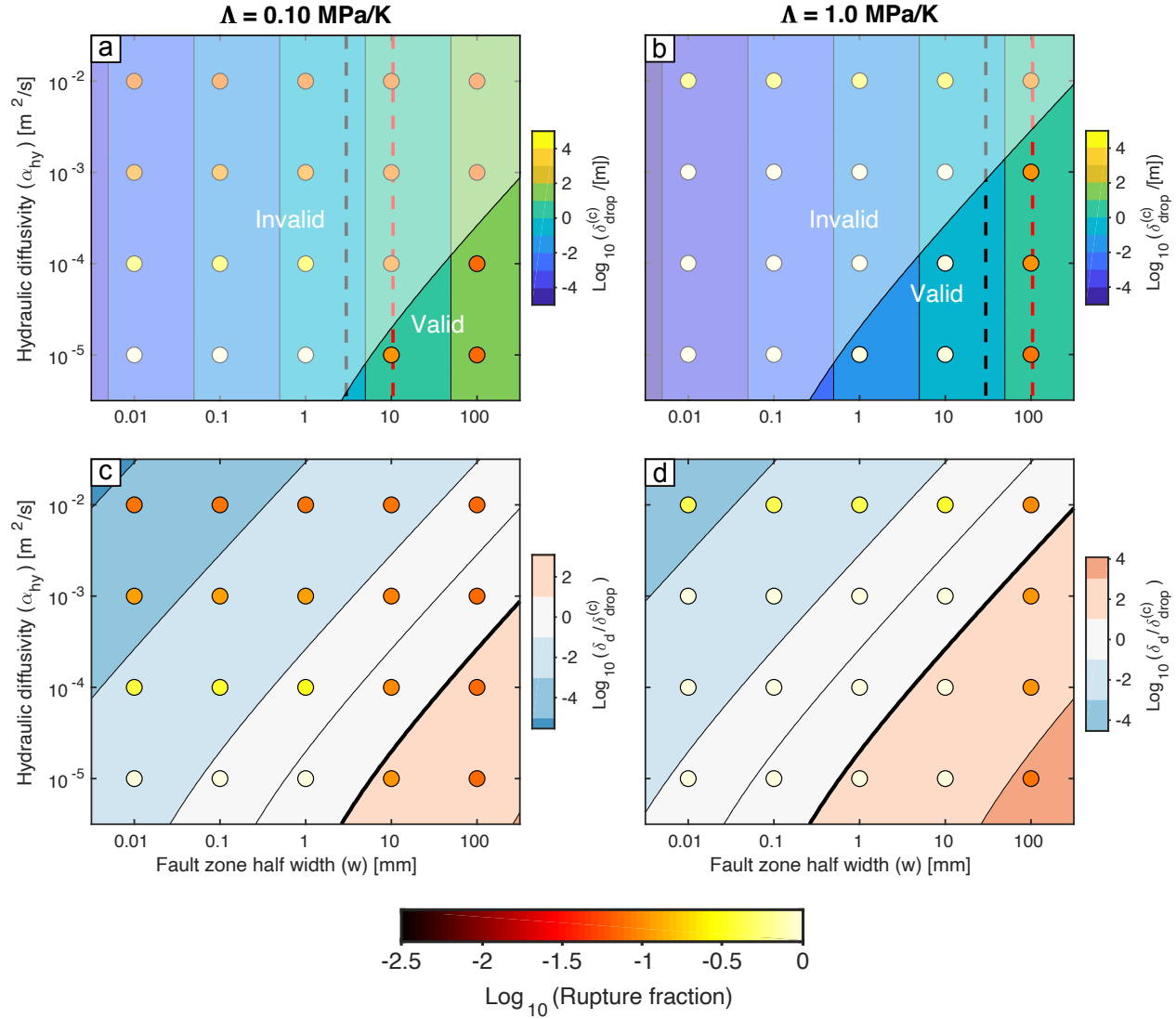


Figure 4.7: Same as Figure 4.6, except for the adiabatic, undrained fault solution (i.e. replacing δ_{drop}^* with $\delta_{drop}^{(c)}$). Note that the order of magnitude change in Λ between (a) and (b) causes an order of magnitude change in $\delta_{drop}^{(c)}$, compared to two orders of magnitude for δ_{drop}^* (shown in Figure 4.6). As the adiabatic, undrained fault solution assumes no transfer of fluid or heat out of the fault zone, $\delta_{drop}^{(c)}$ depends only on the fault zone width. The thick black contour in (c) and (d) now delineates $\delta_{drop}^{(c)} = 0.1\delta_d$; models to the right of this line satisfy $\delta_{drop}^{(c)} < 0.1\delta_d$, the range for which we take the adiabatic, undrained solution to be a good approximation to our models. As we go from (c) to (d), the increasing efficiency of TP results in a shorter $\delta_{drop}^{(c)}$, reducing the diffusion length scales and making more parameter combinations well approximated by the adiabatic, undrained regime. Similar to Figure 4.6, the red dashed line in (a) and (b) shows where $\delta_{drop}^{(c)} = \delta_{event}$ and marks a transition in the fraction of the creeping section ruptured.

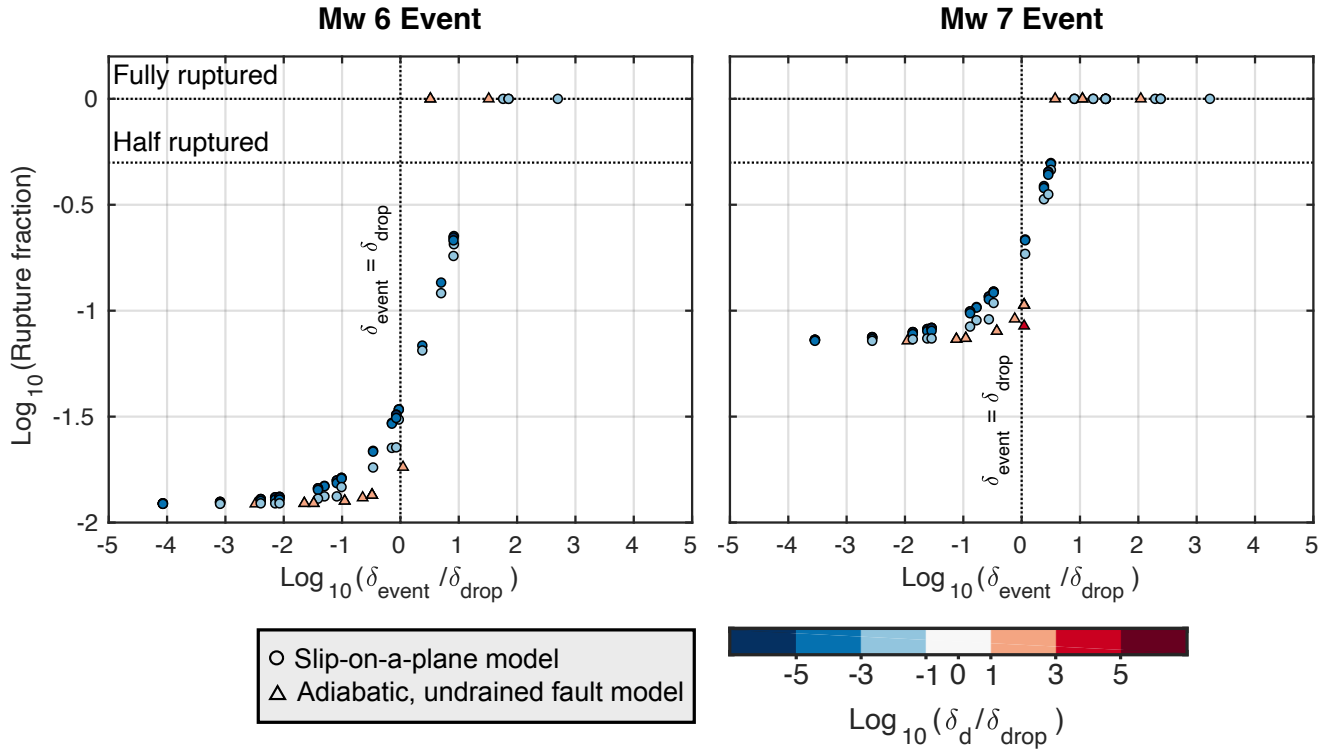


Figure 4.8: Logarithm of the fraction of the creeping section ruptured against the logarithm of the ratio between the total slip for the event in the absence of TP, δ_{event} , and the estimated slip required for a positive stress drop due to TP, δ_{drop} . For each combination of TP parameters, δ_{drop} is estimated for the slip-on-a-plane solution (δ_{drop}^*) and the adiabatic, undrained solution ($\delta_{\text{drop}}^{(c)}$). Only points for estimates that are expected to be good are plotted, defined by $\delta_{\text{drop}}^{(c)} < 0.1\delta_d$ and $\delta_{\text{drop}}^* > 10\delta_d$, as established based on Figures 4.6 and 4.7. Points with estimates of δ_{drop}^* outside of these validity bounds are plotted in Figure 4.9. The color of the symbols gives the log of the ratio between the hydrothermal diffusion distance, δ_d , and δ_{drop} , for the relevant end-member model. Values increasingly greater than 1 indicate that the fault model is increasingly closer to the adiabatic, undrained regime over the slip δ_{drop} , while values less than -1 indicate increasing closeness to the slip-on-a-plane fault regime. The line $\delta_{\text{event}} = \delta_{\text{drop}}$ captures the transition from ruptures arresting in the creeping region to increasingly large fraction of the creeping section being ruptured in our models.

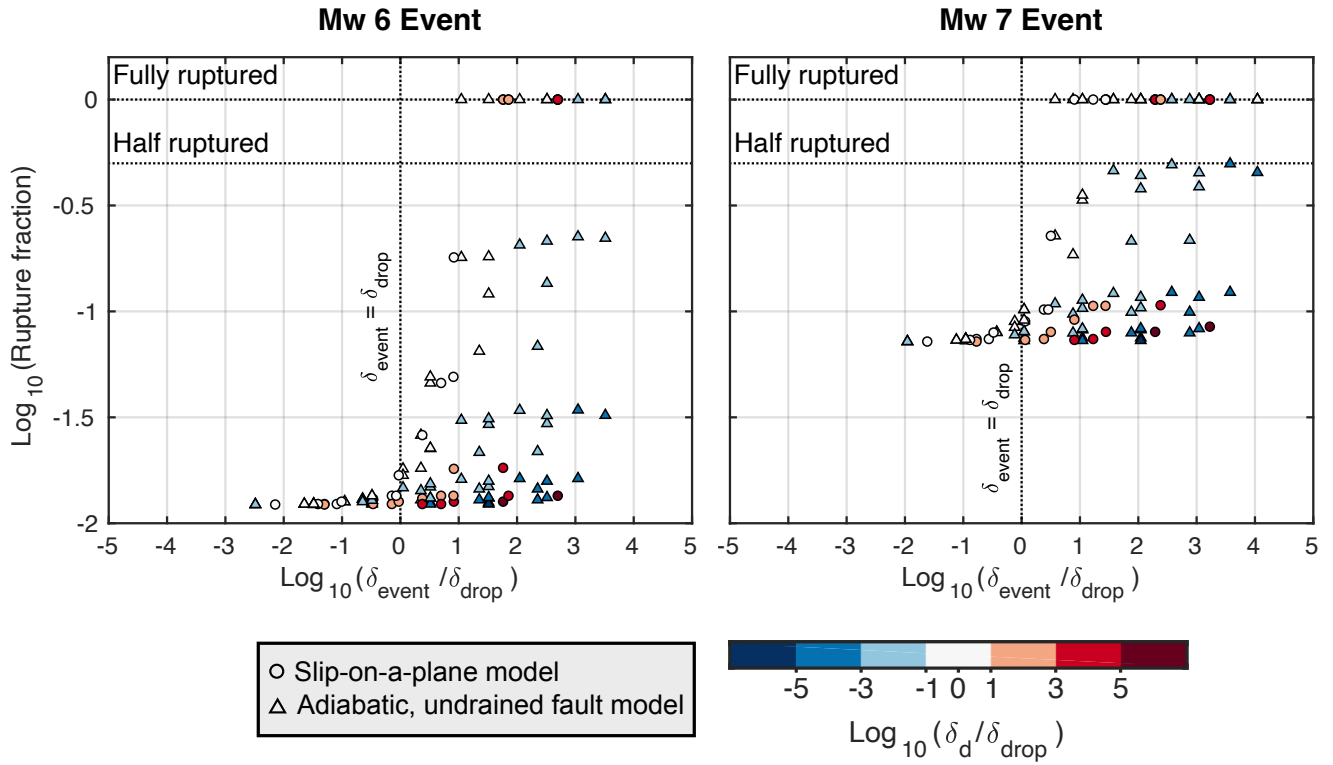


Figure 4.9: Same as Figure 4.8, except for estimates of δ_{drop} that are outside of the validity bounds we choose (i.e. $\delta_{\text{drop}}^{(c)} > 0.1\delta_d$ and $\delta_{\text{drop}}^* < 10\delta_d$).

For events that propagate into the creeping section in our simulations, the peak values of temperature can vary by two orders of magnitude. Lower TP efficiency, that is still sufficient to allow ruptures in the creeping section, promotes larger temperature rises. For some parameter combinations, the temperature change can peak at thousands of degrees, well past the melting threshold of around 1000 °C (Rice, 2006). Such high temperatures are inconsistent with the lack of melting observed in SAFOD cores (Coffey et al., 2022). The assumed fault constitutive law does not incorporate melting, so the onset of melting leads to a lack of self-consistency in our calculations (Noda & Lapusta, 2010). However, as long as the melting temperature is crossed towards the end of the slip, incorporating melting would likely not affect our overall results with regards to the fraction of the creeping section ruptured. Analyzing all our simulation results for peak and average temperatures reached is a subject of ongoing work.

Note that all of our simulations are consistent with the absence of heat anomaly around the SAF (Lachenbruch & Sass, 1980), since the average shear stresses acting

on the creeping section in our models are within the corresponding constraint of 20 MPa. This stress is ensured by the choice of 0.15 as the reference friction coefficient and 120 MPa as the initial effective normal stress, which results in a steady-state value of shear stress at the plate rate in the creeping segment of about 15 MPa, with even lower values during thermal pressurization of pore fluids (e.g., Figure 4.2(a)).

4.5.2 Limitations of our Modeling

Our simplified modeling allows us to consider, in a tractable way, variations of several orders of magnitude in parameters that govern TP. However, the modeling does not account for a number of additional mechanisms that can affect rupture propagation and need to be investigated in future modeling.

As mentioned in the introduction, other dynamic weakening processes have been proposed, including flash heating, gel lubrication, nanoparticle lubrication, decarbonation, dehydration, and fault melting (Di Toro et al., 2011). Not all of these processes would necessarily occur in the clay rich fault gouge found in the creeping section in the SAF, but the conditions under which each mechanism is significant is an important component of future work.

Our models do not include fault dilatancy, which can accompany an increase in slip rate and lead to a drop in pore pressure, resulting in a reduced effective normal stress and increased fault strength (e.g., Segall and Rice (1995)). This effect would contribute to the peak shear resistance of the creeping section as the dynamic rupture enters it, requiring more efficient TP to produce positive stress drops for a given slip. Experimental results and theories of dilatancy are limited to the slow-slip regime and remain subject of active ongoing research; its role in shear, especially rapid, of the mature fault gouge found in the creeping section of the SAF remains uncertain (Brantut, 2020; Rice, 2006).

In order to capture different event styles in 2-D, we use two separate models for the M_w 6 and M_w 7-style ruptures. As a result of the different loading phases before rupture nucleation, the different event styles have different stress heterogeneities at the edge of the creeping section. A more realistic approach would be to capture both the M_w 6 and M_w 7-style events in a single 3-D model, allowing multiple M_w 6 events to rupture before a larger M_w 7 event occurred. 3-D models would also allow us to incorporate depth dependent properties. Such simulations would be much more computationally expensive, potentially requiring weeks of run time, making it challenging to explore a wide range of TP parameters given current computational

constraints. The results of our 2-D parameter exploration can be the starting point for 3-D simulation, such as simulating the cases that arrest our simulated M_w 6-style event and allow our M_w 7-style event to propagate in 2-D.

Our simulated events have average slip rate of about 6 m/s, which is higher than the average of 1 m/s often cited for large events (Heaton, 1990). Higher slip rates increase the rate of shear heating and hence promote TP in models in which heat diffusion is important compared to the layer thickness, such as our fault models with very narrow shear layers that are well approximated by slip-on-a-plane solution (Figure 4.6). However, the effect would be more minor in fault models that are in the adiabatic, undrained regime (Figure 4.7), in which the weakening depends on total shear heat input and hence slip, and not as much on slip rate. Note that the fault model that provides a good match to the field results of Coffey et al. (2022) discussed in Section 4.5.1 is in the latter regime, and hence should not be much affected by its specific average slip rate. Considering the effect of incoming ruptures with different average slip rate (but the same slip) is a subject of ongoing work.

Our models assume a planar fault geometry, although the transition between the locked and creeping sections of the SAF is more complex, with the upper ~ 6 km of the fault showing a warp over a 50 km long section (Simpson et al., 2006). This geometry could act as a barrier to earthquake rupture, even in the presence of efficient TP. The fault is likely more planar at depth (Simpson et al., 2006), and even large stepovers in faults will not necessarily prevent rupture (H. Wang et al., 2020), so while the geometry may reduce the likelihood of rupture propagation for a given set of TP parameters, it would likely not eliminate it entirely.

Measurements at the SAFOD drill hole show that the fault is divided into two actively creeping strands, separated by less than 100 meters (Lockner et al., 2011). Coffey et al. (2022) found that the temperature rise from seismic slip was confined to a zone adjacent to one of the creeping strands of the fault, in a region previously identified as potentially having hosted seismic slip (Bradbury et al., 2011). Experimental tests by French et al. (2014) of SAFOD samples from one of the creeping strands showed dynamic weakening at seismic slip rates, which they attributed to TP. These observations highlight two possibilities, not considered by our model: 1) that a dynamic rupture could break two or more fault strands simultaneously, and 2) that the rupture could break a fault strand that is not actively creeping. If multiple, parallel, faults ruptured simultaneously that may result in lower slip rates and slips on each fault, reducing the amount of TP weakening. If the rupture follows an

alternative path to the currently creeping sections, the rate-and-state parameters inferred for the properties of the creeping section, that we have assumed here, may not apply to the propagation of dynamic ruptures.

We keep the TP parameters constant in time; however, the evolution of TP parameters during an event can lead to significant changes in fault behavior. Rice et al. (2014) and Platt et al. (2014) argued that relatively wide shearing zones would be unstable, resulting in rapid shear localization to a narrow (5-40 μm) principle slip zone during seismic slip. That would imply evolving, and potentially decreasing, shear-zone width during an event, which could promote TP. Recent work by Stathas and Stefanou (2022) has suggested that the principle slip zone would not be stationary within the fault gouge at large slips, potentially resulting in oscillatory behavior in the frictional strength.

The coupling coefficient, Λ , and hydraulic diffusivity, α_{hy} , likely evolve during a rupture, both as a result of damage due to high stresses at the rupture tip, and changing pressure-temperature conditions (e.g., Aben et al. (2020), Badt et al. (2020), Brantut and Mitchell (2018), and Rice (2006)). Our parameters partially account for the effect of damage, as they include values taken from studies that approximated the effect of damage by modifications including increases to the fault gouge permeability and pore space compressibility from the lab-measured values, resulting in larger α_{hy} and lower Λ values, thus less efficient TP (e.g., Rice (2006)). Later work by Brantut and Mitchell (2018) suggested that these modifications gave a reasonable approximation for the impact of damage in faults which showed evidence of melting. To take account of parameter variation with pressure and temperature, Rice (2006) computed quantities averaged along representative pressure-temperature paths. Based on comparisons to solutions with pressure and temperature dependent properties in Rempel and Rice (2006), Rice (2006) argues that these path averages are a reasonable approximation.

The likely time scale and magnitude of the variation in parameters is important when considering the applicability of the models presented in this work. If the initial stress concentration at the rupture tip causes rapid damage-induced changes in TP properties, and further evolution, e.g., due to changes in pressure and temperature, is mild, then our use of fixed parameters during rupture could be justified. However, a larger event may create more significant damage due to higher stress concentrations, causing different events to be governed by different TP parameters. Fully incorporating

these effects into modeling efforts would require a detailed understanding of how material properties evolve during rupture (e.g., Aben et al. (2020)).

4.5.3 Applicability of the Analytical Solutions

We apply analytical solutions for both slip on a mathematical plane and adiabatic, undrained conditions to gain physical insight into our simulations (Lachenbruch, 1980; Mase & Smith, 1987; Rice, 2006). Both solutions assume a constant coefficient of friction, f , and the thin fault solution assumes a constant seismic velocity, v_{seis} . To apply the analytical solutions, we set $f = 0.15$ and $v_{seis} = 6$ m/s. In the following, we discuss these choices and their limitations.

It is common to use a v_{seis} of 1 m/s in evaluating L^* , based on the results of Heaton (1990) (e.g., Noda and Lapusta (2013), Rempel and Rice (2006), and Rice (2006)). In our work, we have chosen 6 m/s, based on the average slip velocity on the edge of the velocity-weakening patch (discussed above). However, this average conceals substantial variation, with the slip velocity rapidly increasing to the limit of 15 m/s in under 5 cm of slip, remaining at the peak slip velocity for around 10 cm and 40 cm for the M_w 6 and M_w 7-style events, respectively, then decaying back down over the remaining slip.

The variation of slip velocity in the creeping section is highly dependent on the strength of the TP weakening. For strong TP, the slip velocity increase is even more rapid than within the velocity-weakening section; when TP is weaker the slip rate evolution is less abrupt; and there is no seismic slip over most of the creeping section when the TP falls below a certain threshold. The total slip of the incoming event, δ_{event} , the variation of the slip velocity during the event, and the ratio of δ_{event} to δ_{drop} , will all affect the most representative velocity to use when we look to understand when an event will propagate (see Figure 4.2 for an example). As the average slip rate over our events without TP is approximately 6 m/s, we use a value of $v_{seis} = 6$ m/s in the calculation of L^* ; however, we acknowledge that our choice, and the assumption of constant velocity, are highly approximate.

The evolution of the rate-and-state friction coefficient, $f = \tau/\sigma_{eff}$, is governed by both by the slip rate and the rate-and-state characteristic slip distance D_{RS} . The most significant change is during the initial evolution from creep to seismic slip, where the friction coefficient jumps to a peak value (the direct effect), then evolves to a steady state value. By using a constant coefficient of friction when calculating L^* and δ_c , we assume both a constant slip velocity, and that the rate-and-state evolution

of the friction coefficient occurs rapidly compared to the TP weakening, therefore allowing us to calculate δ_{drop} assuming steady-state sliding.

The slip over which the rate-and-state friction coefficient evolves from steady-state creep to steady-state seismic slip can be estimated from the equation (Bizzarri & Cocco, 2003):

$$D = D_{RS} \ln \left(\frac{v_{seis}}{v_{creep}} \right). \quad (4.21)$$

The creeping section moves at the plate rate, $v_{creep} = v_{pl} = 23$ mm/yr in steady state, and we set $v_{seis} = 6$ m/s, giving a value of $D \approx 20D_{RS}$, or around 8 cm for our D_{RS} value of 4 mm.

We show an example of the evolution of stress within the creeping section in Figure 4.2(a). This simulation lies within the adiabatic, undrained regime, and the slip δ_{drop} needed to achieve positive stress drop in the creeping region should be well approximated by $\delta_{drop}^{(c)}$, which is ~ 56 cm. Figure 4.2(a) shows an initial quasi-linear weakening over < 10 cm of slip (governed by the state variable shown in Figure 4.2(e)), consistent with the estimate D of Equation 4.21, followed by ~ 40 cm of exponential decay in shear stress until it falls below its initial value, τ_{creep} . Hence the δ_{drop} from the numerical simulation is indeed well approximated by the analytically determined $\delta_{drop}^{(c)}$. In this case, RS friction evolves sufficiently faster than TP weakening to make $\delta_{drop}^{(c)}$ a reasonable approximation.

For parameter combinations leading to stronger TP, the TP weakening length scale can be comparable to, or less than, the rate-and-state evolution distance D . For example, for $\Lambda = 1$ MPa/K, $\alpha_{hy} = 10^{-5}$ m²/s, and assuming $v_{seis} = 6$ m/s, $L^* \approx 4$ mm. When the TP and rate-and-state length scales are comparable, we can no longer assume that all rate-and-state friction evolution occurs before TP becomes relevant. The rate-and-state direct effect would therefore cause the TP weakening to begin from a higher stress, with a higher friction coefficient, than we assume in applying the analytical solution. TP with such a small slip-weakening scale as to be comparable to the rate-and-state evolution scales—much smaller than the incoming seismic slip—also implies very efficient TP, and all such fault models result in rupture propagation through the entire creeping section. The analytically inferred δ_{drop}^* would also be quite small, and hence the analytical solution would still predict propagation through the creeping region. Hence, this potential discrepancy

between the actual and assumed friction coefficient for TP evolution does not have a significant impact on the results presented here.

4.6 Conclusions

We have investigated whether dynamic ruptures can propagate through the creeping section of the San Andreas Fault due to co-seismic thermal pressurization (TP) of pore fluids using a highly simplified 1-D fault model embedded in a 2-D elastodynamic medium. Our results show that, within the highly simplified framework of our model, a wide range of TP parameters cause a M_w 6-style event to rapidly arrest in the creeping section, but a M_w 7-style event to propagate partially or totally through.

Analytical solutions for TP show that the point at which events begin to substantially rupture the creeping section in our simulations is well explained by the requirement that the seismic slip in the incoming dynamic event is greater than the analytical estimates of slip needed for a positive stress drop due to TP in the otherwise velocity-strengthening creeping region. The full rupture of the creeping section always occurs in our simulations when the incoming seismic slip is 10 or more times larger than the analytical estimate of the slip needed for a positive stress drop. Our results suggest an approach for considering the potential effect of different model parameters, including TP parameters, and different incoming seismic ruptures, without dynamic modeling, by comparing potential slip during an event (δ_{event}) to an analytically estimated slip required for a positive stress drop (δ_{drop}). When applying the analytical solutions, it is important to consider the width of the diffusive zones in relation to the width of the shearing layer, to determine whether the slip-on-a-plane regime—in which heat and fluid diffusion are essential and accounted for—or the adiabatic, undrained regime is a good approximation.

These results suggest a mechanism for the creeping section of the San Andreas Fault to rupture during a major earthquake, offering a possible explanation for the inferred ruptures in the creeping section presented by Coffey et al. (2022). Furthermore, we identify models with physically plausible parameters that have temperature increases in the range suggested by Coffey et al. (2022) and find that such ruptures can accumulate even larger seismic slip in the creeping section than estimated by Coffey et al. (2022). The parameters of such models include conservative choices for the efficiency of TP, such as low rate-and-state friction coefficients with the reference friction value of $f_* = 0.15$ (and the resulting “creeping” shear stress of

15 MPa; higher values would promote TP), time-independent TP parameters more appropriate for damaged rocks, such as hydraulic diffusivity of 10^{-4} to 10^{-5} m²/s (using initial values for undamaged rocks may promote TP), and time-independent shear-layer width of 10 mm (whereas shear localization to narrower zones with slip may further promote TP). At the same time, our modeling does not include several ingredients that can reduce the efficiency of TP, including inelastic shear layer dilatancy and complex fault structure with multiple fault strands in the creeping region, which need to be investigated in future work.

Our work highlights the importance of further constraining fault physics through lab experiments coupled with computational and theoretical work, especially the evolution of hydraulic diffusivity, dilatancy, and shear-layer structure/width with ongoing seismic slip. Improved modeling of this problem would include expanding these simulations to 3-D in order to incorporate depth-dependent properties; incorporating dilatancy and evolution of TP parameters during the dynamic rupture; studying more realistic fault geometry; and including the possibility and effects of melting and other dynamic weakening mechanisms.

Table 4.1: Summary of simulation parameters. Parameters that we vary in this study are indicated in **bold** at the bottom, along with their ranges. DW: dynamic weakening. VW: velocity-weakening. VS: velocity-strengthening.

Parameter	Symbol	Value
Fault length	λ	200 km
Frictional fault length	λ_{fric}	180 km
Rate-and-state VW section length for M_w 6-style events	λ_{vw}	25 km
Crustal plane width for M_w 6-style events	H_{seis}	2 km
Rate-and-state VW section length for M_w 7-style events	λ_{vw}	80 km
Crustal plane width for M_w 7-style events	H_{seis}	8 km
Cell size	Δx	5 m
Plate loading rate	v_{pl}	23 mm/yr
Slip velocity limit	v_{lim}	15 m/s
P wave speed	c_p	5.2 km/s
S wave speed	c_s	3.0 km/s
Shear modulus	G	30 GPa
Poisson ratio	ν	0.25
Initial effective normal stress	$\sigma - p_0$	120 MPa
Rate-and-state reference slip velocity	V^*	10^{-6} m/s
Rate-and-state critical slip distance	D_{RS}	4 mm
Rate-and-state properties in VW region	a	0.005
	b	0.007
	f^*	0.3
Rate-and-state properties in VS regions	a	0.01
	b	0.006
	f^*	0.15
Specific heat	ρc	2.7 MPa/K
Thermal diffusivity in DW region	α_{th}	10^{-6} m ² /s
Hydraulic diffusivity in DW region	α_{hyd}	$10^{-5} - 10^{-2}$ m ² /s
Undrained dp/dT in DW region	Λ	0.01-1.0 MPa/K
Half-width of shear zone in DW region	w	0.01-100 mm

CONCLUSIONS AND OUTLOOK

This thesis has illustrated how we can draw on a diverse range of techniques to study the earthquake cycle over multiple temporal and spatial scales. We combined large InSAR data sets with modern processing techniques to study both the deformation of the Earth and the damage that earthquakes do over wide areas. Additionally, we used high-performance computational simulations of earthquakes, and analytical models, to explore how dynamic weakening via thermal pressurization can allow creeping sections of faults to rapidly rupture in earthquakes. Below, we review some suggestions for further work in each project, before providing some closing thoughts.

Our work in Chapter 1 showed how InSAR measurements can be corrected for the signal of plate motion. Until now, it has been common practice to either fit and remove the long-wavelength component of InSAR data, or use GNSS measurements to constrain the long-wavelength (e.g., Jolivet et al. (2015) and Weiss et al. (2020)). With the plate motion correction, combined with previously developed correction methods, it will be useful to explore how well InSAR can constrain very long-wavelength deformation without the aid of GNSS in a variety of different settings. The troposphere will likely be the dominant source of noise for time series spanning a few years (Parizzi et al., 2021), but with data spanning a decade or more the reduction in noise may be sufficient to recover long-wavelength secular signals below the level of 1 mm/yr, allowing greatly improved resolution in calculations of fault coupling.

The sensitivity to InSAR of plate motion also suggests that InSAR may be able to supplement GNSS data when constraining plate motion models, particularly in regions where GNSS is sparse. Given the small size of the plate motion signal, this would likely require long time series of InSAR data in order to get measurements of sufficient quality.

Using plate motion models to adjust InSAR data is roughly equivalent to expressing the InSAR measurements in a reference frame that is moving with the relevant plate. This correction may be less useful in zones of distributed deformation, for example, at plate motion boundaries, and how best to adjust InSAR measurements in this context, particularly without dense GNSS, is another important question that requires further study.

The results we presented for the Makran subduction zone in Chapter 2 show that, with plate motion corrections, as well as other adjustments, particularly for the ionosphere and troposphere, we are able to resolve tectonic deformation over a wide area with InSAR alone. While we were able to make broad statements about the degree of coupling on the megathrust, before such data can be fully interpreted we need more rigorous estimates of the noise levels (predominantly from residual troposphere) for our velocity measurements. Methods for estimating the noise in InSAR time series have already been developed (e.g., Agram and Simons, 2015; Lohman and Simons, 2005; Parizzi et al., 2021) and could be straightforwardly applied to our data, although we may need to take account of spatial variations in the noise. We also need to quantify the uncertainties introduced when we merge adjacent tracks in order to understand how well our merged velocity field is able to measure deformation spanning several tracks. It would also be sensible to experiment with more sophisticated approaches to merging tracks, for example incorporating data uncertainties and masking areas of known deformation when merging.

Our model of the subduction zone could also be improved, for example by using a layered elastic structure and updated fault geometry. There is some suggestion in the comparison between ascending and descending InSAR tracks that the convergence direction varies along the fault, with some amount of westward motion at the western end of the subduction zone, although this cannot be claimed with any certainty given our noise levels. Incorporating a variable convergence direction into our model might therefore be necessary. With the combination of an improved model and better estimates of the noise, it should be possible to invert the surface velocities for coupling on the megathrust. Using a Bayesian approach would allow a probabilistic coupling map to be derived (e.g., Lin et al. (2015)), revealing the impact of data and model uncertainties on the final coupling map, and the resulting uncertainties in the estimated size of potential future earthquakes.

The presence of significant post-seismic deformation at the eastern end of the subduction zone will likely make it challenging to invert for coupling in that region. Recent work has suggested that post-seismic deformation from the Balochistan earthquake is driven by creep on the megathrust (Lv & Shao, 2022), suggesting that the fault is fully decoupled north of 220 km from the deformation front of the accretionary prism. The eastern end is the only part of the megathrust to have hosted large earthquakes in historic time, making understanding coupling there an important issue. It may be possible to simultaneously invert for post-seismic deformation

from the Khash and Balochistan earthquakes, as well as the coupling signal on the megathrust, possibly making use of the different temporal characteristics of post-seismic deformation and interseismic coupling, and the multiple different look directions provided by overlapping InSAR tracks. At the very least, a probabilistic inversion may be able to show us that the coupling cannot be resolved given the currently available data and ongoing post-seismic deformation.

Beyond our specific work, the clear hazard presented by the Makran subduction zone necessitates greater study from a variety of different angles. More geodetic work, ideally including the deployment of a dense network of GNSS stations on shore, and potentially ocean bottom geodesy, would allow for greater resolution of the megathrust coupling in space and time. Further work on measuring paleotsunamis is also necessary to piece together the seismic history of the western end of the subduction zone.

The damage mapping work of Chapter 3 showed how we could exploit part of the InSAR signal generally regarded as noise to rapidly map damage over wide areas. In our work, we quantify the quality of our damage classification using building damage data from the 2016 Amatrice earthquake. For this data set, every building in the town was inspected, and its level of damage recorded, allowing us to compute the precision and recall of our damage classifier. However, the data is limited to a small geographic region, and we cannot be sure if our method will maintain similar performance in other areas of the world. To ensure the reliability of this method, and test potential improvements, future work should focus on the development of high quality benchmarks—a global data set of natural disasters that are well covered by SAR data, and also have comprehensive damage assessments against which the classifier can be tested. Such benchmarks would allow our method to be tested across a wide range of conditions, providing more assurance that the method will perform well on future disasters. Organizations such as the Copernicus Emergency Management Service (<https://emergency.copernicus.eu/>), the Geotechnical Extreme Events Reconnaissance Association (<https://geerassociation.org/>), and the Humanitarian OpenStreetMap Team (<https://www.hotosm.org/>) may be useful places to look for damage mapping products and expertise.

There are several potential methodological developments that could also be explored, both in the input data, and damage classification algorithm, all of which could be tested against the benchmarks described above. The goal of the damage mapping algorithm is to get the best estimate of the probability distribution of the coherence

under normal conditions, so that anomalies can be detected. To this end, the InSAR data could be augmented with spatial and temporal baselines, as well as potentially precipitation records, as these variables will all affect coherence, and could allow the deep learning algorithm to improve its forecast. Adding extra connections to the interferometric network, so that we have more than just sequential coherence images, may also be useful. Ultimately the best option may be to work directly with the amplitude and phase of the SLCs, rather than computing coherence.

On the algorithm side, it will be important to take advantage of the rapid progress in machine learning. The current method only makes use of individual time series when forecasting, meaning that potentially useful spatial correlations between adjacent time series are ignored. Architectures such as convolutional LSTMs allow spatial and temporal information to be incorporated when making forecasts (Shi et al., 2015); however, more recent developments in transformer architectures offer advantages over LSTM-based approaches (Vaswani et al., 2017; Zhou et al., 2021).

Rapid calculation of damage maps will benefit from quick coregistration of SAR images and computation of coherence. Currently, this process can be slow, particularly for large volumes of data, both because of processing and download times. Future work could therefore also go towards making these stages faster by improving the code, and deploying the code on high performance computers that can process many SAR images in parallel.

Along with technical work, efforts should also be made to better understand the requirements of end users, and how these damage maps can be integrated into existing emergency response pipelines. The ability to deliver timely products, in an understandable format, to emergency managers with whom there is a pre-existing relationship, is vital if these damage maps are to have practical applications. Important steps could include conducting interviews with disaster response experts, testing different data visualization approaches, and putting products on an easily accessible platform. It would be sensible to build on the knowledge of disaster response organizations that bridge the gap between scientists and emergency managers, such as the USGS Emergency Response (<https://www.usgs.gov/emergency-operations-portal>) and NASA's Disasters program (<https://appliedsciences.nasa.gov/what-we-do/disasters>). With these connections made, much of the mapping process could be automated, and products directly supplied to first responders with minimal outside intervention. The increasing availability of SAR data will hopefully make this possible within hours in the coming years.

Our dynamic rupture simulations presented in Chapter 4 highlighted the potential role of thermal pressurization in allowing the creeping section of the San Andreas Fault (SAF) to rupture co-seismically. Such results have important implications for the hazard posed by the SAF, and so require further investigation. The immediate development of this project could consist of extending the models to 3-D, and incorporating a wider variety of physical processes in to the model. A 3-D model would allow M_w 6 and M_w 7-style events to be simulated together, although this would present substantial computational challenges. It may be necessary to reduce the computational burden, for example by substantially truncating the size of the creeping section and M_w 7 patch, in order to explore parameter combinations in a manageable amount of time. The use of 2-D models and analytical solutions will be important in narrowing the parameter space.

Direct experimental observations of thermal pressurization in laboratory settings remain limited, the parameters controlling its efficiency are highly uncertain, and the the situations under which other weakening mechanisms could dominate are unclear (e.g., Aben et al., 2020; Badt et al., 2020; Brantut and Mitchell, 2018). Modeling efforts rely on physical observations to determine the relevant processes, but can also inform future experiments by indicating which parameter uncertainties have the greatest influence on the model behavior. Our modeling work has highlighted several particularly important parameters, and future experimental work should be able to place better constraints on these.

The coming decade will see a continual increase in the availability of data and computing power, promising exciting new results, but also many potential pitfalls. In 1810, German polymath Johann Wolfgang von Goethe stated:

“The modern age has a false sense of superiority, because of the great mass of data at its disposal. But the valid criterion of distinction is rather the extent to which [humanity] knows how to form and master the material at [its] command.”¹

Goethe’s words are a reminder that it is not just humanity’s ability to collect and process large volumes of data that pushes forward our understanding of the world,

¹This quote sits on my advisor’s wall, and is generally attributed to Goethe’s “Theory of Colors,” published in 1810 in German, with an English translation appearing in 1840. I have been unable to find a direct version of this quote in the 1840 translation, but the quote is attributed to Goethe by the 1949 book “Goethe: Wisdom and Experience.” I therefore include the attribution with some caution. As Abraham Lincoln famously said: “Remember that not everything you read on the internet is true.”

but also our ability to intelligently “form and master” the material at our command, combining insights from the smallest to the largest scales to tell a coherent story about our Earth.

The scientific advances to come promise profound benefits to humanity, but it is ultimately a question of politics and policy that determines how they will be used. In this thesis we have examined natural processes that can lead to enormous destruction, but there are no natural disasters. Who lives and who dies, which buildings collapse and which remain standing—these are questions that depend on the decisions made by those with power. These decisions must be informed by the best science, *and* made for the benefit of everyone. It is to this endeavor that I now turn my attention.

BIBLIOGRAPHY

- 3Blue1Brown. (2017). But what is a Neural Network? Deep learning, Chapter 1 - YouTube [Accessed June 22, 2020]. Retrieved June 23, 2020, from https://www.youtube.com/watch?v=aircAruvnKk&list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi
- Aben, F. M., Doan, M. L., & Mitchell, T. M. (2020). Variation of Hydraulic Properties Due to Dynamic Fracture Damage: Implications for Fault Zones. *Journal of Geophysical Research: Solid Earth*, 125(4), 1–22. <https://doi.org/10.1029/2019JB018919>
- Agram, P. S., Jolivet, R., Riel, B., Lin, Y. N., Simons, M., Hetland, E., Doin, M.-P., & Lasserre, C. (2013). New Radar Interferometric Time Series Analysis Toolbox Released. *Eos, Transactions American Geophysical Union*, 94(7), 69–70. <https://doi.org/10.1002/2013EO070001>
- Agram, P. S., & Simons, M. (2015). A noise model for InSAR time series. *Journal of Geophysical Research: Solid Earth*, 120(4), 2752–2771. <https://doi.org/10.1002/2014JB011271>
- Altamimi, Z., Métivier, L., & Collilieux, X. (2012). ITRF2008 plate motion model. *Journal of Geophysical Research: Solid Earth*, 117(7), 1–14. <https://doi.org/10.1029/2011JB008930>
- Altamimi, Z., Métivier, L., Rebischung, P., Rouby, H., & Collilieux, X. (2017). ITRF2014 plate motion model. *Geophysical Journal International*, 209(3), 1906–1912. <https://doi.org/10.1093/gji/ggx136>
- Altamimi, Z., Rebischung, P., Métivier, L., & Collilieux, X. (2016). ITRF2014: A new release of the International Terrestrial Reference Frame modeling nonlinear station motions. *Journal of Geophysical Research: Solid Earth*, 121(8), 6109–6131. <https://doi.org/10.1002/2016JB013098>
- Amelung, F., Galloway, D. L., Bell, J. W., Zebker, H. A., & Lacznik, R. J. (1999). Sensing the ups and downs of Las Vegas: InSAR reveals structural control of land subsidence and aquifer-system deformation. *Geology*, 27(6), 483. [https://doi.org/10.1130/0091-7613\(1999\)027<0483:STUADO>2.3.CO;2](https://doi.org/10.1130/0091-7613(1999)027<0483:STUADO>2.3.CO;2)
- Anantrasirichai, N., Biggs, J., Albino, F., & Bull, D. (2019). A deep learning approach to detecting volcano deformation from satellite imagery using synthetic datasets. *Remote Sensing of Environment*, 230, 111179. <https://doi.org/10.1016/j.rse.2019.04.032>
- Andrews, D. J. (2005). Rupture dynamics with energy loss outside the slip zone. *Journal of Geophysical Research: Solid Earth*, 110(1), 1–14. <https://doi.org/10.1029/2004JB003191>

- Ansari, H., De Zan, F., & Parizzi, A. (2021). Study of Systematic Bias in Measuring Surface Deformation with SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 59(2), 1285–1301. <https://doi.org/10.1109/TGRS.2020.3003421>
- Argus, D. F., Gordon, R. G., Heflin, M. B., Ma, C., Eanes, R. J., Willis, P., Peltier, W. R., & Owen, S. E. (2010). The angular velocities of the plates and the velocity of Earth's centre from space geodesy. *Geophysical Journal International*, 180(3), 913–960. <https://doi.org/10.1111/j.1365-246X.2009.04463.x>
- Avouac, J.-P. (2015). From Geodetic Imaging of Seismic and Aseismic Fault Slip to Dynamic Modeling of the Seismic Cycle. *Annual Review of Earth and Planetary Sciences*, 43(1), 233–271. <https://doi.org/10.1146/annurev-earth-060614-105302>
- Avouac, J.-P., Ayoub, F., Wei, S., Ampuero, J.-P., Meng, L., Leprince, S., Jolivet, R., Duputel, Z., & Helmberger, D. (2014). The 2013, Mw 7.7 Balochistan earthquake, energetic strike-slip reactivation of a thrust fault. *Earth and Planetary Science Letters*, 391, 128–134. <https://doi.org/10.1016/j.epsl.2014.01.036>
- Badt, N. Z., Tullis, T. E., Hirth, G., & Goldsby, D. L. (2020). Thermal Pressurization Weakening in Laboratory Experiments. *Journal of Geophysical Research: Solid Earth*, 125(5), 1–21. <https://doi.org/10.1029/2019JB018872>
- Bähr, H. (2013). *Orbital effects in spaceborne synthetic aperture radar interferometry* (Doctoral dissertation). <https://doi.org/10.5445/KSP/1000037166>
- Bähr, H., Samiei-Esfahany, S., & Hanssen, R. F. (2012). On The Effect Of Reference Frame Motion On InSAR Deformation Estimates. *Proceedings of Fringe 2011*.
- Bakun, W. H., Aagaard, B., Dost, B., Ellsworth, W. L., Hardebeck, J. L., Harris, R. A., Ji, C., Johnston, M. J. S., Langbein, J., Lienkaemper, J. J., Michael, A. J., Murray, J. R., Nadeau, R. M., Reasenber, P. A., Reichle, M. S., Roeloffs, E. A., Shakal, A., Simpson, R. W., & Waldhauser, F. (2005). Implications for prediction and hazard assessment from the 2004 Parkfield earthquake. *Nature*, 437(7061), 969–974. <https://doi.org/10.1038/nature04067>
- Bakun, W. H., & Lindh, A. G. (1985). The Parkfield, California, Earthquake Prediction Experiment. *Science*, 229(4714), 619–624. <https://doi.org/10.1126/science.229.4714.619>
- Bamler, R., & Hartl, P. (1998). Synthetic aperture radar interferometry. *Inverse Problems*, 14(4), R1–R54. <https://doi.org/10.1088/0266-5611/14/4/001>
- Barbot, S., Fialko, Y., & Bock, Y. (2009). Postseismic deformation due to the Mw 6.0 2004 Parkfield earthquake: Stress-driven creep on a fault with spatially variable rate-and-state friction parameters. *Journal of Geophysical Research: Solid Earth*, 114(7), 1–26. <https://doi.org/10.1029/2008JB005748>

- Barbot, S., Lapusta, N., & Avouac, J.-p. (2012). Under the hood of the earthquake machine: toward predictive modeling of the seismic cycle. *Science*, 336(6082), 707–10. <https://doi.org/10.1126/science.1218796>
- Barnhart, W. D., Hayes, G. P., Briggs, R. W., Gold, R. D., & Bilham, R. (2014). Ball-and-socket tectonic rotation during the 2013 Mw7.7 Balochistan earthquake. *Earth and Planetary Science Letters*, 403, 210–216. <https://doi.org/10.1016/j.epsl.2014.07.001>
- Barnhart, W. D., Hayes, G. P., Samsonov, S. V., Fielding, E. J., & Seidman, L. E. (2014). Breaking the oceanic lithosphere of a subducting slab: The 2013 Khash, Iran earthquake. *Geophysical Research Letters*, 41(1), 32–36. <https://doi.org/10.1002/2013GL058096>
- Bekaert, D., Walters, R., Wright, T., Hooper, A., & Parker, D. (2015). Statistical comparison of InSAR tropospheric correction techniques. *Remote Sensing of Environment*, 170, 40–47. <https://doi.org/10.1016/j.rse.2015.08.035>
- Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2), 157–166. <https://doi.org/10.1109/72.279181>
- Berardino, P., Fornaro, G., Lanari, R., & Sansosti, E. (2002). A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms. *IEEE Transactions on Geoscience and Remote Sensing*, 40(11), 2375–2383. <https://doi.org/10.1109/TGRS.2002.803792>
- Berrar, D., & Flach, P. (2012). Caveats and pitfalls of ROC analysis in clinical microarray research (and how to avoid them). *Briefings in Bioinformatics*, 13(1), 83–97. <https://doi.org/10.1093/bib/bbr008>
- Biggs, J., Wright, T., Lu, Z., & Parsons, B. (2007). Multi-interferogram method for measuring interseismic deformation: Denali Fault, Alaska. *Geophysical Journal International*, 170(3), 1165–1179. <https://doi.org/10.1111/j.1365-246X.2007.03415.x>
- Bizzarri, A., & Cocco, M. (2003). Slip-weakening behavior during the propagation of dynamic ruptures obeying rate- and state-dependent friction laws. *Journal of Geophysical Research*, 108(B8), 2373. <https://doi.org/10.1029/2002jb002198>
- Blewitt, G. (2015). *GPS and Space-Based Geodetic Methods* (Vol. 3). Elsevier B.V. <https://doi.org/10.1016/B978-0-444-53802-4.00060-9>
- Bouaraba, A., Younsi, A., Belhadj Aissa, A., Acheroy, M., Milisavljevic, N., & Closson, D. (2012). Robust Techniques for Coherent Change Detection Using Cosmo-SkyMed Sar Images. *Progress In Electromagnetics Research M*, 22(November 2011), 219–232. <https://doi.org/10.2528/PIERM11110707>

- Bradbury, K. K., Evans, J. P., Chester, J. S., Chester, F. M., & Kirschner, D. L. (2011). Lithology and internal structure of the San Andreas fault at depth based on characterization of Phase 3 whole-rock core in the San Andreas Fault Observatory at Depth (SAFOD) borehole. *Earth and Planetary Science Letters*, 310(1-2), 131–144. <https://doi.org/10.1016/j.epsl.2011.07.020>
- Brandenberg, S. J., Wang, P., Nweke, C. C., Hudson, K., Mazzoni, S., Bozorgnia, Y., Hudnut, K. W., Davis, C. A., Ahdi, S. K., Zareian, F., Fayaz, J., Koehler, R. D., Chupik, C., Pierce, I., Williams, A., Akciz, S., Hudson, M. B., Kishida, T., Brooks, B. A., . . . Conway, E. (2019). *Preliminary report on engineering and geological effects of the July 2019 Ridgecrest earthquake sequence* (J. P. Stewart, Ed.; Report). <https://doi.org/10.18118/G6H66K>
- Brantut, N. (2020). Dilatancy-induced fluid pressure drop during dynamic rupture: Direct experimental evidence and consequences for earthquake dynamics. *Earth and Planetary Science Letters*, 538, 116179. <https://doi.org/10.1016/j.epsl.2020.116179>
- Brantut, N., & Mitchell, T. M. (2018). Assessing the Efficiency of Thermal Pressurization Using Natural Pseudotachylite-Bearing Rocks. *Geophysical Research Letters*, 45(18), 9533–9541. <https://doi.org/10.1029/2018GL078649>
- Brantut, N., & Platt, J. D. (2017). Dynamic Weakening and the Depth Dependence of Earthquake Faulting. *Geophysical monograph series* (pp. 171–194). <https://doi.org/10.1002/9781119156895.ch9>
- Brunner, D., Lemoine, G., & Bruzzone, L. (2010). Earthquake damage assessment of buildings using VHR optical and SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 48(5), 2403–2420. <https://doi.org/10.1109/TGRS.2009.2038274>
- Burgmann, R. (2000). Earthquake Potential Along the Northern Hayward Fault, California. *Science*, 289(5482), 1178–1182. <https://doi.org/10.1126/science.289.5482.1178>
- Byrne, D. E., Sykes, L. R., & Davis, D. M. (1992). Great thrust earthquakes and aseismic slip along the plate boundary of the Makran Subduction Zone. *Journal of Geophysical Research*, 97(B1), 449. <https://doi.org/10.1029/91JB02165>
- Cao, Y., Jónsson, S., & Li, Z. (2021). Advanced InSAR Tropospheric Corrections From Global Atmospheric Models that Incorporate Spatial Stochastic Properties of the Troposphere. *Journal of Geophysical Research: Solid Earth*, 126(5), 1–20. <https://doi.org/10.1029/2020JB020952>
- Chang, S. H., Avouac, J. P., Barbot, S., & Lee, J. C. (2013). Spatially variable fault friction derived from dynamic modeling of aseismic afterslip due to the 2004 Parkfield earthquake. *Journal of Geophysical Research: Solid Earth*, 118(7), 3431–3447. <https://doi.org/10.1002/jgrb.50231>

- Chaussard, E., Johnson, C. W., Fattahi, H., & Bürgmann, R. (2016). Potential and limits of InSAR to characterize interseismic deformation independently of GPS data: Application to the southern San Andreas Fault system. *Geochemistry, Geophysics, Geosystems*, 17(3), 1214–1229. <https://doi.org/10.1002/2015GC006246>
- Chen, C. W., & Zebker, H. A. (2002). Phase unwrapping for large SAR interferograms: Statistical segmentation and generalized network models. *IEEE Transactions on Geoscience and Remote Sensing*, 40(8), 1709–1719. <https://doi.org/10.1109/TGRS.2002.802453>
- Chester, F. M., & Chester, J. S. (1998). Ultracataclasite structure and friction processes of the Punchbowl fault, San Andreas system, California. *Tectonophysics*, 295(1-2), 199–221. [https://doi.org/10.1016/S0040-1951\(98\)00121-8](https://doi.org/10.1016/S0040-1951(98)00121-8)
- Cho, K., van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the Properties of Neural Machine Translation: Encoder-Decoder Approaches [arXiv: 1409.1259]. *arXiv:1409.1259 [cs, stat]*. Retrieved March 20, 2019, from <http://arxiv.org/abs/1409.1259>
- Coffey, G. L., Savage, H. M., Polissar, P. J., Cox, S. E., Hemming, S. R., Winckler, G., & Bradbury, K. K. (2022). History of earthquakes along the creeping section of the San Andreas fault, California, USA. *Geology*, 50(4), 516–521. <https://doi.org/10.1130/G49451.1>
- Cong, X., Balss, U., Eineder, M., & Fritz, T. (2012). Imaging Geodesy—Centimeter-Level Ranging Accuracy With TerraSAR-X: An Update. *IEEE Geoscience and Remote Sensing Letters*, 9(5), 948–952. <https://doi.org/10.1109/LGRS.2012.2187042>
- Copernicus Emergency Management Service. (2016). EMSR177: Earthquake in Central Italy [Accessed June 22, 2020]. Retrieved June 23, 2020, from <https://emergency.copernicus.eu/mapping/list-of-components/EMSR177>
- Cubas, N., Lapusta, N., Avouac, J. P., & Perfettini, H. (2015). Numerical modeling of long-term earthquake sequences on the NE Japan megathrust: Comparison with observations and implications for fault friction. *Earth and Planetary Science Letters*, 419, 187–198. <https://doi.org/10.1016/j.epsl.2015.03.002>
- Dalaison, M., Jolivet, R., van Rijsingen, E. M., & Michel, S. (2021). The Interplay Between Seismic and Aseismic Slip Along the Chaman Fault Illuminated by InSAR. *Journal of Geophysical Research: Solid Earth*, 126(12), 1–23. <https://doi.org/10.1029/2021JB021935>
- Dalla Mura, M., Benediktsson, J., Bovolo, F., & Bruzzone, L. (2008). An Unsupervised Technique Based on Morphological Filters for Change Detection in Very High Resolution Images. *IEEE Geoscience and Remote Sensing Letters*, 5(3), 433–437. <https://doi.org/10.1109/LGRS.2008.917726>

- Davis, J., & Goadrich, M. (2006). The relationship between Precision-Recall and ROC curves. *Proceedings of the 23rd international conference on Machine learning - ICML '06*, 233–240. <https://doi.org/10.1145/1143844.1143874>
- De Paola, N., Collettini, C., Faulkner, D. R., & Trippetta, F. (2008). Fault zone architecture and deformation processes within evaporitic rocks in the upper crust. *Tectonics*, 27(4), 1–21. <https://doi.org/10.1029/2007TC002230>
- De Zan, F., Prats-Iraola, P., Scheiber, R., & Rucci, A. (2014). Interferometry with TOPS: Coregistration and azimuth shifts. *Proceedings of the European Conference on Synthetic Aperture Radar, EUSAR, Proceeding(0)*, 949–952.
- DeMets, C., Gordon, R. G., & Argus, D. F. (2010). Geologically current plate motions. *Geophysical Journal International*, 181(1), 1–80. <https://doi.org/10.1111/j.1365-246X.2009.04491.x>
- Di Toro, G., Han, R., Hirose, T., De Paola, N., Nielsen, S., Mizoguchi, K., Ferri, F., Cocco, M., & Shimamoto, T. (2011). Fault lubrication during earthquakes. *Nature*, 471(7339), 494–498. <https://doi.org/10.1038/nature09838>
- Dicaprio, C. J., Simons, M., Kenner, S. J., & Williams, C. A. (2008). Post-seismic reloading and temporal clustering on a single fault. *Geophysical Journal International*, 172(2), 581–592. <https://doi.org/10.1111/j.1365-246X.2007.03622.x>
- Dieterich, J. H. (1978). Preseismic fault slip and earthquake prediction. *Journal of Geophysical Research*, 83(B8), 3940. <https://doi.org/10.1029/jb083ib08p03940>
- Doin, M. P., Lasserre, C., Peltzer, G., Cavalié, O., & Doubre, C. (2009). Corrections of stratified tropospheric delays in SAR interferometry: Validation with global atmospheric models. *Journal of Applied Geophysics*, 69(1), 35–50. <https://doi.org/10.1016/j.jappgeo.2009.03.010>
- Dziewonski, A. M., Chou, T. A., & Woodhouse, J. H. (1981). Determination of earthquake source parameters from waveform data for studies of global and regional seismicity. *Journal of Geophysical Research*, 86(B4), 2825–2852. <https://doi.org/10.1029/JB086iB04p02825>
- Eineder, M., Minet, C., Steigenberger, P., Cong, X., & Fritz, T. (2011). Imaging geodesy - Toward centimeter-level ranging accuracy with TerraSAR-X. *IEEE Transactions on Geoscience and Remote Sensing*, 49(2), 661–671. <https://doi.org/10.1109/TGRS.2010.2060264>
- Ekström, G., Nettles, M., & Dziewoński, A. M. (2012). The global CMT project 2004–2010: Centroid-moment tensors for 13,017 earthquakes. *Physics of the Earth and Planetary Interiors*, 200–201, 1–9. <https://doi.org/10.1016/j.pepi.2012.04.002>

- Emardson, T. R., Simons, M., & Webb, F. H. (2003). Neutral atmospheric delay in interferometric synthetic aperture radar applications: Statistical description and mitigation. *Journal of Geophysical Research: Solid Earth*, 108(B5), 1–8. <https://doi.org/10.1029/2002JB001781>
- Endo, Y., Adriano, B., Mas, E., & Koshimura, S. (2018). New Insights into Multiclass Damage Classification of Tsunami-Induced Building Damage from SAR Images. *Remote Sensing*, 10(12), 2059–2059. <https://doi.org/10.3390/rs10122059>
- Erickson, B. A., Jiang, J., Barall, M., Lapusta, N., Dunham, E. M., Harris, R., Abrahams, L. S., Allison, K. L., Ampuero, J. P., Barbot, S., Cattania, C., Elbanna, A., Fialko, Y., Idini, B., Kozdon, J. E., Lambert, V., Liu, Y., Luo, Y., Ma, X., . . . Wei, M. (2020). The community code verification exercise for Simulating Sequences of Earthquakes and Aseismic Slip (SEAS). *Seismological Research Letters*, 91(2A), 874–890. <https://doi.org/10.1785/0220190248>
- European Space Agency. (n.d.). Single Look Complex-Sentinel-1 SAR Technical Guide-Sentinel Online [Accessed August 31, 2020]. Retrieved August 31, 2020, from <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-1-sar/products-algorithms/level-1-algorithms/single-look-complex>
- Fattahi, H., Agram, P., & Simons, M. (2017). A Network-Based Enhanced Spectral Diversity Approach for TOPS Time-Series Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 55(2), 777–786. <https://doi.org/10.1109/TGRS.2016.2614925>
- Fattahi, H., & Amelung, F. (2013). DEM Error Correction in InSAR Time Series. *Geoscience and Remote Sensing, IEEE Transactions on*, 51(7), 4249–4259. <https://doi.org/10.1109/TGRS.2012.2227761>
- Fattahi, H., & Amelung, F. (2014). InSAR uncertainty due to orbital errors. *Geophysical Journal International*, 199(1), 549–560. <https://doi.org/10.1093/gji/ggu276>
- Fattahi, H., & Amelung, F. (2015). InSAR bias and uncertainty due to the systematic and stochastic tropospheric delay. *Journal of Geophysical Research: Solid Earth*, 120(12), 8758–8773. <https://doi.org/10.1002/2015JB012419>
- Fattahi, H., & Amelung, F. (2016). InSAR observations of strain accumulation and fault creep along the Chaman Fault system, Pakistan and Afghanistan. *Geophysical Research Letters*, 43(16), 8399–8406. <https://doi.org/10.1002/2016GL070121>
- Fattahi, H., Simons, M., & Agram, P. (2017). InSAR Time-Series Estimation of the Ionospheric Phase Delay: An Extension of the Split Range-Spectrum Technique. *IEEE Transactions on Geoscience and Remote Sensing*, 55(10), 5984–5996. <https://doi.org/10.1109/TGRS.2017.2718566>

- Faulkner, D. R., Mitchell, T. M., Behnsen, J., Hirose, T., & Shimamoto, T. (2011). Stuck in the mud? Earthquake nucleation and propagation through accretionary forearcs. *Geophysical Research Letters*, 38(18), 1–5. <https://doi.org/10.1029/2011GL048552>
- Ferretti, A., Fumagalli, A., Novali, F., Prati, C., Rocca, F., & Rucci, A. (2011). A New Algorithm for Processing Interferometric Data-Stacks: SqueeSAR. *IEEE Transactions on Geoscience and Remote Sensing*, 49(9), 3460–3470. <https://doi.org/10.1109/TGRS.2011.2124465>
- Fialko, Y. (2006). Interseismic strain accumulation and the earthquake potential on the southern San Andreas fault system. *Nature*, 441(7096), 968–971. <https://doi.org/10.1038/nature04797>
- Fialko, Y., Simons, M., & Agnew, D. (2001). The complete (3-D) surface displacement field in the epicentral area of the 1999 Mw 7.1 Hector Mine earthquake, California, from space geodetic observations. *Geophysical Research Letters*, 28(16), 3063–3066. <https://doi.org/10.1029/2001GL013174>
- Fielding, E. J., Talebian, M., Rosen, P. A., Nazari, H., Jackson, J. A., Ghorashi, M., & Walker, R. (2005). Surface ruptures and building damage of the 2003 Bam, Iran, earthquake mapped by satellite synthetic aperture radar interferometric correlation. *Journal of Geophysical Research: Solid Earth*, 110(3), 1–15. <https://doi.org/10.1029/2004JB003299>
- Fournier, T., Pritchard, M. E., & Finnegan, N. (2011). Accounting for Atmospheric Delays in InSAR Data in a Search for Long-Wavelength Deformation in South America. *IEEE Transactions on Geoscience and Remote Sensing*, 49(10), 3856–3867. <https://doi.org/10.1109/TGRS.2011.2139217>
- French, M. E., Kitajima, H., Chester, J. S., Chester, F. M., & Hirose, T. (2014). Displacement and dynamic weakening processes in smectite-rich gouge from the Central Deforming Zone of the San Andreas Fault. *Journal of Geophysical Research: Solid Earth*, 119(3), 1777–1802. <https://doi.org/10.1002/2013JB010757>
- Frohling, E., & Szeliga, W. (2016). GPS constraints on interplate locking within the Makran subduction zone. *Geophysical Journal International*, 205(1), 67–76. <https://doi.org/10.1093/gji/ggw001>
- Fruehn, J., White, R. S., & Minshull, T. A. (1997). Internal deformation and compaction of the Makran accretionary wedge. *Terra Nova*, 9(3), 101–104. <https://doi.org/10.1046/j.1365-3121.1997.d01-13.x>
- Garagash, D. I. (2012). Seismic and aseismic slip pulses driven by thermal pressurization of pore fluid. *Journal of Geophysical Research: Solid Earth*, 117(4), 1–37. <https://doi.org/10.1029/2011JB008889>

- Geudtner, D., Winter, R., & Vachon, P. (1996). Flood monitoring using ERS-1 SAR interferometry coherence maps. *IGARSS '96. 1996 International Geoscience and Remote Sensing Symposium*, 2, 966–968. <https://doi.org/10.1109/IGARSS.1996.516536>
- Gomba, G., Parizzi, A., De Zan, F., Eineder, M., & Bamler, R. (2016). Toward operational compensation of ionospheric effects in SAR interferograms: The split-spectrum method. *IEEE Transactions on Geoscience and Remote Sensing*, 54(3), 1446–1461. <https://doi.org/10.1109/TGRS.2015.2481079>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning* [Google-Books-ID: omivDQAAQBAJ]. MIT Press.
- Haghipour, N., Burg, J. P., Kober, F., Zeilinger, G., Ivy-Ochs, S., Kubik, P. W., & Faridi, M. (2012). Rate of crustal shortening and non-Coulomb behaviour of an active accretionary wedge: The folded fluvial terraces in Makran (SE, Iran). *Earth and Planetary Science Letters*, 355–356, 187–198. <https://doi.org/10.1016/j.epsl.2012.09.001>
- Hanssen, R. F. (2001). *Radar Interferometry: Data Interpretation and Error Analysis* (Vol. 2). Springer Netherlands. <https://doi.org/10.1007/0-306-47633-9>
- Harris, R. A. (2017). Large earthquakes and creeping faults. *Reviews of Geophysics*, 55(1), 169–198. <https://doi.org/10.1002/2016RG000539>
- Hayes, G. P., Moore, G. L., Portner, D. E., Hearne, M., Flamme, H., Furtney, M., & Smoczyk, G. M. (2018). Slab2, a comprehensive subduction zone geometry model. *Science*, 362(6410), 58–61. <https://doi.org/10.1126/science.aat4723>
- Heaton, T. H. (1990). Evidence for and implications of self-healing pulses of slip in earthquake rupture. *Physics of the Earth and Planetary Interiors*, 64(1), 1–20. [https://doi.org/10.1016/0031-9201\(90\)90002-F](https://doi.org/10.1016/0031-9201(90)90002-F)
- Heermance, R., Shipton, Z. K., & Evans, J. P. (2003). Fault structure control on fault slip and ground motion during the 1999 rupture of the Chelungpu fault, Taiwan. *Bulletin of the Seismological Society of America*, 93(3), 1034–1050. <https://doi.org/10.1785/0120010230>
- Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., . . . Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, 146(730), 1999–2049. <https://doi.org/10.1002/qj.3803>
- Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>

- Hoffmann, G., Rupprechter, M., Balushi, N. A., Grützner, C., & Reicherter, K. (2013). The impact of the 1945 Makran tsunami along the coastlines of the Arabian Sea (Northern Indian Ocean) – a review. *Zeitschrift für Geomorphologie, Supplementary Issues*, 57(4), 257–277. <https://doi.org/10.1127/0372-8854/2013/S-00134>
- Hoffmann, G., Grützner, C., Schneider, B., Preusser, F., & Reicherter, K. (2020). Large Holocene tsunamis in the northern Arabian Sea. *Marine Geology*, 419(September 2019). <https://doi.org/10.1016/j.margeo.2019.106068>
- Hooper, A., Bekaert, D., Spaans, K., & Arikan, M. (2012). Recent advances in SAR interferometry time series analysis for measuring crustal deformation. *Tectonophysics*, 514-517, 1–13. <https://doi.org/10.1016/j.tecto.2011.10.013>
- Hough, S. E., Yun, S.-H., Jung, J., Thompson, E., Parker, G. A., & Stephenson, O. (2020). Near-field Ground Motions and Shaking from the 2019 Mw 7.1 Ridgecrest, California, Mainshock: Insights from Instrumental, Macro-seismic Intensity, and Remote-Sensing Data. *Bulletin of the Seismological Society of America*. <https://doi.org/10.1785/0120200045>
- Hsu, Y. J., Simons, M., Avouac, J. P., Galetka, J., Sieh, K., Chlieh, M., Natawidjaja, D., Prawirodirdjo, L., & Bock, Y. (2006). Frictional afterslip following the 2005 Nias-Simeulue earthquake, Sumatra. *Science*, 312(5782), 1921–1926. <https://doi.org/10.1126/science.1126960>
- Johanson, I. A., Fielding, E. J., Rolandone, F., & Bürgmann, R. (2006). Coseismic and Postseismic Slip of the 2004 Parkfield Earthquake from Space-Geodetic Data. *Bulletin of the Seismological Society of America*, 96(4B), 269–282. <https://doi.org/10.1785/0120050818>
- Johnson, K. M. (2013). Slip rates and off-fault deformation in Southern California inferred from GPS data and models. *Journal of Geophysical Research: Solid Earth*, 118(10), 5643–5664. <https://doi.org/10.1002/jgrb.50365>
- Johnston, P. J., Filmer, M. S., & Fuhrmann, T. (2021). Evaluation of methods for connecting InSAR to a terrestrial reference frame in the Latrobe Valley, Australia. *Journal of Geodesy*, 95(10). <https://doi.org/10.1007/s00190-021-01560-2>
- Jolivet, R., Duputel, Z., Riel, B., Simons, M., Rivera, L., Minson, S. E., Zhang, H., Aivazis, M. a. G., Ayoub, F., Leprince, S., Samsonov, S., Motagh, M., & Fielding, E. J. (2014). The 2013 Mw 7.7 Balochistan Earthquake: Seismic Potential of an Accretionary Wedge. *Bulletin of the Seismological Society of America*, 104(2), 1020–1030. <https://doi.org/10.1785/0120130313>
- Jolivet, R., Grandin, R., Lasserre, C., Doin, M. P., & Peltzer, G. (2011). Systematic InSAR tropospheric phase delay corrections from global meteorological reanalysis data. *Geophysical Research Letters*, 38(17), 1–6. <https://doi.org/10.1029/2011GL048757>

- Jolivet, R., Simons, M., Agram, P. S., Duputel, Z., & Shen, Z.-K. (2015). Aseismic slip and seismogenic coupling along the central San Andreas Fault. *Geophysical Research Letters*, 42(2), 297–306. <https://doi.org/10.1002/2014GL062222>
- Jolivet, R., Agram, P. S., Lin, N. Y., Simons, M., Doin, M.-p., Peltzer, G., & Li, Z. (2014). Improving InSAR geodesy using Global Atmospheric Models. *Journal of Geophysical Research: Solid Earth*, 119(3), 2324–2341. <https://doi.org/10.1002/2013JB010588>
- Jordan, T. E., Lohman, R. B., Tapia, L., Pfeiffer, M., Scott, C. P., Amundson, R., Godfrey, L., & Riquelme, R. (2020). Surface materials and landforms as controls on InSAR permanent and transient responses to precipitation events in a hyperarid desert, Chile. *Remote Sensing of Environment*, 237, 111544. <https://doi.org/10.1016/j.rse.2019.111544>
- Jung, J., Kim, D.-j., Lavalle, M., & Yun, S.-H. (2016). Coherent Change Detection Using InSAR Temporal Decorrelation Model: A Case Study for Volcanic Ash Detection. *IEEE Transactions on Geoscience and Remote Sensing*, 54(10), 5765–5775. <https://doi.org/10.1109/TGRS.2016.2572166>
- Jung, J., Yun, S. H., Kim, D. J., & Lavalle, M. (2018). Damage-Mapping Algorithm Based on Coherence Model Using Multitemporal Polarimetric-Interferometric SAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, 56(3), 1520–1532. <https://doi.org/10.1109/TGRS.2017.2764748>
- Jung, J., & Yun, S.-H. (2020). Evaluation of Coherent and Incoherent Landslide Detection Methods Based on Synthetic Aperture Radar for Rapid Response: A Case Study for the 2018 Hokkaido Landslides. *Remote Sensing*, 12(2), 265. <https://doi.org/10.3390/rs12020265>
- Kaneko, Y., & Lapusta, N. (2008). Variability of earthquake nucleation in continuum models of rate-and-state faults and implications for aftershock rates. *Journal of Geophysical Research: Solid Earth*, 113(12), 1–25. <https://doi.org/10.1029/2007JB005154>
- Kaneko, Y., Avouac, J. P., & Lapusta, N. (2010). Towards inferring earthquake patterns from geodetic observations of interseismic coupling. *Nature Geoscience*, 3(5), 363–369. <https://doi.org/10.1038/ngeo843>
- Karimzadeh, S., Matsuoka, M., Miyajima, M., Adriano, B., Fallahi, A., & Karashi, J. (2018). Sequential SAR coherence method for the monitoring of buildings in Sarpole-Zahab, Iran. *Remote Sensing*, 10(8), 1–17. <https://doi.org/10.3390/rs10081255>
- Kaufman, S., Rosset, S., & Perlich, C. (2011). Leakage in data mining: Formulation, detection, and avoidance. *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining - KDD '11*, 556. <https://doi.org/10.1145/2020408.2020496>

- Kendrick, K., Akciz, S., Angster, S., Avouac, J., Bachhuber, J., Bennett, S., Blake, K., Bork, S., Brooks, B., Burgess, P., et al. (2019). Geologic observations of surface fault rupture associated with the Ridgecrest M6.4 and M7. 1 earthquake sequence by the Ridgecrest Rupture Mapping Group. *Southern California Earthquake Center Annual Meeting Proceedings (SCEC, 2019), contribution, 9779*.
- Khaledzadeh, M., & Ghods, A. (2022). Estimation of size of megathrust zone in the Makran subduction system by thermal modelling. *Geophysical Journal International*, 228(3), 1530–1540. <https://doi.org/10.1093/gji/ggab417>
- Khorrami, F., Vernant, P., Masson, F., Nilfouroushan, F., Mousavi, Z., Nankali, H., Saadat, S. A., Walpersdorf, A., Hosseini, S., Tavakoli, P., Aghamohammadi, A., & Alijanzade, M. (2019). An up-to-date crustal deformation map of Iran using integrated campaign-mode and permanent GPS velocities. *Geophysical Journal International*, 217(2), 832–843. <https://doi.org/10.1093/gji/ggz045>
- Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization [arXiv: 1412.6980]. *arXiv:1412.6980 [cs]*. Retrieved March 20, 2019, from <http://arxiv.org/abs/1412.6980>
- Kong, Y.-L., Huang, Q., Wang, C., Chen, J., Chen, J., & He, D. (2018). Long Short-Term Memory Neural Networks for Online Disturbance Detection in Satellite Image Time Series. *Remote Sensing*, 10(3), 452. <https://doi.org/10.3390/rs10030452>
- Kuo, L.-w., Sone, H., Luzin, V., Yeh, E.-c., Hsu, Y.-j., & Tan, E. (2022). Gouge fabrics reset by thermal pressurization record stress on faults after earthquakes. *Geology*, 20(20), 1–5. <https://doi.org/10.1130/G50217.1>
- Lachenbruch, A. H. (1980). Frictional heating, fluid pressure, and the resistance to fault motion. *Journal of Geophysical Research*, 85(B11), 6097. <https://doi.org/10.1029/JB085iB11p06097>
- Lachenbruch, A. H., & Sass, J. H. (1980). Heat flow and energetics of the San Andreas Fault Zone. *Journal of Geophysical Research: Solid Earth*, 85(B11), 6185–6222. <https://doi.org/10.1029/JB085iB11p06185>
- Lambert, V., Lapusta, N., & Perry, S. (2021). Propagation of large earthquakes as self-healing pulses or mild cracks. *Nature*, 591(7849), 252–258. <https://doi.org/10.1038/s41586-021-03248-1>
- Langbein, J., Murray, J. R., & Snyder, H. A. (2006). Coseismic and initial postseismic deformation from the 2004 Parkfield, California, earthquake, observed by global positioning system, electronic distance meter, creepmeters, and borehole strainmeters. *Bulletin of the Seismological Society of America*, 96(4 B), 304–320. <https://doi.org/10.1785/0120050823>
- Lapusta, N. (2001). *Elastodynamic Analyses of Sliding With Rate And State Friction* (Doctoral dissertation May). Harvard University.

- Lapusta, N., Rice, J. R., Ben-Zion, Y., & Zheng, G. (2000). Elastodynamic analysis for slow tectonic loading with spontaneous rupture episodes on faults with rate- and state-dependent friction. *Journal of Geophysical Research*, 105(10), 23765. <https://doi.org/10.1029/2000JB900250>
- Lau, N., Borsa, A. A., & Becker, T. W. (2020). Present-Day Crustal Vertical Velocity Field for the Contiguous United States. *Journal of Geophysical Research: Solid Earth*, 125(10). <https://doi.org/10.1029/2020JB020066>
- Lazecky, M., & Hooper, A. (2022). InSAR-derived horizontal velocities in a global reference frame. *ESSOAr (preprint)*, 1–14. <https://doi.org/https://doi.org/10.1002/essoar.10511058.1>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lehner, F. K., Li, V. C., & Rice, J. R. (1981). Stress diffusion along rupturing plate boundaries. *Journal of Geophysical Research*, 86(B7), 6155–6169. <https://doi.org/10.1029/JB086iB07p06155>
- Li, Y., Martinis, S., & Wieland, M. (2019). Urban flood mapping with an active self-learning convolutional neural network based on TerraSAR-X intensity and interferometric coherence. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152, 178–191. <https://doi.org/10.1016/j.isprsjprs.2019.04.014>
- Li, Z., Muller, J. P., Cross, P., & Fielding, E. J. (2005). Interferometric synthetic aperture radar (InSAR) atmospheric correction: GPS, Moderate Resolution Imaging Spectroradiometer (MODIS), and InSAR integration. *Journal of Geophysical Research: Solid Earth*, 110(3), 1–10. <https://doi.org/10.1029/2004JB003446>
- Li, Z., Wright, T., Hooper, A., Crippa, P., Gonzalez, P., Walters, R., Elliott, J., Ebmeier, S., Hatton, E., & Parsons, B. (2016). Towards InSAR everywhere, all the time, with Sentinel-1. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 41(July), 763–766. <https://doi.org/10.5194/isprsarchives-XLI-B4-763-2016>
- Liang, C., Agram, P., Simons, M., & Fielding, E. J. (2019). Ionospheric Correction of InSAR Time Series Analysis of C-band Sentinel-1 TOPS Data. *IEEE Transactions on Geoscience and Remote Sensing*, 57(9), 6755–6773. <https://doi.org/10.1109/tgrs.2019.2908494>
- Liang, C., Liu, Z., Fielding, E. J., & Bürgmann, R. (2018). InSAR time series analysis of L-Band wide-swath SAR data acquired by ALOS-2. *IEEE Transactions on Geoscience and Remote Sensing*, 56(8), 4492–4506. <https://doi.org/10.1109/TGRS.2018.2821150>
- Lin, Y. N., Jolivet, R., Simons, M., Agram, P. S., Martens, H. R., Li, Z., & Lodi, S. H. (2015). High interseismic coupling in the Eastern Makran (Pakistan) subduction zone. *Earth and Planetary Science Letters*, 420, 116–126. <https://doi.org/10.1016/j.epsl.2015.03.037>

- Lipton, Z. C., Berkowitz, J., & Elkan, C. (2015). A Critical Review of Recurrent Neural Networks for Sequence Learning [arXiv: 1506.00019]. *arXiv:1506.00019 [cs]*. Retrieved March 20, 2019, from <http://arxiv.org/abs/1506.00019>
- Lockner, D. A., Morrow, C., Moore, D., & Hickman, S. (2011). Low strength of deep San Andreas fault gouge from SAFOD core. *Nature*, 472(7341), 82–85. <https://doi.org/10.1038/nature09927>
- Lohman, R. B., & Simons, M. (2005). Some thoughts on the use of InSAR data to constrain models of surface deformation: Noise structure and data down-sampling. *Geochemistry, Geophysics, Geosystems*, 6(1), 1–12. <https://doi.org/10.1029/2004GC000841>
- Loos, S., Lallemand, D., Baker, J., McCaughey, J., Yun, S.-H., Budhathoki, N., Khan, F., & Singh, R. (2020). G-DIF: A geospatial data integration framework to rapidly estimate post-earthquake damage. *Earthquake Spectra*, 36(4), 1695–1718. <https://doi.org/10/gjs6s9>
- Lv, X., Amelung, F., & Shao, Y. (2022). Widespread Aseismic Slip Along the Makran Megathrust Triggered by the 2013 Mw 7.7 Balochistan Earthquake. *Geophysical Research Letters*, 1–11. <https://doi.org/10.1029/2021gl097411>
- Lv, X., & Shao, Y. (2022). Rheology of the Northern Tibetan Plateau Lithosphere Inferred from the Post-Seismic Deformation Resulting from the 2001 Mw 7.8 Kokoxili Earthquake. *Remote Sensing*, 14(5), 1207. <https://doi.org/10.3390/rs14051207>
- Mahapatra, P., der Marel, H. v., van Leijen, F., Samiei-Esfahany, S., Klees, R., & Hanssen, R. (2018). InSAR datum connection using GNSS-augmented radar transponders. *Journal of Geodesy*, 92(1), 21–32. <https://doi.org/10.1007/s00190-017-1041-y>
- Malhotra, P., Vig, L., Shroff, G., & Agarwal, P. (2015). Long Short Term Memory Networks for Anomaly Detection in Time Series. *23rd European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2015 - Proceedings* (pp. 89–94).
- Marinkovic, P., & Larsen, Y. (2015). On Resolving the Local Oscillator Drift Induced Phase Ramps in ASAR and ERS1/2 Interferometric Data—The Final Solution. *Fringe 2015 Workshop (ESA SP - 731)*, (1), 20.
- Mase, C. W., & Smith, L. (1987). Effects of frictional heating on the thermal, hydrologic, and mechanical response of a fault. *Journal of Geophysical Research*, 92(B7), 6249–6272. <https://doi.org/10.1029/JB092iB07p06249>
- Massonnet, D., & Feigl, K. L. (1998). Radar interferometry and its application to changes in the Earth's surface. *Reviews of Geophysics*, 36(4), 441. <https://doi.org/10.1029/97RG03139>

- Massonnet, D., Rossi, M., Carmona, C., Adragna, F., Peltzer, G., Feigl, K., & Rabaute, T. (1993). The displacement field of the Landers earthquake mapped by radar interferometry. *Nature*, 364(6433), 138–142. <https://doi.org/10.1038/364138a0>
- Maurer, J., & Johnson, K. (2014). Fault coupling and potential for earthquakes on the creeping section of the central San Andreas Fault. *Journal of Geophysical Research: Solid Earth*, 119(5), 4414–4428. <https://doi.org/10.1002/2013JB010741>
- McKenzie, D. P., & Parker, R. L. (1967). The North Pacific: an Example of Tectonics on a Sphere. *Nature*, 216(5122), 1276–1280. <https://doi.org/10.1038/2161276a0>
- Merryman Boncori, J. P. (2019). Measuring Coseismic Deformation With Spaceborne Synthetic Aperture Radar: A Review. *Frontiers in Earth Science*, 7(February), 1–20. <https://doi.org/10.3389/feart.2019.00016>
- Milbert, D. (2018). solid: Solid Earth Tide. <https://geodesyworld.github.io/SOFTS/solid.htm>
- Mitchell, T. M., & Faulkner, D. R. (2009). The nature and origin of off-fault damage surrounding strike-slip fault zones with a wide range of displacements: A field study from the Atacama fault system, northern Chile. *Journal of Structural Geology*, 31(8), 802–816. <https://doi.org/10.1016/j.jsg.2009.05.002>
- Monti-Guarnieri, A. V., Brovelli, M. A., Manzoni, M., Mariotti d'Alessandro, M., Molinari, M. E., & Oxoli, D. (2018). Coherent Change Detection for Multi-pass SAR. *IEEE Transactions on Geoscience and Remote Sensing*, 56(11), 6811–6822. <https://doi.org/10.1109/TGRS.2018.2843560>
- Morishita, Y., Lazecky, M., Wright, T., Weiss, J., Elliott, J., & Hooper, A. (2020). LiCSBAS: An Open-Source InSAR Time Series Analysis Package Integrated with the LiCSAR Automated Sentinel-1 InSAR Processor. *Remote Sensing*, 12(3), 424. <https://doi.org/10.3390/rs12030424>
- Motagh, M., Shamshiri, R., Haghshenas Haghighi, M., Wetzel, H. U., Akbari, B., Nahavandchi, H., Roessner, S., & Arabi, S. (2017). Quantifying groundwater exploitation induced subsidence in the Rafsanjan plain, southeastern Iran, using InSAR time-series and in situ measurements. *Engineering Geology*, 218, 134–151. <https://doi.org/10.1016/j.enggeo.2017.01.011>
- Motohka, T., Kankaku, Y., Miura, S., & Suzuki, S. (2019). Alos-4 L-Band SAR Mission and Observation. *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 5271–5273. <https://doi.org/10.1109/IGARSS.2019.8898169>

- Murray, K. D., Lohman, R. B., & Bekaert, D. P. (2021). Cluster-Based Empirical Tropospheric Corrections Applied to InSAR Time Series Analysis. *IEEE Transactions on Geoscience and Remote Sensing*, 59(3), 2204–2212. <https://doi.org/10.1109/TGRS.2020.3003271>
- Musson, R. M. (2009). Subduction in the Western Makran: The historian's contribution. *Journal of the Geological Society*, 166(3), 387–391. <https://doi.org/10.1144/0016-76492008-119>
- Ndikumana, E., Ho Tong Minh, D., Baghdadi, N., Courault, D., & Hossard, L. (2018). Deep Recurrent Neural Network for Agricultural Classification using multitemporal SAR Sentinel-1 for Camargue, France. *Remote Sensing*, 10(8), 1217. <https://doi.org/10.3390/rs10081217>
- Needham, J. (1959). *Science and Civilisation in China: Volume 3, Mathematics and the Sciences of the Heavens and the Earth*. Cambridge University Press. <https://doi.org/10.2307/2051584>
- Neely, W. R., Borsa, A. A., & Silverii, F. (2020). GInSAR: A cGPS Correction for Enhanced InSAR Time Series. *IEEE Transactions on Geoscience and Remote Sensing*, 58(1), 136–146. <https://doi.org/10.1109/TGRS.2019.2934118>
- Nemati, M. (2019). Seismotectonic and seismicity of Makran, a bimodal subduction zone, SE Iran. *Journal of Asian Earth Sciences*, 169(January 2018), 139–161. <https://doi.org/10.1016/j.jseae.2018.08.009>
- Noda, H., & Lapusta, N. (2010). Three-dimensional earthquake sequence simulations with evolving temperature and pore pressure due to shear heating: Effect of heterogeneous hydraulic diffusivity. *Journal of Geophysical Research: Solid Earth*, 115(12), 1–24. <https://doi.org/10.1029/2010JB007780>
- Noda, H., & Lapusta, N. (2013). Stable creeping fault segments can become destructive as a result of dynamic weakening. *Nature*, 493(7433), 518–521. <https://doi.org/10.1038/nature11703>
- Okada, Y. (1985). Surface deformation due to shear and tensile faults in a half-space. *Bulletin of the Seismological Society of America*, 75(4), 1135–1154. <https://doi.org/10.1785/BSSA0750041135>
- Olah, C. (2015). Understanding LSTM Networks [Accessed June 22, 2020]. Retrieved June 22, 2020, from <http://colah.github.io/posts/2015-08-Understanding-LSTMs/>
- Olen, S., & Bookhagen, B. (2018). Mapping Damage-Affected Areas after Natural Hazard Events Using Sentinel-1 Coherence Time Series. *Remote Sensing*, 10(8), 1272. <https://doi.org/10.3390/rs10081272>
- Parizzi, A., Brcic, R., & De Zan, F. (2021). InSAR Performance for Large-Scale Deformation Measurement. *IEEE Transactions on Geoscience and Remote Sensing*, 59(10), 8510–8520. <https://doi.org/10.1109/TGRS.2020.3039006>

- Parizzi, A., Gonzalez, F. R., & Brcic, R. (2020). A covariance-based approach to merging InSAR and GNSS displacement rate measurements. *Remote Sensing*, 12(2). <https://doi.org/10.3390/rs12020300>
- Penney, C., Copley, A., & Oveisi, B. (2015). Subduction tractions and vertical axis rotations in the Zagros-Makran transition zone, SE Iran: The 2013 May 11 Mw 6.1 Minab earthquake. *Geophysical Journal International*, 202(2), 1122–1136. <https://doi.org/10.1093/gji/ggv202>
- Penney, C., Tavakoli, F., Saadat, A., Nankali, H. R., Sedighi, M., Khorrami, F., Sobouti, F., Rafi, Z., Copley, A., Jackson, J., & Priestley, K. (2017). Megathrust and accretionary wedge properties and behaviour in the Makran subduction zone. *Geophysical Journal International*, 209(3), 1800–1830. <https://doi.org/10.1093/gji/ggx126>
- Perfettini, H., & Avouac, J. P. (2004). Stress transfer and strain rate variations during the seismic cycle. *Journal of Geophysical Research: Solid Earth*, 109(6), 1–8. <https://doi.org/10.1029/2003JB002917>
- Pesaresi, M., Gerhardinger, A., & Haag, F. (2007). Rapid damage assessment of built-up structures using VHR satellite data in tsunami-affected areas. *International Journal of Remote Sensing*, 28(13-14), 3013–3036. <https://doi.org/10.1080/01431160601094492>
- Peter, H. (2021). *Copernicus POD Product Handbook: Copernicus Sentinel-1, -2 and -3 Precise Orbit Determination Service (CPOD)* (tech. rep.). GMV. <https://sentinel.esa.int/documents/247904/4599719/Copernicus-POD-Product-Handbook.pdf>
- Peterson, K. E., Barnhart, W. D., & Li, S. (2018). Viscous Accretionary Prisms: Viscoelastic Relaxation of the Makran Accretionary Prism Following the 2013 Baluchistan, Pakistan Earthquake. *Journal of Geophysical Research: Solid Earth*, 123(11), 107–10. <https://doi.org/10.1029/2018JB016057>
- Petit, G., & Luzum, B. (2010). IERS Conventions (IERS Technical Note No. 36). <http://www.iers.org/TN36/>
- Platt, J. D., Rudnicki, J. W., & Rice, J. R. (2014). Stability and localization of rapid shear in fluid-saturated fault gouge: 2. Localized zone width and strength evolution. *Journal of Geophysical Research: Solid Earth*, 119(5), 4334–4359. <https://doi.org/10.1002/2013JB010711>
- Plattner, C., Parizzi, A., Carena, S., Rieger, S. M., Friedrich, A. M., Abolghasem, A. M., & DeZan, F. (2021). Long-lived afterslip of the 2013 Mw 6.1 Minab earthquake detected by Persistent Scatterer Interferometry along the Ier fault (western Makran-Zagros transition zone, Iran). *Geophysical Journal International*, 229(1), 171–185. <https://doi.org/10.1093/gji/ggab456>

- Ponti, D. J., Blair, J. L., Rosa, C. M., Thomas, K., Pickering, A. J., Akciz, S., Angster, S., Avouac, J.-P., Bachhuber, J., Bacon, S., Barth, N., Bennett, S., Blake, K., Bork, S., Brooks, B., Bullard, T., Burgess, P., Chupik, C., Dawson, T., . . . Zinke, R. (2020). Documentation of Surface Fault Rupture and Ground-Deformation Features Produced by the 4 and 5 July 2019 Mw 6.4 and Mw 7.1 Ridgecrest Earthquake Sequence [Publisher: GeoScienceWorld]. *Seismological Research Letters*, 91(5), 2942–2959. <https://doi.org/10.1785/0220190322>
- Priestley, K., Sobouti, F., Mokhtarzadeh, R., Irandoust, M., Ghods, R., Motaghi, K., & Ho, T. (2022). New Constraints for the On-Shore Makran Subduction Zone Crustal Structure. *Journal of Geophysical Research: Solid Earth*, 127(1). <https://doi.org/10.1029/2021JB022942>
- Prizomwala, S. P., Gandhi, D., Bhatt, N., Winkler, W., Kumar, M. R., Makwana, N., & Bhatt, N. (2018). Geological evidence for AD 1008 tsunami along the Kachchh coast, Western India: Implications for hazard along the Makran Subduction Zone. *Scientific Reports*, 8(1), 1–8. <https://doi.org/10.1038/s41598-018-35193-x>
- Qiu, Q., Zhou, Z., Lin, J., Zhang, F., Chen, Z., & Yang, X. (2022). Great earthquake and tsunami potential in the eastern Makran subduction zone: New insights from geodetic and structural constraints. *Tectonophysics*, 837(September 2021), 229462. <https://doi.org/10.1016/j.tecto.2022.229462>
- Rajendran, C. P., Heidarzadeh, M., Sanwal, J., Karthikeyan, A., & Rajendran, K. (2021). The Orphan Tsunami of 1524 on the Konkan Coast, Western India, and Its Implications. *Pure and Applied Geophysics*, 178(12), 4697–4716. <https://doi.org/10.1007/s00024-020-02575-0>
- Regard, V., Bellier, O., Thomas, J. C., Boursès, D., Bonnet, S., Abbassi, M. R., Braucher, R., Mercier, J., Shabanian, E., Soleymani, S., & Feghhi, K. (2005). Cumulative right-lateral fault slip rate across the Zagros-Makran transfer zone: Role of the Minab-Zendan fault system in accommodating Arabia-Eurasia convergence in southeast Iran. *Geophysical Journal International*, 162(1), 177–203. <https://doi.org/10.1111/j.1365-246X.2005.02558.x>
- Reid, H. F. (1910). The California Earthquake of April 18, 1906: Volume II (The Mechanics of the Earthquake). *State Earthquake Investigation Commission*, 192.
- Reinhardt, O., & Oldroyd, D. (1983). Kant's theory of earthquakes and volcanic action. *Annals of Science*, 40(3), 247–272. <https://doi.org/10.1080/00033798300200221>
- Rempel, A. W., & Rice, J. R. (2006). Thermal pressurization and onset of melting in fault zones. *Journal of Geophysical Research: Solid Earth*, 111(9). <https://doi.org/10.1029/2006JB004314>

- Rice, J. R., & Ruina, A. L. (1983). Stability of steady frictional slipping. *Journal of Applied Mechanics, Transactions ASME*, 50(2), 343–349. <https://doi.org/10.1115/1.3167042>
- Rice, J. R. (2006). Heating and weakening of faults during earthquake slip. *Journal of Geophysical Research: Solid Earth*, 111(5), 1–29. <https://doi.org/10.1029/2005JB004006>
- Rice, J. R., Rudnicki, J. W., & Platt, J. D. (2014). Stability and localization of rapid shear in fluid-saturated fault gouge: 1. Linearized stability analysis. *Journal of Geophysical Research: Solid Earth*, 119(5), 4311–4333. <https://doi.org/10.1002/2013JB010710>
- Riddell, A. R., King, M. A., & Watson, C. S. (2020). Present-Day Vertical Land Motion of Australia From GPS Observations and Geophysical Models. *Journal of Geophysical Research: Solid Earth*, 125(2). <https://doi.org/10.1029/2019JB018034>
- Rodriguez, E., & Martin, J. (1992). Theory and design of interferometric synthetic aperture radars. *IEEE Proceedings F Radar and Signal Processing*, 139(2), 147. <https://doi.org/10.1049/ip-f-2.1992.0018>
- Rosen, P. A., Gurrola, E., Sacco, G. F., & Zebker, H. (2012). The InSAR scientific computing environment. *Proceedings of the European Conference on Synthetic Aperture Radar, EUSAR, 2012-April*, 730–733.
- Rosen, P. A., Hensley, S., Joughin, I., Li, F., Madsen, S., Rodriguez, E., & Goldstein, R. (2000). Synthetic aperture radar interferometry. *Proceedings of the IEEE*, 88(3), 333–382. <https://doi.org/10.1109/5.838084>
- Ross, Z. E., Idini, B., Jia, Z., Stephenson, O. L., Zhong, M., Wang, X., Zhan, Z., Simons, M., Fielding, E. J., Yun, S.-H., Hauksson, E., Moore, A. W., Liu, Z., & Jung, J. (2019). Hierarchical interlocked orthogonal faulting in the 2019 Ridgecrest earthquake sequence. *Science*, 366(6463), 346–351. <https://doi.org/10.1126/science.aaz0109>
- Ross, Z. E., Yue, Y., Meier, M.-A., Hauksson, E., & Heaton, T. H. (2019). PhaseLink: A Deep Learning Approach to Seismic Phase Association. *Journal of Geophysical Research: Solid Earth*, 124(1), 856–869. <https://doi.org/10.1029/2018JB016674>
- Rubin, A. M., & Ampuero, J. P. (2005). Earthquake nucleation on (aging) rate and state faults. *Journal of Geophysical Research: Solid Earth*, 110(11), 1–24. <https://doi.org/10.1029/2005JB003686>
- Rubino, V., Lapusta, N., & Rosakis, A. J. (2022). Intermittent lab earthquakes in dynamically weakening fault gouge. *Nature*, 606(April 2021). <https://doi.org/10.1038/s41586-022-04749-3>

- Ruina, A. (1983). Slip instability and state variable friction laws. *Journal of Geophysical Research: Solid Earth*, 88(B12), 10359–10370. <https://doi.org/10.1029/JB088iB12p10359>
- Salah, P., Sasaki, J., & Soltanpour, M. (2021). Comprehensive Probabilistic Tsunami Hazard Assessment in the Makran Subduction Zone. *Pure and Applied Geophysics*, 178(12), 5085–5107. <https://doi.org/10.1007/s00024-021-02725-y>
- Savage, J. C. (1983). A dislocation model of strain accumulation and release at a subduction zone. *Journal of Geophysical Research*, 88(B6), 4984–4996. <https://doi.org/10.1029/JB088iB06p04984>
- Schlüter, H. U., Prexl, A., Gaedicke, C., Roeser, H., Reichert, C., Meyer, H., & Von Daniels, C. (2002). The Makran accretionary wedge: Sediment thicknesses and ages and the origin of mud volcanoes. *Marine Geology*, 185(3-4), 219–232. [https://doi.org/10.1016/S0025-3227\(02\)00192-5](https://doi.org/10.1016/S0025-3227(02)00192-5)
- Scholz, C. (1998). Earthquakes and friction laws. *Nature*, 391, 37–42. <https://doi.org/10.1038/34097>
- Scott, C., Bunds, M., Shirzaei, M., & Toke, N. (2020). Creep Along the Central San Andreas Fault From Surface Fractures, Topographic Differencing, and InSAR. *Journal of Geophysical Research: Solid Earth*, 125(10), 1–16. <https://doi.org/10.1029/2020JB019762>
- Segall, P., & Rice, J. R. (1995). Dilatancy, compaction, and slip instability of a fluid-infiltrated fault. *Journal of Geophysical Research: Solid Earth*, 100(B11), 22155–22171. <https://doi.org/10.1029/95JB02403>
- Shah-hosseini, M., Morhange, C., Naderi Beni, A., Marriner, N., Lahijani, H., Hamzeh, M., & Sabatier, F. (2011). Coastal boulders as evidence for high-energy waves on the Iranian coast of Makran. *Marine Geology*, 290(1-4), 17–28. <https://doi.org/10.1016/j.margeo.2011.10.003>
- Sharma, P. (2019). Updates in Commissioning Timeline for NASA-ISRO Synthetic Aperture Radar (NISAR). *2019 IEEE Aerospace Conference*, 1–12. <https://doi.org/10.1109/AERO.2019.8741891>
- Shen, Z.-K., Jackson, D. D., Feng, Y., Cline, M., Kim, M., Fang, P., & Bock, Y. (1994). Postseismic deformation following the Landers earthquake, California, 28 June 1992. *Bulletin of the Seismological Society of America*, 84(3), 780–791. <https://doi.org/10.1785/BSSA0840030780>
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-k., & Woo, W.-c. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting, 1–12. <https://doi.org/10.1074/jbc.M200827200>
- Shirzaei, M. (2015). A seamless multitrack multitemporal InSAR algorithm. *Geochemistry, Geophysics, Geosystems*, 16(5), 1656–1669. <https://doi.org/10.1002/2015GC005759>

- Sibson, R. H. (1973). Interactions between Temperature and Pore-Fluid Pressure during Earthquake Faulting and a Mechanism for Partial or Total Stress Relief. *Nature*, 243, 66–68. <https://doi.org/10.1038/physci243066a0>
- Sibson, R. H. (2003). Thickness of the seismic slip zone. *Bulletin of the Seismological Society of America*, 93(3), 1169–1178. <https://doi.org/10.1785/0120020061>
- Sieh, K. E. (1978a). Central California Foreshocks of the Great 1857 Earthquake. *Bulletin of the Seismological Society of America*, 68(6), 1731–1749. <https://doi.org/10.1785/BSSA0680061731>
- Sieh, K. E. (1978b). Slip Along the San Andreas Fault Associated with the Great 1857 Earthquake. *Bulletin of the Seismological Society of America*, 68(5), 1421–1448. <https://doi.org/10.1785/BSSA0680051421>
- Sieh, K. E., Stuiver, M., & Brillinger, D. (1989). A more precise chronology of earthquakes produced by the San Andreas Fault in southern California. *Journal of Geophysical Research*, 94(B1), 603. <https://doi.org/10.1029/JB094iB01p00603>
- Simons, M., Fialko, Y., & Rivera, L. (2002). Coseismic Deformation from the 1999 Mw 7.1 Hector Mine, California, Earthquake as Inferred from InSAR and GPS Observations. *Bulletin of the Seismological Society of America*, 92(4), 1390–1402. <https://doi.org/10.1785/0120000933>
- Simons, M., Minson, S. E., Sladen, A., Ortega, F., Jiang, J., Owen, S. E., Meng, L., Ampuero, J.-P., Wei, S., Chu, R., Helmberger, D. V., Kanamori, H., Hetland, E., Moore, A. W., & Webb, F. H. (2011). The 2011 Magnitude 9.0 Tohoku-Oki Earthquake: Mosaicking the Megathrust from Seconds to Centuries. *Science*, 332(6036), 1421–1425. <https://doi.org/10.1126/science.1206731>
- Simons, M., & Rosen, P. (2015). Interferometric Synthetic Aperture Radar Geodesy. *Treatise on geophysics* (pp. 339–385). Elsevier. <https://doi.org/10.1016/B978-0-444-53802-4.00061-0>
- Simpson, R. W., Barall, M., Langbein, J., Murray, J. R., & Rymer, M. J. (2006). San Andreas fault geometry in the Parkfield, California, region. *Bulletin of the Seismological Society of America*, 96(4 B), 28–37. <https://doi.org/10.1785/0120050824>
- Smith, G. L., McNeill, L. C., Wang, K., He, J., & Henstock, T. J. (2013). Thermal structure and megathrust seismogenic potential of the Makran subduction zone. *Geophysical Research Letters*, 40(8), 1528–1533. <https://doi.org/10.1002/grl.50374>
- Stathas, A., & Stefanou, I. (2022). Fault friction under thermal pressurization during large coseismic-slip Part II: Expansion to the model of frictional slip. *arXiv Preprint*. <http://arxiv.org/abs/2205.00316>

- Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to Sequence Learning with Neural Networks [arXiv: 1409.3215]. *arXiv:1409.3215 [cs]*. Retrieved March 20, 2019, from <http://arxiv.org/abs/1409.3215>
- Tanikawa, W., & Shimamoto, T. (2009). Frictional and transport properties of the Chelungpu fault from shallow borehole data and their correlation with seismic behavior during the 1999 Chi-Chi earthquake. *Journal of Geophysical Research: Solid Earth*, 114(1), 1–15. <https://doi.org/10.1029/2008JB005750>
- Tarayre, H., & Massonnet, D. (1996). Atmospheric Propagation heterogeneities revealed by ERS-1 interferometry. *Geophysical Research Letters*, 23(9), 989–992. <https://doi.org/10.1029/96GL00622>
- Titus, S. J. (2006). Thirty-Five-Year Creep Rates for the Creeping Segment of the San Andreas Fault and the Effects of the 2004 Parkfield Earthquake: Constraints from Alignment Arrays, Continuous Global Positioning System, and Creepmeters. *Bulletin of the Seismological Society of America*, 96(4B), S250–S268. <https://doi.org/10.1785/0120050811>
- Toké, N. A., Arrowsmith, J. R., Rymer, M. J., Landgraf, A., Haddad, D. E., Busch, M., Cayan, J., & Hannah, A. (2011). Late Holocene slip rate of the San Andreas fault and its accommodation by creep and moderate-magnitude earthquakes at Parkfield, California. *Geology*, 39(3), 243–246. <https://doi.org/10.1130/G31498.1>
- Toké, N. A., Arrowsmith, J. R., Young, J. J., & Crosby, C. J. (2006). Paleoseismic and postseismic observations of surface slip along the parkfield segment of the san andreas fault. *Bulletin of the Seismological Society of America*, 96(4 B), S221–S238. <https://doi.org/10.1785/0120050809>
- Toké, N. A., & Arrowsmith, R. (2015). *Examining the Cause of Prehistoric Ground Deformation at the Dry Lake Valley Site along the Creeping Section of the San Andreas Fault, San Benito, California* (tech. rep.). https://files.scec.org/s3fs-public/reports/2013/69f63a37e3d0cac81326dccc45d3e99/13147_report.pdf
- Tsutsumi, A., & Shimamoto, T. (1997). High-velocity frictional properties of gabbro. *Geophysical Research Letters*, 24(6), 699–702. <https://doi.org/10.1029/97GL00503>
- Tullis, T. E. (2007). Friction of Rock at Earthquake Slip Rates. *Treatise on geophysics* (pp. 131–152). <https://doi.org/10.1016/B978-044452748-6.00064-X>
- Ujiie, K., Kameyama, M., & Yamaguchi, A. (2010). Geological record of thermal pressurization and earthquake instability of subduction thrusts. *Tectonophysics*, 485(1-4), 260–268. <https://doi.org/10.1016/j.tecto.2010.01.002>
- Ulaby, F., & Long, D. (2015). *Microwave Radar and Radiometric Remote Sensing*. Artech House. <http://cds.cern.ch/record/2197387>

- United Nations Institute for Training and Research. (2017). Damaged Structures and Related Density Map in Sarpol-e-Zahab, Kermanshah, Iran. Retrieved June 23, 2020, from <https://unitar.org/maps/map/2732>
- United States Geological Survey. (2016). M 6.2 - 10km SE of Norcia, Italy [Accessed June 22, 2020]. Retrieved June 23, 2020, from <https://earthquake.usgs.gov/earthquakes/eventpage/us10006g7d/executive>
- United States Geological Survey. (2017). M 7.3 - 29km S of Halabjah, Iraq [Accessed June 22, 2020]. Retrieved June 23, 2020, from <https://earthquake.usgs.gov/earthquakes/eventpage/us2000bmcg/executive>
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention Is All You Need. (Nips). <https://doi.org/10.1017/S0952523813000308>
- Viesca, R. C., & Garagash, D. I. (2015). Ubiquitous weakening of faults due to thermal pressurization. *Nature Geoscience*, 8(11), 875–879. <https://doi.org/10.1038/ngeo2554>
- Voigt, S., Giulio-Tonolo, F., Lyons, J., Kučera, J., Jones, B., Schneiderhan, T., Platzeck, G., Kaku, K., Hazarika, M. K., Czarán, L., Li, S., Pedersen, W., James, G. K., Proy, C., Muthike, D. M., Bequignon, J., & Guha-Sapir, D. (2016). Global trends in satellite-based emergency mapping. *Science*, 353(6296), 247–252. <https://doi.org/10.1126/science.aad8728>
- Wang, H., Liu, M., Duan, B., & Cao, J. (2020). Rupture propagation along stepovers of strike-slip faults: Effects of initial stress and fault geometry. *Bulletin of the Seismological Society of America*, 110(3), 1011–1024. <https://doi.org/10.1785/0120190233>
- Wang, K., Zhu, Y., Nissen, E., & Shen, Z. K. (2021). On the Relevance of Geodetic Deformation Rates to Earthquake Potential. *Geophysical Research Letters*, 48(11), 1–9. <https://doi.org/10.1029/2021GL093231>
- Washaya, P., Balz, T., & Mohamadi, B. (2018). Coherence Change-Detection with Sentinel-1 for Natural and Anthropogenic Disaster Monitoring in Urban Areas. *Remote Sensing*, 10(7), 1026. <https://doi.org/10.3390/rs10071026>
- Weiss, J. R., Walters, R. J., Morishita, Y., Wright, T. J., Lazecky, M., Wang, H., Hussain, E., Hooper, A. J., Elliott, J. R., Rollins, C., Yu, C., González, P. J., Spaans, K., Li, Z., & Parsons, B. (2020). High-Resolution Surface Velocities and Strain for Anatolia From Sentinel-1 InSAR and GNSS Data. *Geophysical Research Letters*, 47(17). <https://doi.org/10.1029/2020GL087376>
- Wibberley, C. A., & Shimamoto, T. (2003). Internal structure and permeability of major strike-slip fault zones: the Median Tectonic Line in Mie Prefecture, Southwest Japan. *Journal of Structural Geology*, 25(1), 59–78. [https://doi.org/10.1016/S0191-8141\(02\)00014-7](https://doi.org/10.1016/S0191-8141(02)00014-7)

- Wibberley, C. A., & Shimamoto, T. (2005). Earthquake slip weakening and asperities explained by thermal pressurization. *Nature*, 436(7051), 689–692. <https://doi.org/10.1038/nature03901>
- Wieland, M., Liu, W., & Yamazaki, F. (2016). Learning change from Synthetic Aperture Radar images: Performance evaluation of a Support Vector Machine to detect earthquake and tsunami-induced changes. *Remote Sensing*, 8(10). <https://doi.org/10.3390/rs8100792>
- Wright, T. J., Parsons, B. E., & Lu, Z. (2004). Toward mapping surface deformation in three dimensions using InSAR. *Geophysical Research Letters*, 31(1), 1–5. <https://doi.org/10.1029/2003GL018827>
- Xu, X., & Sandwell, D. T. (2020). Toward Absolute Phase Change Recovery with InSAR: Correcting for Earth Tides and Phase Unwrapping Ambiguities. *IEEE Transactions on Geoscience and Remote Sensing*, 58(1), 726–733. <https://doi.org/10.1109/TGRS.2019.2940207>
- Xu, X., Sandwell, D. T., Klein, E., & Bock, Y. (2021). Integrated Sentinel-1 InSAR and GNSS Time-Series Along the San Andreas Fault System. *Journal of Geophysical Research: Solid Earth*, 126(11), 1–14. <https://doi.org/10.1029/2021JB022579>
- Xu, Z.-W., Wu, J., & Wu, Z.-S. (2004). A survey of ionospheric effects on space-based radar. *Waves in Random Media*, 14(2), S189–S273. <https://doi.org/10.1088/0959-7174/14/2/008>
- Yang, X., Qiu, Q., Feng, W., Lin, J., Zhang, J., Zhou, Z., & Zhang, F. (2022). Mechanism of the 2017 Mw 6.3 Pasni earthquake and its significance for future major earthquakes in the eastern Makran. *Geophysical Journal International*, 1–39. <https://doi.org/10.1093/gji/ggac257>
- Yu, C., Penna, N. T., & Li, Z. (2020). Ocean Tide Loading Effects on InSAR Observations Over Wide Regions. *Geophysical Research Letters*, 47(15). <https://doi.org/10.1029/2020GL088184>
- Yun, S.-H., Hudnut, K., Owen, S., Webb, F., Simons, M., Sacco, P., Gurrola, E., Manipon, G., Liang, C., Fielding, E., Milillo, P., Hua, H., & Coletta, A. (2015). Rapid Damage Mapping for the 2015 Mw 7.8 Gorkha Earthquake Using Synthetic Aperture Radar Data from COSMO–SkyMed and ALOS-2 Satellites. *Seismological Research Letters*, 86(6), 1549–1556. <https://doi.org/10.1785/0220150152>
- Yunjun, Z., Fattahi, H., & Amelung, F. (2019). Small baseline InSAR time series analysis: Unwrapping error correction and noise reduction. *Computers and Geosciences*, 133(May), 104331. <https://doi.org/10.1016/j.cageo.2019.104331>

- Yunjun, Z., Fattahi, H., Pi, X., Rosen, P., Simons, M., Agram, P., & Aoki, Y. (2022). Range Geolocation Accuracy of C/L-band SAR and its Implications for Operational Stack Coregistration. *IEEE Transactions on Geoscience and Remote Sensing*, 2892(100), 1–1. <https://doi.org/10.1109/TGRS.2022.3168509>
- Zebker, H. A., & Villasenor, J. (1992). Decorrelation in interferometric radar echoes. *IEEE Transactions on Geoscience and Remote Sensing*, 30(5), 950–959. <https://doi.org/10.1109/36.175330>
- Zheng, Y., Fattahi, H., Agram, P., Simons, M., & Rosen, P. (2022). On Closure Phase and Systematic Bias in Multilooked SAR Interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–11. <https://doi.org/10.1109/TGRS.2022.3167648>
- Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2021). Informer: Beyond Efficient Transformer for Long Sequence Time-Series Forecasting. *35th AAAI Conference on Artificial Intelligence, AAAI 2021, 12B*, 11106–11115.
- Zielke, O., Arrowsmith, J. R., Ludwig, L. G., & Akciz, S. O. (2012). High-resolution topography-derived offsets along the 1857 Fort Tejon earthquake rupture trace, San Andreas fault. *Bulletin of the Seismological Society of America*, 102(3), 1135–1154. <https://doi.org/10.1785/0120110230>
- Zimmaro, P., Nweke, C. C., Hernandez, J. L., Hudson, K. S., Hudson, M. B., Ahdi, S. K., Boggs, M. L., Davis, C. A., Goulet, C. A., Brandenburg, S. J., Hudnut, K. W., & Stewart, J. P. (2020). Liquefaction and Related Ground Failure from July 2019 Ridgecrest Earthquake Sequence. *Bulletin of the Seismological Society of America*. <https://doi.org/10.1785/0120200025>

Appendix A

SUPPLEMENTARY MATERIALS FOR CHAPTER 1: THE IMPACT OF PLATE MOTIONS ON LONG-WAVELENGTH INSAR-DERIVED VELOCITY FIELDS

A.1 Introduction

Here, we present more details on the SAR data sources (Table A.1) and processing steps (Section A.2) from Chapter 1. In Section A.3, we give further details on the corrections that we apply to our data and the possible sources of residual signals after these corrections are applied. In Figures A.1-A.5, we present the velocity corrections for all tracks other than Makran track 86, which is shown in the main text (Figure 1.2). In Section A.4, we provide further information on how bulk vertical and horizontal motions impact InSAR velocity fields, and in Figure A.6, we present a figure that can be used to estimate the scale of velocity ramp for ascending (ASC) and descending (DSC) tracks for a given plate velocity. In Section 1.5, we stated that plate motion could bias calculations of 3D velocities from multiple overlapping InSAR tracks, which we illustrate in Section A.5 and Figure A.7.

A.2 InSAR Processing Details

We download Sentinel-1 Single Look Complex (SLC) data from the Alaska Satellite Facility. We process our data using the InSAR Scientific Computing Environment (Rosen et al. (2012), and see <https://github.com/isce-framework/isce2>). For the Makran data (tracks 86 and 20), we use the topsApp processing chain, where each interferometric pair is processed separately. Other tracks are processed using topsStack, which coregisters all SLCs to a single reference image at the start of the processing. topsStack therefore avoids the need to repeat the coregistration for every interferometric pair, reducing computational expense compared to topsApp.

Each acquisition is used to form interferograms with the next three SLCs. We use satellite orbits from the Copernicus Precise Orbit Determination service and digital elevation models from the Shuttle Radar Topography Mission to remove the phase contribution from the viewing geometry. To correct for azimuth misregistration we apply enhanced spectral diversity (ESD) (De Zan et al., 2014) for each pair. When processing data using topsStack, we estimate azimuth misregistration for each date in a network sense based on ESD results for each pair (Fattahi, Agram, et al., 2017),

but do not use the network approach for topsApp. The different approaches are unlikely to have a significantly different impact on the long wavelength velocity field from long time series. Two dimensional phase unwrapping is performed using the SNAPHU algorithm for each interferogram (Chen & Zebker, 2002).

We use the range split-spectrum method of Liang et al. (2019) to estimate the ionospheric phase screen in each interferogram. For topsStack data, the ionospheric phase is estimated for each interferogram, then for each date by network inversion (Fattahi, Simons, et al., 2017). The ionosphere correction is then applied in the time series domain. For topsApp data, the ionospheric phase screen is estimated and removed before final phase unwrapping. As long as we remove interferograms where there are clear unwrapping errors in the ionosphere, we expect the difference between these approaches to be insignificant.

We calculate InSAR time series from the interferometric network using the SBAS methodology (Berardino et al., 2002), implemented in MintPy (Yunjun et al., 2019). The secular velocity is obtained from a linear, least-squares fit to the time series. We use the PyAPS software package to apply corrections for the troposphere with ERA-5 weather models (Jolivet et al., 2011; Jolivet, Agram, et al., 2014). Solid Earth tides removal is performed using PySolid (Milbert, 2018; Yunjun et al., 2022), and the Digital Elevation Model (DEM) error is estimated from the correlation between the residual time series and perpendicular baseline, as described in Fattahi and Amelung (2013). All of the above corrections are applied to the deformation time series. We apply the plate motion adjustment in the velocity domain. As we are performing linear least-squares fits to the time series, the plate motion correction could be straightforwardly applied to the time series (assuming a constant plate rate) with identical results. All other corrections could similarly be applied either in the velocity domain or the time series domain and result in the same final velocity, assuming that we obtain the velocities from a linear fit to the time series.

For the Makran and Aqaba data sets we mask out unreliable data based on the union of connected component masks from every interferogram (i.e., areas identified as unreliably unwrapped by the SNAPHU algorithm (Chen & Zebker, 2002)). This is a conservative approach to masking that removes a large amount of the data, but allows us to focus on long-wavelength features. The lower coherence for our Australian tracks means that the connected component mask removes too much data to view the spatial velocity pattern, so we instead mask out points with an average

spatial coherence below 0.82. This less conservative masking strategy leads us to include noisier data, possibly biasing our estimation of the across track ramp.

A.3 Further Details on InSAR Corrections and Residuals

To illustrate the influence of the corrections on the final secular velocity, we perform a linear, least-squares fit to the time series of each of the applied corrections (these fits are shown in the second row of Figures 2 and [A.1-A.5](#)). While we fit a linear velocity term for the figures to show how the corrections affect the measured secular velocity, we do not expect these corrections to have a linear variation in time. As these corrections are not perfect, and we have not accounted for all potential contributors to our velocity fields, we also need to consider what other residuals may remain in our time series. Of particular interest is the extent to which the velocity ramps before plate motion correction can be attributed to sources other than plate motion. Below, we present more details on these corrections and the residuals that may remain in our data.

The strength of ionospheric signals is controlled by the density of charged particles in the ionosphere (Gomba et al., [2016](#)). Ionization is mainly due to solar radiation, meaning the ionospheric signal depends on factors such as the time of day, the geomagnetic latitude, and the approximately 11-year solar cycle (Liang et al., [2019](#)). Sentinel-1 ascending tracks are acquired at dusk (around 6 pm local solar time), and suffer from much stronger ionospheric impacts than the Sentinel-1 descending tracks which are acquired at dawn (around 6 am local solar time). The most recent peak of solar activity occurred in April 2014, decreasing to a minimum in December 2019, meaning that data earlier in our time series (which start in late 2014 for Aqaba and Makran) have much stronger ionospheric effects than later dates. This long-term variation in the ionosphere means that we cannot average out the ionospheric signal by fitting to several years of data, making the ionosphere the most significant correction to the secular velocity for the ascending tracks we study.

Ionospheric corrections clearly substantially reduce long-wavelength velocity ramps for the ascending tracks we present here, but it is hard to quantify what residual ionospheric signals may be left without comparing to external data. Liang et al. ([2018](#)) compared the long-wavelength signal with GNSS velocities after ionosphere correction using the split-spectrum method. They found good agreement between the long-wavelength InSAR and GNSS velocities, suggesting that residual iono-

sphere was not having a significant impact on the long-wavelength InSAR signals after correction.

Errors in the ionosphere estimation can be introduced by unwrapping errors in the sub-band interferograms (Gomba et al., 2016; Liang et al., 2019) and before time series processing we examine the ionospheric phase estimates and remove those that have obvious unwrapping errors. Unwrapping the sub-band interferograms requires strong filtering (Liang et al., 2019), meaning that the split-spectrum method is not able to capture short-wavelength variations in the ionospheric signal. We observe the strength of the calculated long-wavelength ionospheric signal to vary with solar activity, which declined from a peak in April 2014 to a minimum in December 2019, and we would expect the short wavelength ionosphere to follow the same pattern. If we have a substantial component of short wavelength ionospheric noise, this might then show up as larger residuals in the time series for the earlier dates (2014-15) compared to dates in late 2019. To test this, we first fit and remove from every pixel a functional form (a linear term plus the amplitude and phase of annual and semi-annual sinusoids), then calculate the root mean square (RMS) of the residuals for every acquisition. The primary variation in the residual RMS is seasonal (e.g., varying between about 10 mm and 45 mm for track 86 (ASC) in the Makran), likely due to incompletely removed tropospheric delays. There is not a strong trend in the RMS over longer timescales, suggesting the short wavelength ionosphere remaining in our data is not having a substantial effect, particularly compared with residual troposphere.

The geometry of the Sentinel-1 bursts can result in ionospheric phase discontinuities at the burst boundaries (see Section III.C in Liang et al. (2019)). The topsApp code is able to compute and remove these discontinuities as part of the ionosphere estimation. The removal of burst discontinuities due to the ionosphere can be most clearly seen in the ladder-like pattern of the velocity in Figure A.1(f), which is the calculated ionospheric signal removed from A.1(a). Here, the ionosphere is contributing ramps of up to 0.5 mm/yr over the 20 km along-track width of the bursts, with sharp discontinuities of up to 0.5 mm/yr at the burst boundaries. These ramps and discontinuities are removed by the ionosphere correction. The burst ramps and discontinuities are also present in the ionosphere estimation of the ascending track (Figure 1.2(f)), which is also performed using topsApp, but cannot be seen in the figure due to the wider dynamic range of the ionosphere.

The topsStack code is not yet able to take account of the burst geometry when computing the ionospheric phase screen. For the tracks processed using topsStack (Aqaba and Australia), the lack of accounting for the burst geometry results in small discontinuities in the velocity field at the burst boundaries in the final velocity field, which can be seen on close inspection of Figures A.2-A.5. These discontinuities are below the 1 mm/yr level, and are short-wavelength ramps within the bursts, meaning they do not affect the long-wavelength velocity field.

If there are errors in the long-wavelength ionospheric phase removal, and these errors are proportional to the size of the ionospheric signal, they could be revealed by differences in the long-wavelength residual velocity field between ascending and descending tracks covering the same region. For both Makran and Aqaba, there do not seem to be substantial differences in the long-wavelength residuals between the ascending and descending tracks that could reasonably be attributed to uncompensated ionosphere. These results suggest that the ionosphere corrections are able to remove a substantial fraction of the long-wavelength ionospheric signal, but we are not able to precisely quantify the residual long-wavelength ionosphere.

Unlike ionospheric signals, the tropospheric phase is dominated by seasonal variability, meaning longer time series will reduce the tropospheric effect on the estimated secular velocity (Fattahi & Amelung, 2015; Parizzi et al., 2021). The troposphere signal varies depending on the location, but can have an amplitude equivalent to tens of centimeters of ground deformation. This signal can be mitigated, but not eliminated, by tropospheric models (Bekaert et al., 2015; Fattahi & Amelung, 2015; Parizzi et al., 2021). Fattahi and Amelung (2015) examined the tropospheric noise just to the west of the Makran subduction zone, finding that, after corrections with the ERA-Interim weather model, the tropospheric delay would lead to uncertainties in the InSAR-derived velocities of 2 mm/yr over 100km and 4 mm/yr over 400km with 7 years of Envisat data. Parizzi et al. (2021) evaluated the tropospheric contribution to InSAR-derived velocity uncertainties for locations including the Makran and Gulf of Aqaba. They estimated that uncertainties were in the range 2-4 mm/yr over 150 km for the Makran, and 1-2 mm/yr over 150 km in the Gulf of Aqaba, using at least four years of Sentinel-1 data and corrections from the ERA5 weather model from ECMWF. They stated that tropospheric signals are the limiting factor for measuring large-scale deformation using InSAR, as the residuals from the ionosphere and solid Earth tides are negligible after correction, and the contribution of orbital errors is also not significant compared with the troposphere.

Table A.1: Summary of Sentinel-1 SAR data used. ASC: Ascending track. DSC: Descending track. Aqn. no.: Number of acquisitions. Makran and Aqaba use only Sentinel-1A data, Australia uses only Sentinel-1B. The plates used for plate motion corrections are listed in the final column and are taken from Table 1 of Altamimi et al. (2017).

Region	Track	Direction	Start date	End date	Aqn. no.	Plate
Makran	86	ASC	20141023	20210401	158	EURA
Makran	20	DSC	20141007	20210328	150	EURA
Aqaba	87	ASC	20141104	20220102	203	ARAB
Aqaba	21	DSC	20150815	20210503	160	ARAB
Australia	46	DSC	20161003	20211212	157	AUST
Australia	119	DSC	20160926	20211217	158	AUST

Our data spans 5-7 years, with 150-200 acquisitions per track (Table A.1), and we use the ERA5 weather model in our corrections (Hersbach et al., 2020). We therefore expect the contribution of the troposphere to the velocity field to be at or below the levels outlined by Parizzi et al. (2021) for Makran and Aqaba, but we have not directly estimated this for our data. We also do not have comparable estimates for the Australian tracks.

Tracks in the same region acquired at the same time of day should have statistically similar tropospheric signals in terms of the amplitude and spatial correlation of the phase in each acquisition. Tropospheric signals can therefore create similar InSAR velocity uncertainties for nearby tracks (e.g., Fattahi and Amelung (2015)). However, the troposphere is uncorrelated at timescales longer than a day (Emardson et al., 2003), meaning tracks in the same region whose acquisitions are separated by more than a day (as is the case for all of our tracks in the same region) will have different realisations of the tropospheric phase in each acquisition. We would therefore expect the long-wavelength velocity residuals from the troposphere to be different for each track in a given region, rather than systematically contributing velocity ramps in a particular direction. This expectation also applies to tropospheric residuals that remain after troposphere corrections have been applied.

Orbital errors can also create long-wavelength artefacts in individual interferograms, which can impact the velocity estimate. Fattahi and Amelung (2014) found that for Sentinel-1, orbital errors would lead to velocity uncertainties at the level of 0.5 mm/yr over 100 km, assuming 15 acquisitions per year for 8 years. Similarly to the troposphere, we would expect the orbital error to have different realisations in each

track, meaning that the velocity ramps from orbital errors would not be expected to occur systematically in any particular direction.

The solid Earth tides (SET) are another long-wavelength signal. SET can contribute along-track InSAR phase ramps, and the varying satellite line-of-sight (LOS) will also mean that SET deformation creates ramps in the range direction as well (X. Xu & Sandwell, 2020), in a similar fashion to the LOS projection of plate motion. The SET are periodic, and the Sentinel-1 sampling frequency causes diurnal and semi-diurnal tidal periods to be aliased to longer periods of up to a year (X. Xu & Sandwell, 2020). Using multi-year time series reduces the influence of these periodic terms on the secular velocity estimation. The SET estimate is accurate to better than 1 mm (Petit & Luzum, 2010), so after correction we expect its residual contribution to the long-wavelength velocity to be much smaller than other signals discussed here.

We do not correct for the deformation due to ocean tidal loading (OTL) in this work (Dicaprio et al., 2008). OTL is a periodic, long-wavelength signal, much like the SET, and the satellite LOS variation causes changing sensitivity to this deformation in the range direction. The periodic nature of the signal means that it will cause a smaller bias in the estimated secular velocity in longer time series (Yu et al., 2020), similarly to the SET. We expect deformation due to OTL to decay with distance from the ocean, meaning the direction and magnitude of any long-wavelength velocity residuals due to OTL will be different for each of our regions. Yu et al. (2020) identified regions of the Earth where OTL corrections would be significant for reducing long-wavelength residuals (see Figure 1 in Yu et al. (2020)). As all of our study areas lie outside these regions, we believe that the OTL signal will have a smaller influence on the velocity than corrections such as the ionosphere, troposphere and plate motion, but may be comparable to the SET, particularly for the Makran.

The DEM error signal is generally short wavelength, as well as depending on the satellite baseline (Fattahi & Amelung, 2013), which is well controlled for Sentinel-1 (Z. Li et al., 2016) and varies largely randomly in time. The DEM error signal is estimated from correlation between the perpendicular baseline and time series residuals, meaning that large residuals from other sources, such as the troposphere, can potentially bias the DEM error estimation (Fattahi & Amelung, 2013). It is possible that we see this biasing in Figure 1.2(i) at around (26°N, 60°E), but the impact is still small, at around 0.5 mm/yr over 100 km. Even with the biasing from

residual troposphere, we expect residual long-wavelength noise from DEM error to be well below residuals from other sources after correction.

Decorrelation is an additional source of noise in InSAR measurements (Zebker & Villasenor, 1992). Changes in the surface properties can alter the interferometric phase, which causes noise in the estimates of the deformation time series and thus greater errors in the calculated secular velocity. We do not expect this noise source to vary systematically over long distances, so it is unlikely to create significant biases in the long-wavelength velocity field. The impact of decorrelation can be seen by comparing tracks from Australia and the Makran. Lower coherence of the Australian interferograms compared with the Makran causes the Australian track velocity profiles to have a larger scatter in the velocities (Figure 1.4).

Another source of noise comes from the closure phase introduced as a result of multilooking combined with phase-changing physical processes on the Earth's surface (Ansari et al., 2021; Zheng et al., 2022). While this can bias velocity estimates, we would not expect to see systematic spatial biases in the velocity over hundreds of kilometers, so we do not expect this signal to contribute substantially to the long-wavelength velocity field.

Even when a term does not contribute to the long-wavelength velocity field (say, varying on a scale of well under 100 km), it can still bias the estimate of the across track ramp, with the bias being more pronounced for shorter tracks. It is possible that this explains the larger residual across track ramp for track 46 (Australia). This track is the shortest in our data set, and also has the largest across track residual after plate motion correction, at 1.5 mm/yr/track. For longer tracks, the bias on the ramp estimation by short-wavelength terms is reduced. This reduction in bias is similar to longer time series having a secular velocity that is less biased by periodic terms.

The expected amplitude and spatial wavelength of the residuals discussed above, as well as the consistency between multiple tracks, including ascending and descending tracks covering the same area, gives additional credence to the claim that the long-wavelength range-dependent ramps that we observe in our data after applying corrections are primarily due to plate motion. Plate motion and ionosphere are the two signals that have secular trends over several years, whereas signals such as the troposphere and SET are dominated by seasonal variations. It is this temporal variation that explains why the ionosphere and plate motion are dominant contributors to the long-wavelength velocities in our multi-year time series, while the large amplitude of troposphere variation results in it also contributing substantially.

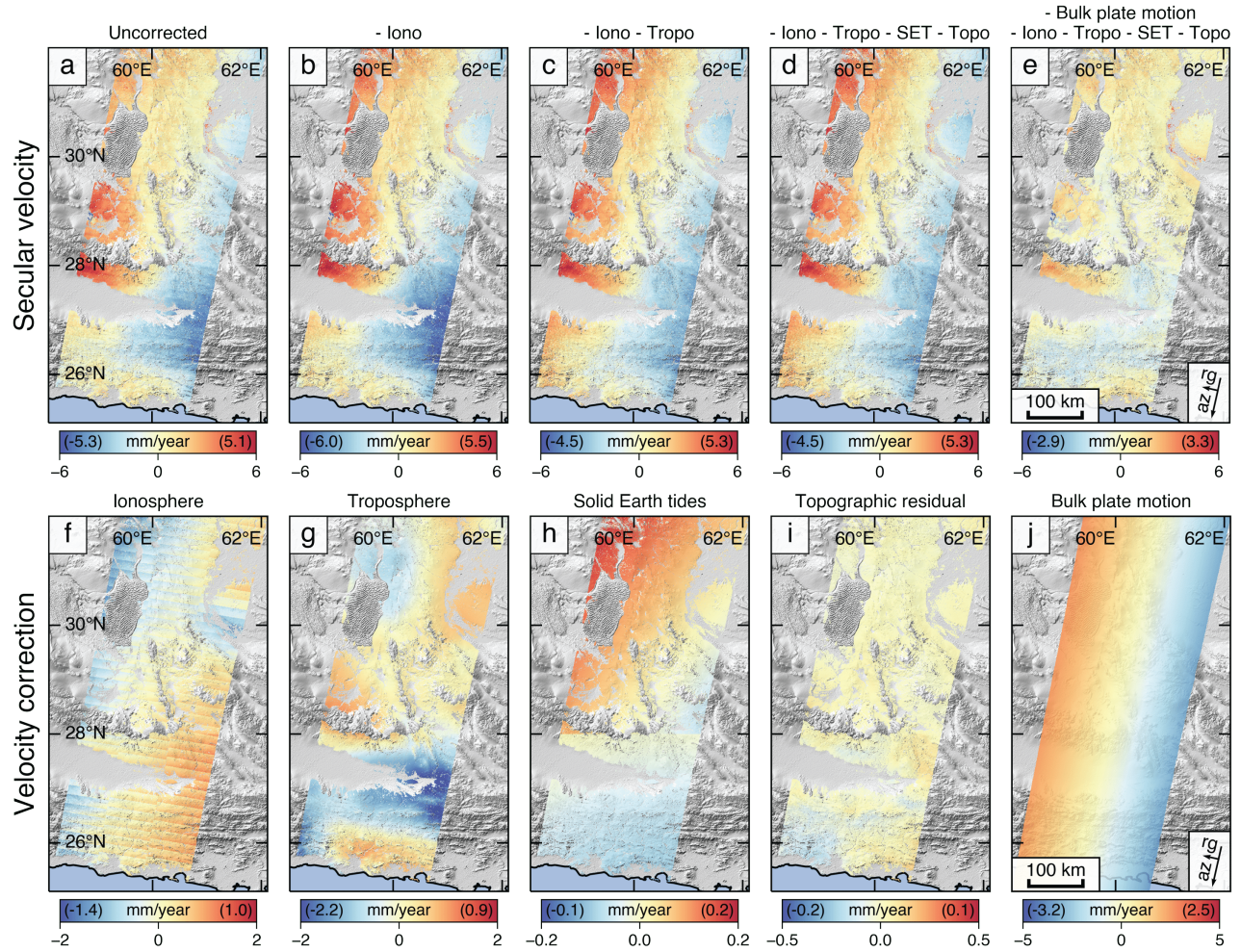


Figure A.1: Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 20 (DSC) over the Makran subduction zone. Panels are as described in Figure 1.2. Note the substantial difference in the long-wavelength ionospheric correction between ASC (Figure 1.2(f)) and DSC tracks ((f)). The burst discontinuities due to the ionosphere can be clearly seen in (f) (Liang et al., 2019).

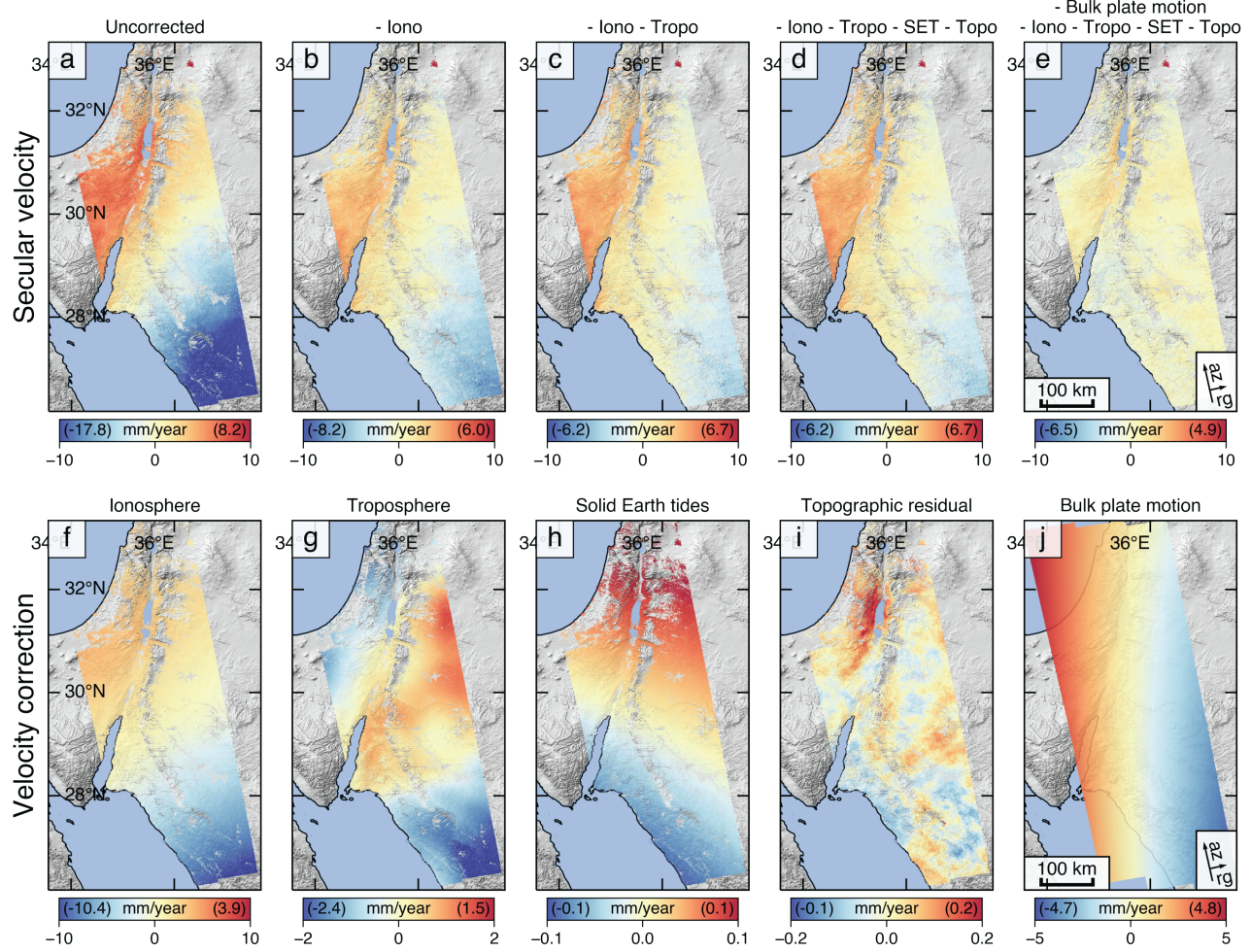


Figure A.2: Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 87 (ASC) over the Gulf of Aqaba. Panels are as described in Figure 1.2.

A.4 Further Details on the Signal of Vertical and Horizontal Plate Motions in InSAR Measurements

In Figure A.6, we show estimates of the scale of the velocity ramp in the satellite range direction for given vertical or horizontal plate translation. The values are calculated by subtracting the LOS projection of plate motion in the satellite far range from the projection in the satellite near range (see Figure 1, but note that the ramps in Figure 1 are plotted as a function of distance along the ground (ground range), not distance from the satellite (slant range)). These plots can be used for an estimate of the expected InSAR-derived velocity ramps given values of the plate motion velocity. Figure A.6 shows that, for a given plate speed, the impact on the LOS velocity is largest when the plate is moving parallel to the heading angle

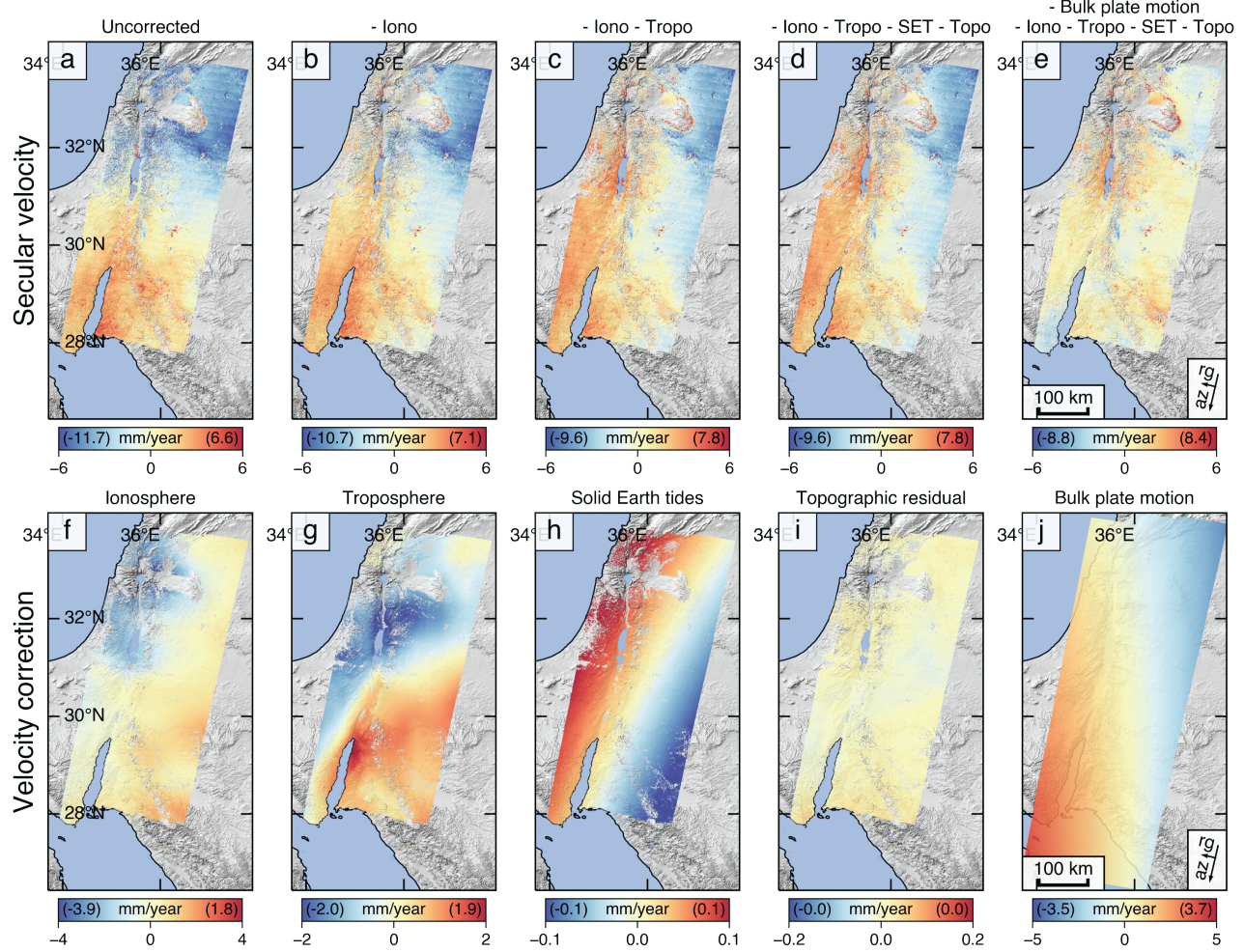


Figure A.3: Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 21 (DSC) over the Gulf of Aqaba. Panels are as described in Figure 1.2.

of the LOS vector (i.e., approximately perpendicular to the satellite direction of motion). Note that in Figure A.6 we use the line-of-sight vectors from the ascending and descending tracks over Aqaba (track 87 and track 21, respectively). While the pattern of LOS incidence angle variation across the track will be very similar for all tracks acquired in interferometric wide-swath mode, the satellite orbit heading angle (and therefore the LOS heading angle) will vary with latitude due to the near-polar orbit of the Sentinel-1 satellites. The LOS heading angle also has a small variation across the track, which can be seen from the LOS geometry files (not shown here). Users should therefore examine the true LOS vector for their study area, and only use Figure A.6 for a quick estimate of the ramps they could expect in their study regions.

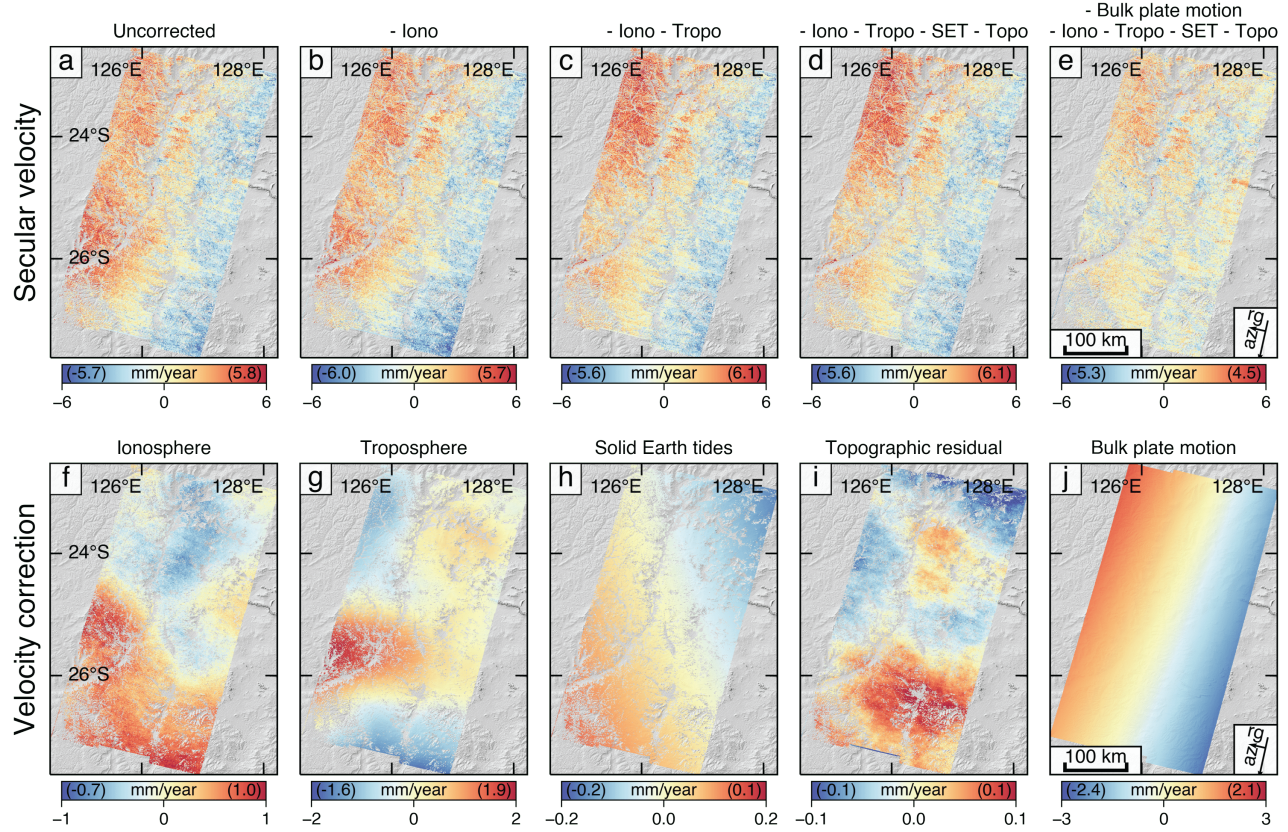


Figure A.4: Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 46 (DSC) over western Australia. Panels are as described in Figure 1.2.

Vertical secular bulk motion will have a different signature in ASC and DSC tracks compared to horizontal secular velocity. For horizontal plate motion to the east, for example, the ASC track will have a line of sight velocity gradient that is negative with increasing range and the descending track will have the opposite, assuming a right-looking satellite such as Sentinel-1 (Figure A.6). When translated into geographic coordinates, this will result in both ASC and DSC tracks having negative velocity gradients to the east (Figure 1). In contrast, a constant, positive, vertical velocity will appear as negative LOS velocity gradients in range for ASC and DSC, which translates to a negative LOS velocity gradient to the east in the ASC track and a positive gradient to the east in the DSC track. The gradients would be reversed for negative vertical velocities. This effect is illustrated in Figure 1(c), and we show examples of the ramp sizes for varying horizontal and vertical plate motion in Figure A.6.

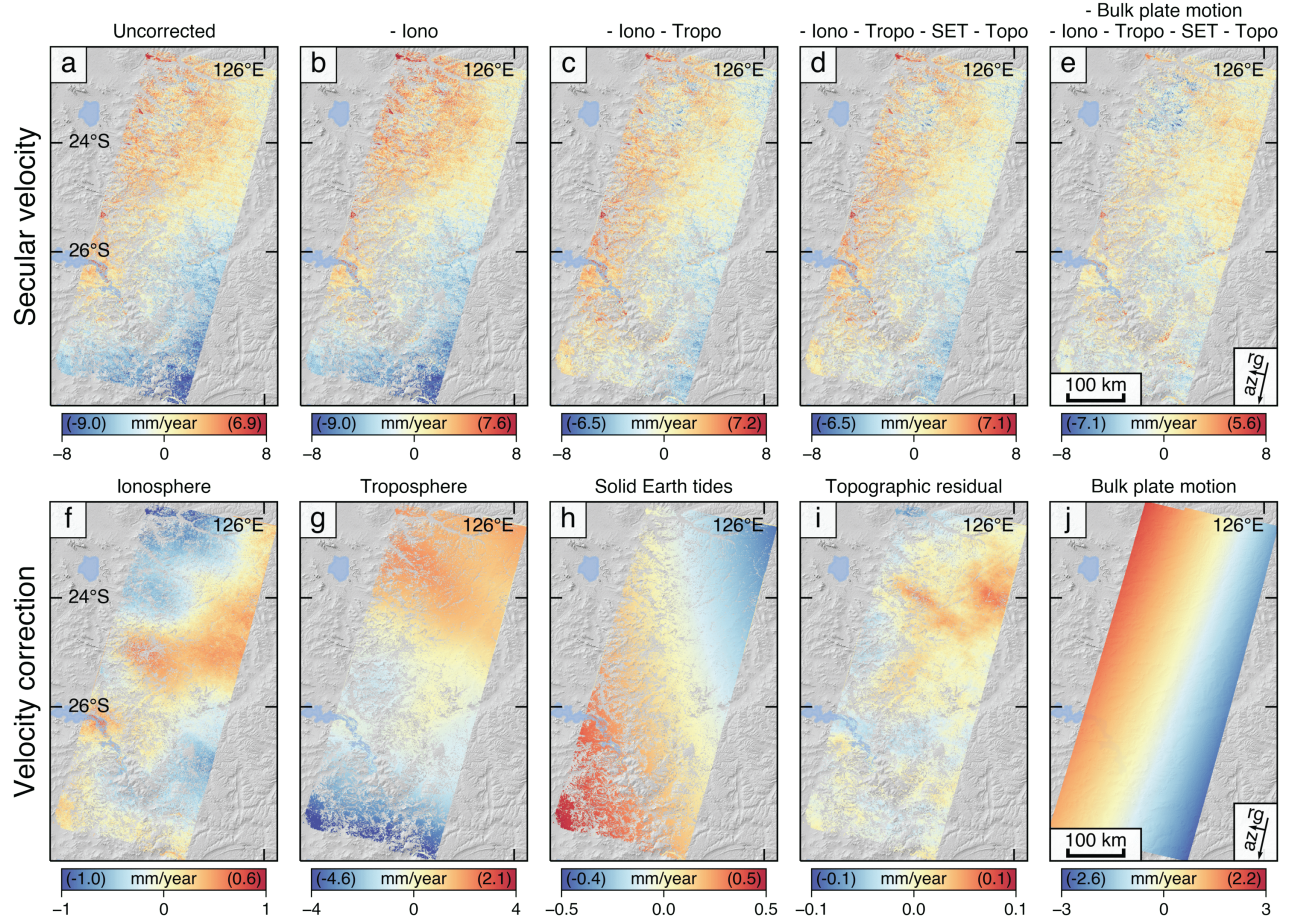


Figure A.5: Cumulative impact of corrections on the InSAR-derived velocity field for Sentinel-1, track 119 (DSC) over western Australia. Panels are as described in Figure 1.2.

Altamimi et al. (2017) and Altamimi et al. (2012) have noted that plate motion models may be contaminated by the inclusion of stations undergoing glacial isostatic adjustment (GIA). Therefore it may not be appropriate to investigate GIA by subtracting a plate motion model from a velocity field, and a more careful investigation of the local velocity field might be warranted.

As suggested in the Conclusion, the signal of plate motion in InSAR data could be used to better constrain plate motion models. This could be limited to constraining horizontal plate motions, but if the direction of horizontal plate motion is known, then the velocity gradients of ascending and descending tracks combined could potentially be used to constrain the amplitude of horizontal and vertical plate motions together. Such constraints could prove useful for improving plate models, particularly where GNSS points are sparse. However, this approach could prove challenging

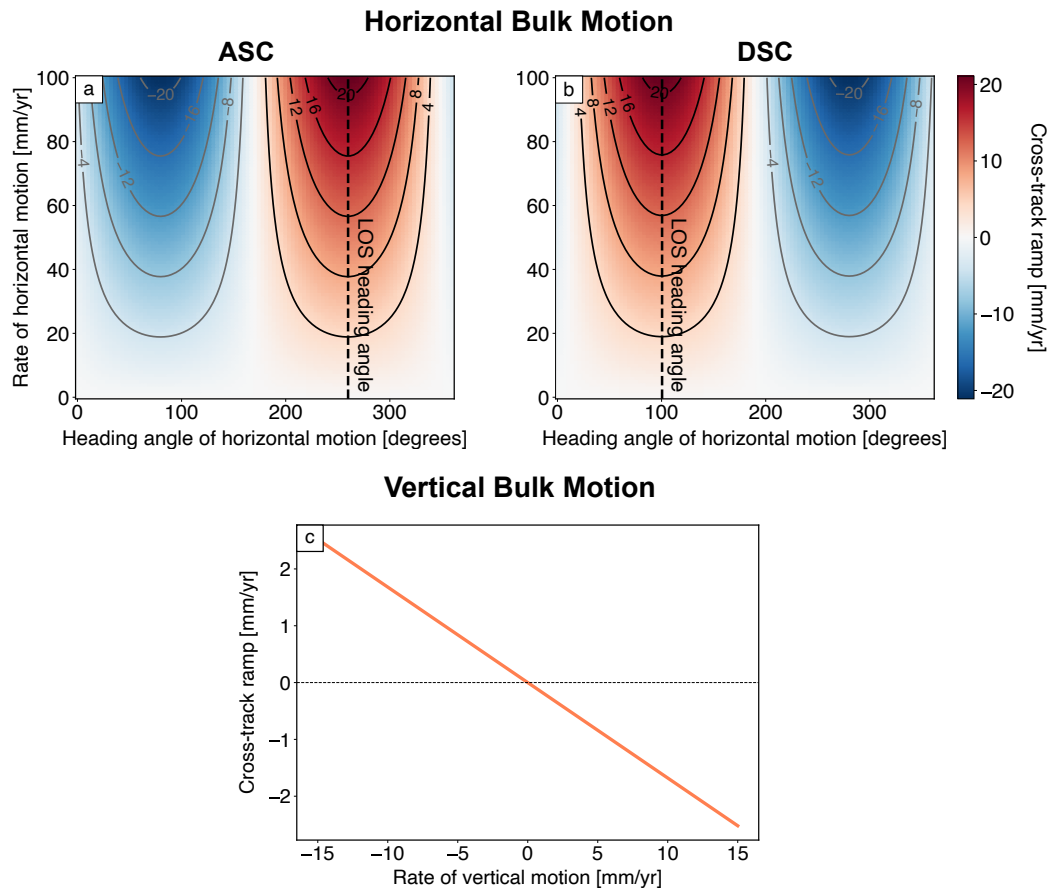


Figure A.6: Size of velocity ramp in range for the 250 km wide Sentinel-1 swath for varying horizontal and vertical bulk motion. These calculations assume a constant plate velocity (i.e., no rotation), ignore the slight curvature in the across-track ramps (Figure 1), and take the LOS geometry from the ASC and DSC Sentinel-1 tracks over Aqaba (87 and 21, respectively). The LOS vector points from the ground to the satellite. Heading angles are measured clockwise from north. “LOS heading angle” is the heading angle of the horizontal projection of the LOS vector. “Heading angle of horizontal motion” is the heading angle of the bulk plate motion. A positive cross-track ramp means that LOS velocity increases with increasing satellite range. **(a)** Size of velocity ramp for horizontal motion in an ASC track. **(b)** Size of velocity ramp for horizontal motion in a DSC track. **(c)** Size of velocity ramp for vertical velocity motion. For vertical motion, the ramp has the same variation with range for ASC and DSC tracks.

if there were substantial contributions from noise sources and tectonic deformation, and would likely require long InSAR time series to reduce biasing of the velocity by non-secular terms, and long tracks to avoid biasing by short-wavelength velocity residuals.

A.5 The Impact of Plate Motion on 3D Velocity Fields Derived from InSAR

While individual InSAR tracks only provide one line-of-sight for each point on the ground, multiple tracks with different LOS angles can be combined to constrain the 3D velocity field (Fialko et al., 2001; Wright et al., 2004). When an ascending and descending track covering the same region are available, these can be used to derive the horizontal and vertical velocities by assuming the direction of the horizontal velocity, and choosing the same reference point for both of these tracks, which is implicitly assumed to be stable in the reference frame of the satellite (e.g., Wright et al. (2004)). By assuming that this point is stable, we further assume that the ascending and descending track velocities can be explained purely by motions relative to this point, which means our solution cannot contain any overall translation. The assumption of a constant horizontal velocity direction also prohibits any rotational component in our inferred velocities. If there is in fact an overall translation or rotation of our observed region, this could lead to a biasing of our calculated horizontal and vertical velocities.

In Figure A.7, we illustrate how plate motion biases our estimation of the 3D velocity field. We use the overlapping region of tracks 86 and 20 over the Makran subduction zone, and first apply the corrections shown in Figures 2 and A.1, (f)-(i), to each tracks 86 and 20, respectively. Using these two lines-of-sight we then calculate the horizontal and vertical velocities, assuming that horizontal velocity is purely due east. When we perform this calculation before plate motion correction, the east and vertical components contain long-wavelength velocity ramps (Figure A.7 (a)-(b)). The vertical velocity has a ramp of around 3 mm/yr/100 km in the east-west direction, and the eastward velocity has a ramp of approximately 1.5 mm/yr/100 km in the north-south direction. These velocity ramps are primarily a bias from plate motion combined with the assumption that the reference point is fixed.

After removing the plate motion contribution to the ascending and descending tracks and repeating the calculation, the long-wavelength velocity ramps in the calculated horizontal and vertical maps are substantially reduced (Figure A.7 (c)-(d)). There

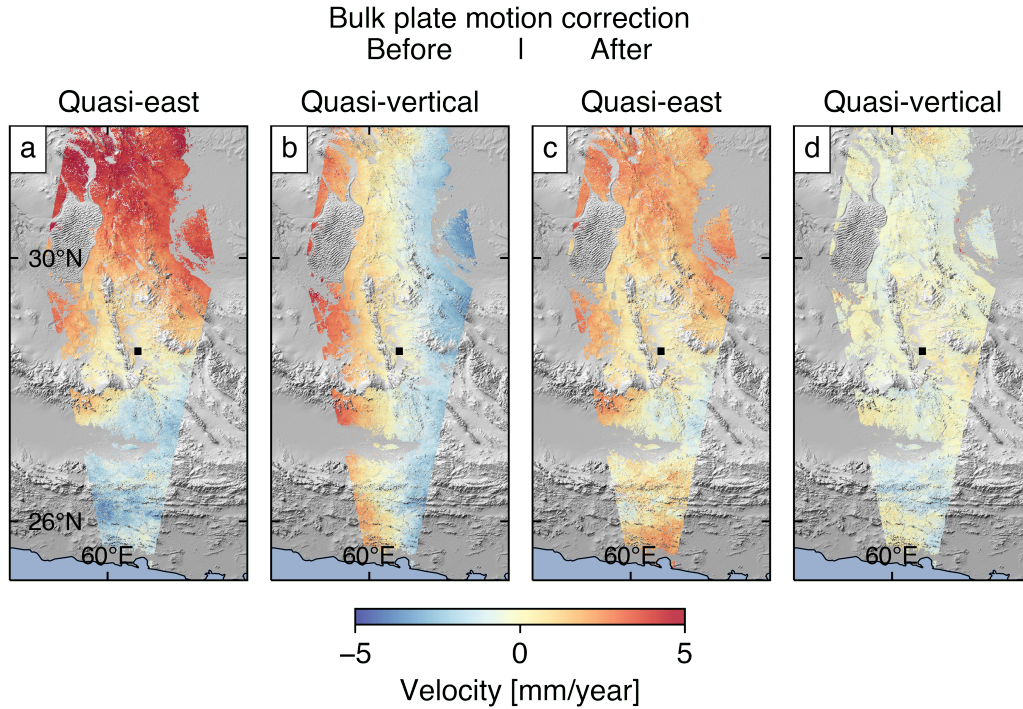


Figure A.7: Impact of accounting for plate motion on the calculation of horizontal and vertical velocities from ASC and DSC tracks. The calculation is performed for the overlapping region between tracks 86 (ASC) and 20 (DSC) over the Makran subduction zone, with horizontal motion assumed to be purely east-west. The black square represents the reference point, which is assumed to be stationary with respect to the satellite. **(a)** Calculated eastward motion before plate motion correction, but after other corrections have been applied. Note the apparent north-south velocity gradient. **(b)** Calculated vertical motion before plate motion correction. Note the apparent east-west velocity gradient. **(c)**, **(d)** Same as (a) and (b), respectively, but with the plate motion correction applied to each track before calculating the horizontal and vertical terms. Long wavelength velocity gradients have been substantially reduced.

may be some remaining bias from the assumption that horizontal motion is purely east-west.

The bias from plate motion would not be present if the InSAR observations had had a ramp fitted and removed before the calculation (as is often done), as this would remove the plate motion velocity ramp. The bias would also be less obvious if the area studied was smaller, or the tectonic signal was larger. The results in Figure A.7 emphasize the importance of accounting for the reference frame of our observations before combining separate tracks.

Appendix B

SUPPLEMENTARY MATERIALS FOR CHAPTER 3: DEEP LEARNING-BASED DAMAGE MAPPING WITH INSAR COHERENCE TIME SERIES

B.1 Phase variance and the logit transform

The Cramer-Rao bound on the variance of the interferometric phase, σ_{phase}^2 , can be written (Rodriguez & Martin, 1992):

$$\sigma_{phase}^2 = \frac{1}{2N_L} \frac{1 - \gamma^2}{\gamma^2}, \quad (\text{B.1})$$

where γ is the coherence between two SLCs and N_L is the number of SLC pixels in the chip used to estimate the coherence. The phase variance asymptotically approaches this limit as the number of looks increases, with the limit being a good approximation for $N_L > 4$. In this work our chip size gives us $N_L = 75$.

Taking the logarithm of the phase variance gives us:

$$\log_e(\sigma_{phase}^2) = \log_e\left(\frac{1 - \gamma^2}{\gamma^2}\right) - \log_e(2N_L). \quad (\text{B.2})$$

We also have the logit transform of coherence squared (Eq. 3.9), which is:

$$S^{-1}(\gamma^2) = \log_e\left(\frac{\gamma^2}{1 - \gamma^2}\right). \quad (\text{B.3})$$

We therefore have the relationship:

$$S^{-1}(\gamma^2) = -\log_e(\sigma^2) - \log_e(2N_L), \quad (\text{B.4})$$

meaning the logit transform of the coherence squared only differs from the logarithm of the phase variance by an additive constant and sign.

B.2 RNN Model and Training Details

The basis of RNNs are *artificial neural networks*. These are layers of matrix-vector multiplications, interspersed with non-linear *activation functions*, that transform input data into a desired output. The matrix values (known as *weights*) are varied, or *trained*, to get the desired mapping between input and output data. For a quick pedagogic introduction to artificial neural networks, we refer the reader to

3Blue1Brown (2017), more detail on artificial neural networks and RNNs can be found in Goodfellow et al. (2016).

So-called *vanilla* (basic) RNNs have trouble learning long-term dependencies, which in our case means learning how to make forecasts when the desired output at a given step depends on data received many steps in the past (Bengio et al., 1994). Long short-term memory RNNs (LSTMs) were created to solve the problem of long-term dependencies (Hochreiter & Schmidhuber, 1997). LSTMs have a series of *gates*—learned weights combined with non-linear activation functions—that allow the network to decide what information to add, retain, and forget when each new observation is fed in, in such a way as to better store information over many time steps compared to a vanilla RNN. For a pedagogic introduction to LSTMs, we refer the reader to Olah (2015). The weights are learned during training, so that the network is able to decide what information is most useful for a given set of time series. A single set of weights are learned for all the time series that are given to the network.

Gated recurrent units (GRUs) are a gating mechanism for RNNs, similar to LSTMs but with fewer parameters and thus less computationally expensive to train. GRUs have been found to have similar performance to LSTMs on a range of tasks (Cho et al., 2014), and are what we use in this work.

As an example for our data, one could imagine trying to forecast the coherence for a region with dry summers, and wet winters. During the dry summer, the ground surface may be relatively undisturbed, leading to high InSAR coherence. However, rain and snow during the wet winters could often disturb the Earth's surface, leading to large drops in coherence. If we had at least one year of coherence time series data, and were trying to forecast the coherence at the beginning of the summer, the recent coherence measurements from the winter months may not be the most useful data to make this forecast. Instead, we would want the network to make a forecast based on similar sequences that it had seen during previous summer periods.

By training the network on a large number of coherence time series containing winter and summer behavior, the gating mechanisms should learn the relevant information to use when forecasting during summer and winter sequences. Note that this is not simply learning a periodicity in the signal. If there was an unusually wet summer or dry winter, the network should be able to adjust its forecast based on new data and its training.

The following equations describe f_ϕ (Eq. 3.2) for the GRU that we use in this work:

$$z_t = S(W_z x_t + U_z h_{t-1} + b_z) \quad (\text{B.5})$$

$$r_t = S(W_r x_t + U_r h_{t-1} + b_r) \quad (\text{B.6})$$

$$\bar{h}_t = \tanh(W_h x_t + U_h (r_t \circ h_{t-1}) + b_h) \quad (\text{B.7})$$

$$h_t = (1 - z_t) \circ h_{t-1} + z_t \circ \bar{h}_t, \quad (\text{B.8})$$

where S is the sigmoid activation function, $\circ$ is the Hadamard (element-wise) product, z_t is the update gate vector, r_t is the reset gate vector, $\bar{h}_t$ is the candidate hidden state, and $\phi = [W, U, b]$ are the learnable parameters of f . We use a hidden state of size 256, with the initial state h_0 set to all zeros. For g_ψ (Eq. 3.5), we use a fully connected, feed-forward neural network to map the RNN hidden state to parameters of a Gaussian distribution. This neural network has 3-layers, with Rectified Linear Unit (ReLU) activation functions and hidden layer sizes of 128. The resulting output is finally passed through two separate linear layers to obtain the mean and natural logarithm of the variance (we output the logarithm for improved numerical stability). Figure B.1 shows how the different components are connected. In total, our model has 265,090 learnable parameters. Hyperparameters of the architecture are chosen according to rules of thumb, and at this stage we have not performed a cross-validation or hyperparameter search. This search can be done to further improve the forecast made by the RNN, for example by varying the size of the hidden state and the number of layers in the fully connected neural network.

We randomly select 80% of sequences in the training set $\mathcal{D}_t$ to train our model ($\mathcal{D}_{t,t}$) and use the remaining 20% as validation sequences for model selection ($\mathcal{D}_{t,v}$). We emphasize that the model does not see any of the validation sequences and only sees pre-event information during training.

We train for 20 epochs with a batch size of 256 using the Adam optimizer with a learning rate of 0.0005 (Kingma & Ba, 2014). Each epoch sees each training sequence once and has a total run time of several minutes on a Tesla P100 GPU (depending on the number and length of training sequences, which varies by region). The trained network is then applied to the forecasting data set $\mathcal{D}_f$ to generate co-event coherence forecasts at each time step and location. Details of how $\mathcal{D}_t$ and $\mathcal{D}_f$ are calculated are included in Section B.3.

Non-determinism can be introduced in the RNN training, for example due to the implementation of GPU algorithms and the random selection of batches during

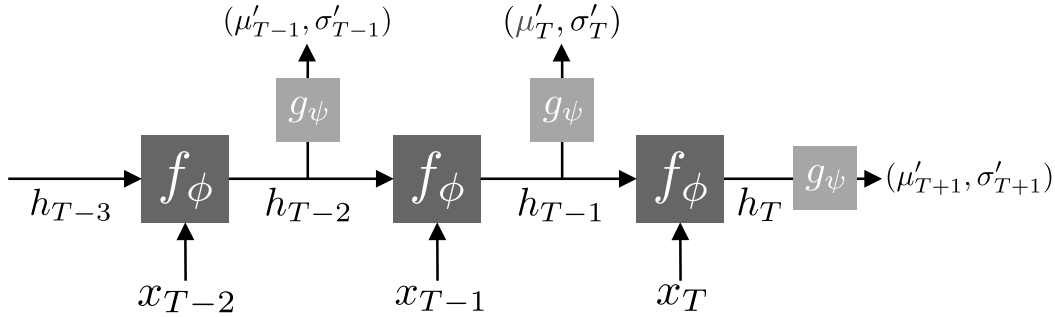


Figure B.1: Graphical model for our recurrent neural network for a series of time steps. We use gated recurrent units for f_ϕ and a feed-forward neural network for g_ψ . For each pixel the hidden state (h) summarizes the coherence information (x) up to that point, and is used to forecast the mean (μ') and standard deviation (σ') of the logit-transformed squared coherence. We initialize the hidden state h_0 as all zeros. The final hidden state, h_T , is used to forecast the co-event coherence distribution parameters $(\mu'_{T+1}, \sigma'_{T+1})$.

training. From limited trials on the Amatrice data set we find that repeated training with the same data and training parameters can lead to small differences in the PR AUC (less than 0.05), but this variability does not affect our overall conclusions. It is possible that the data leakage issue, discussed in Section 3.6, is leading to overfitting, and thus to a larger variance in the co-event forecast for repeated re-training, thus causing a larger variation in the PR AUC. We therefore caution that the specific values indicated in this paper shouldn't be taken as representative of the values likely to be obtained in all circumstances, and reiterate that further testing in a wider variety of cases is necessary.

An implementation of our deep learning model can be found on GitHub: <https://github.com/olliestephenson/dpm-rnn-public>.

B.3 Study Areas and Data

A summary of the SAR data used for each study area is presented in Table B.1, and more details are given in this section. In all cases we use data starting in October 2014 (when the Sentinel-1A satellite began acquiring data) until the first post-event SAR acquisition. We start with Level-1 Single Look Complex (SLC) images acquired in interferometric wideswath mode (European Space Agency, n.d.), then, using the InSAR Scientific Computing Environment (ISCE) (Rosen et al., 2012), we create a coregistered stack of SLC images, which are corrected for the flat Earth and topographic phase contributions. For each region, we calculate $N - 1$ sequential

coherence images from N SAR acquisitions. The post-event and final pre-event SAR acquisitions are used to generate a co-event coherence image, which is not seen during RNN training but is used to calculate the size of the coherence anomaly. While the time between SAR acquisitions varied for each region, in all three cases the first post-event image was taken six days after the final pre-event image. Note that we keep the data in radar coordinates for the coherence calculation and do not interpolate or resample once the SLC stack is coregistered.

To calculate the coherence time series in $\mathcal{D}_t$, we set the stride (distance between adjacent center SLC pixels of the chip, or averaging window) to be equal to the chip width in each direction so that each SLC pixel is only used to estimate coherence in one chip (Fig. B.2(a)). We use data spanning a wide area surrounding the specific region we are searching for damage.

For calculating $\mathcal{D}_f$, we only use SAR data in the area of interest (i.e. the damaged region we would like to map), but set the stride of the coherence chip to one, giving us a larger number of coherence pixels in the same area. Coherence pixels are therefore spaced at approximately 3 m and 14 m in range and azimuth, respectively; however, each coherence chip contains information from a 50 m by 70 m (15 by 5 SLC pixel) chip (Fig. B.2(b)). Note that the exact size of the chip on the ground depends on the local topography. The stride of the chip means that adjacent coherence chips in $\mathcal{D}_f$ will share a large number of SLC pixels, and thus have a high degree of correlation in their time series.

Note that as the area of interest lies within the training region a small fraction of the coherence time series in $\mathcal{D}_t$ will also be in $\mathcal{D}_f$, leading to a small amount of data leakage (i.e. the model is trained on some of the time series that it is trying to forecast). This could lead to some overconfidence in the coherence forecast, reducing the quality of the damage map. However, model selection is done using the validation set ($\mathcal{D}_{t,v}$), i.e. the 20% of $\mathcal{D}_t$ that is not used for training, and the data leakage is a tiny fraction of the total number of time series used for training.

August 24, 2016 M_w 6.2 central Italy earthquake

Our first data set is from the M_w 6.2 August 24, 2016 central Italy earthquake. At least 299 people were killed, with more than 400 injured and major damage done to the Italian town of Amatrice (United States Geological Survey, 2016). For this event, the Copernicus Emergency Management Service produced a damage map by manually reviewing high resolution pre- and post-event optical satellite imagery.

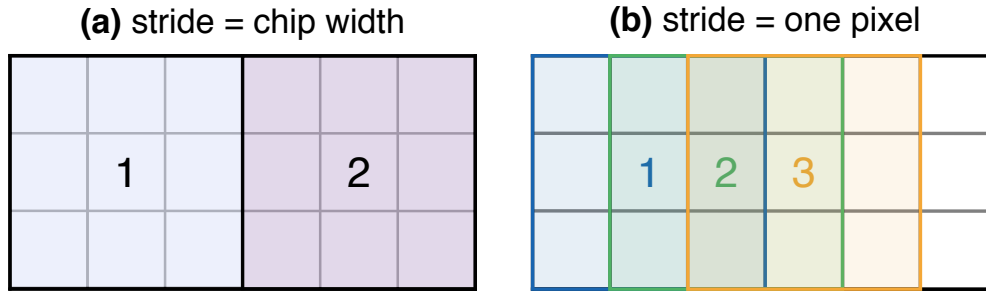


Figure B.2: Schematic of the different coherence calculation methods for (a) $\mathcal{D}_t$, and (b) $\mathcal{D}_f$. The small squares represent SLC pixels and the large, numbered, squares represent coherence chips, here represented as three by three SLC pixels for illustration purposes.

Every building in Amatrice had its geographic footprint determined and each footprint was classified into one of five damage levels, from “unaffected” to “completely destroyed.” This data was made available online (Copernicus Emergency Management Service, 2016). This data set was constructed via visual inspection of high resolution optical imagery taken before and after damage to the town. In our work we further simplify the data set by putting damage levels “not affected” and “negligible to slight damage” into an “undamaged” class, and “moderately damaged,” “highly damaged” and “completely destroyed” into a “damaged” class. The resulting damage map is plotted in Fig. B.3. As every building within the centre of the town was assessed, we can use this as ground truth with which to quantitatively compare the RNN and CCD methods.

The pre-event SAR repeat time was irregular, mostly 12 or 24 days between acquisitions, which creates a source of noise for our RNN method due to variable temporal decorrelation. The data in the training region, $\mathcal{D}_t$, covers the area around and east of Amatrice.

November 12, 2017 M_w 7.3 Iran-Iraq earthquake

The city of Sarpol-e-Zahab lies on the Iran-Iraq border and was hit by a M_w 7.3 earthquake on November 12, 2017, causing substantial damage, over 7000 injured and at least 630 fatalities (United States Geological Survey, 2017). The United Nations Institute for Training and Research (UNITAR) produced an earthquake damage map for Sarpol-e-Zahab in the aftermath of the earthquake. By manually reviewing high resolution post-event optical satellite data, UNITAR mapped the

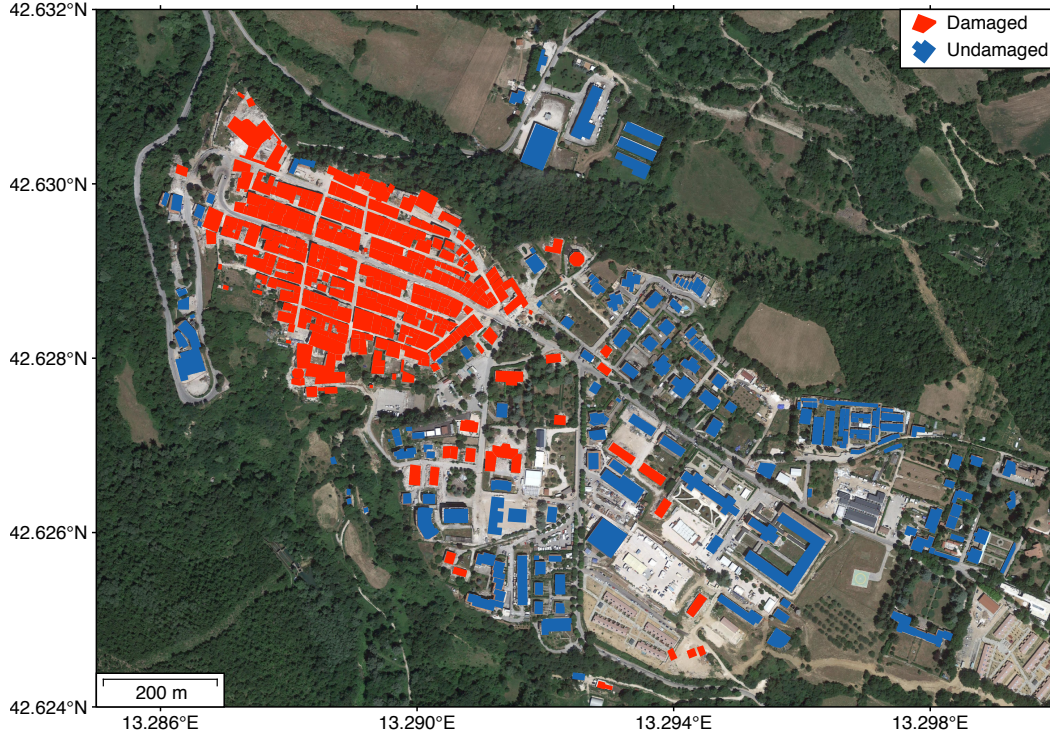


Figure B.3: Ground truth data used for quantifying damage classification performance for the town of Amatrice, damaged during the August 24, 2016 M_w 6.2 central Italy Earthquake. Building polygons and damage levels are supplied by the Copernicus Emergency Management Service (Copernicus Emergency Management Service, 2016), then simplified as described in Section B.3. Note that the optical imagery in this figure was taken a little under a year after the earthquake, and structures have been built that do not appear in the damage assessment. Optical imagery for the figure is from Google, taken July 6th 2017.

location of 683 “potentially damaged” buildings (United Nations Institute for Training and Research, 2017) (Fig. B.4). However, the data does not contain damage levels or building footprints for every building, meaning we are unable to constrain which buildings are undamaged and so can’t perform the same quantification as for Amatrice.

Before the earthquake, the satellite repeat time was 12 days (with six 24 day intervals). The post-event acquisition was six days after the final pre-event image. The data in the training region, $\mathcal{D}_t$, spans the city and the region to the north and west of Sarpol-e-Zahab.

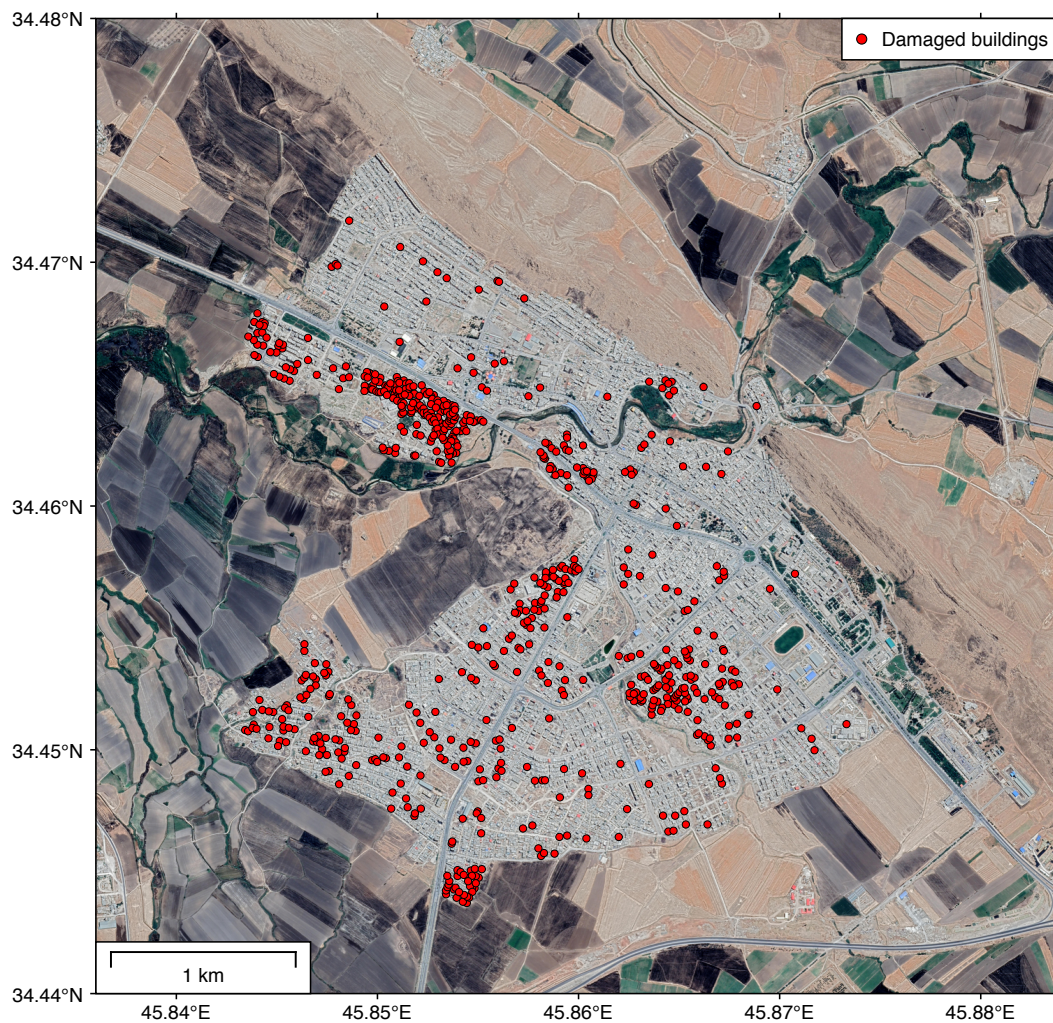


Figure B.4: Location of 683 potentially damaged buildings manually mapped from optical satellite imagery by the United Nations Institute for Training and Research (United Nations Institute for Training and Research, 2017) for the town of Sarpol-e-Zahab, damaged during the November 12, 2017 M_w 7.3 Iran-Iraq earthquake. Optical imagery for the figure is from Google, CNES/Airbus, taken July 27th 2020.

Table B.1: Summary of SAR data used for each study area. N_{aqn} : number of pre-event acquisitions used, $N_{t,f}$: number of pixels in the training ($\mathcal{D}_t$) and forecasting ($\mathcal{D}_f$) sets, respectively, $E_{t,f}$: spatial extent of the training and forecasting regions, Track: Sentinel-1 orbital track and direction of flight.

Event	Date (MM/DD/YY)	N_{aqn}	N_t	E_t (km $\times$ km)	N_f	E_f (km $\times$ km)	Track
Central Italy	08/24/16	48	3,396,828	140×70	70,000	2.4×1.4	Descending 22
Iran-Iraq	11/12/17	89	11,056,389	210×210	750,000	5.2×7.0	Ascending 174
Ridgecrest	07/04/19	97	2,860,000	60×60	2,860,000	60×60	Ascending 64

July 2019 M_w 6.4 and M_w 7.1 Ridgecrest, California, USA earthquakes

Numerous groups conducted mapping of surface ruptures, liquefaction and building damage in the aftermath of the Ridgecrest earthquakes (Brandenberg et al., 2019; Hough et al., 2020; Kendrick et al., 2019; Ponti et al., 2020; Zimmaro et al., 2020), allowing for a qualitative comparison with our damage proxy map (Fig. 3.7). Our area of interest is significantly larger than the previous two, and to cover the entire rupture we need to use the whole of $\mathcal{D}_t$ in $\mathcal{D}_f$. We use strides of 5 and 15 (the same as for $\mathcal{D}_t$) as we are more interested in larger scale signals. The fact that the training and forecasting regions are the same means that the model will have been trained on 80% of $\mathcal{D}_f$; however, as usual the optimum model is chosen from the performance on the 20% of $\mathcal{D}_t$ ($\mathcal{D}_{t,v}$) that is not seen during training.

The pre-event SAR acquisition interval in the Ridgecrest region was more variable than data for the other two areas, starting out at a modal value of 24 days in October 2014, then reducing to 12 days and finally 6 day intervals from March 2019 until the event. Again, the post-event acquisition was six days after the final pre-event image. While the acquisition interval was overall more variable than the other two regions, the several months of 6 day intervals before the event likely improves the quality of the RNN forecast.

B.4 Training with a Variable Amount of Pre-Event Data

In order to explore how the quality of the RNN damage map depends on the amount of data, we repeat the training and forecasting for the Amatrice data set, decreasing the temporal span of the pre-event data used and calculating the Precision-Recall Area Under Curve (PR AUC) for each trained model. We use the Amatrice data set for this exploration, as this is the only location where we have comprehensive ground truth and so are able to obtain precision-recall curves. Training and forecasting data are generated as described in Section B.3, but then only coherence images that use data acquired on or after specific dates are used in training.

These results are presented in Figure B.5(a), along with a comparison to the CCD PR AUC as a baseline. The CCD method uses a single pre-event coherence image along with the co-event image. When we use all available data, we obtain the result previously presented in Figure 3.4 (PR AUC=0.7). As we move the cut-off date later in time, we observe a steady decrease in the PR AUC until late 2015, where there is a sudden drop to around 0.5 (well below CCD performance), followed by an overall increase up to the final point, which only uses two pre-event coherence images.

If every coherence image was adding information that could allow for a better forecast of the coherence distribution, then we would expect that more training data would improve the results. The sudden drop in PR AUC when the data cut-off is in late 2015, followed by the increase in performance as we reduce the amount of training data, conflicts with these expectations. This behavior suggests that the coherence data in the year leading up to the earthquake are potentially less representative of the distribution that we are trying to forecast, meaning that training only on this data decreases performance.

We can gain some indication of the variability of coherence through time by looking at the mean of coherence data (in logit space) used for training, presented in Figure B.5(b). These results show anomalously high coherence values in late 2015, lining up with the decrease in PR AUC. It is important to note that all training with a cut-off date earlier than late 2015 includes these anomalously high coherence values. However, it is possibly the case that as we increase the amount of data beyond a year, the fraction of the training data which is anomalous decreases and so the results improve.

The results presented in this section suggest that performance can be unreliable when training is done on less than a year of coherence data. However this conclusion is likely highly dependent on the specific attributes of the local region. For example, we would probably not observe the same effects in a dry desert area, where coherence is more stable through time. Further testing is necessary to explore how the results depend on the time span of the training data.

As noted in Section B.2, non-determinism in the RNN training leads to variability in the exact PR AUC values when training is repeated for the same data. This variability leads to some scatter in the results; however, it does not affect the overall results presented in this section.

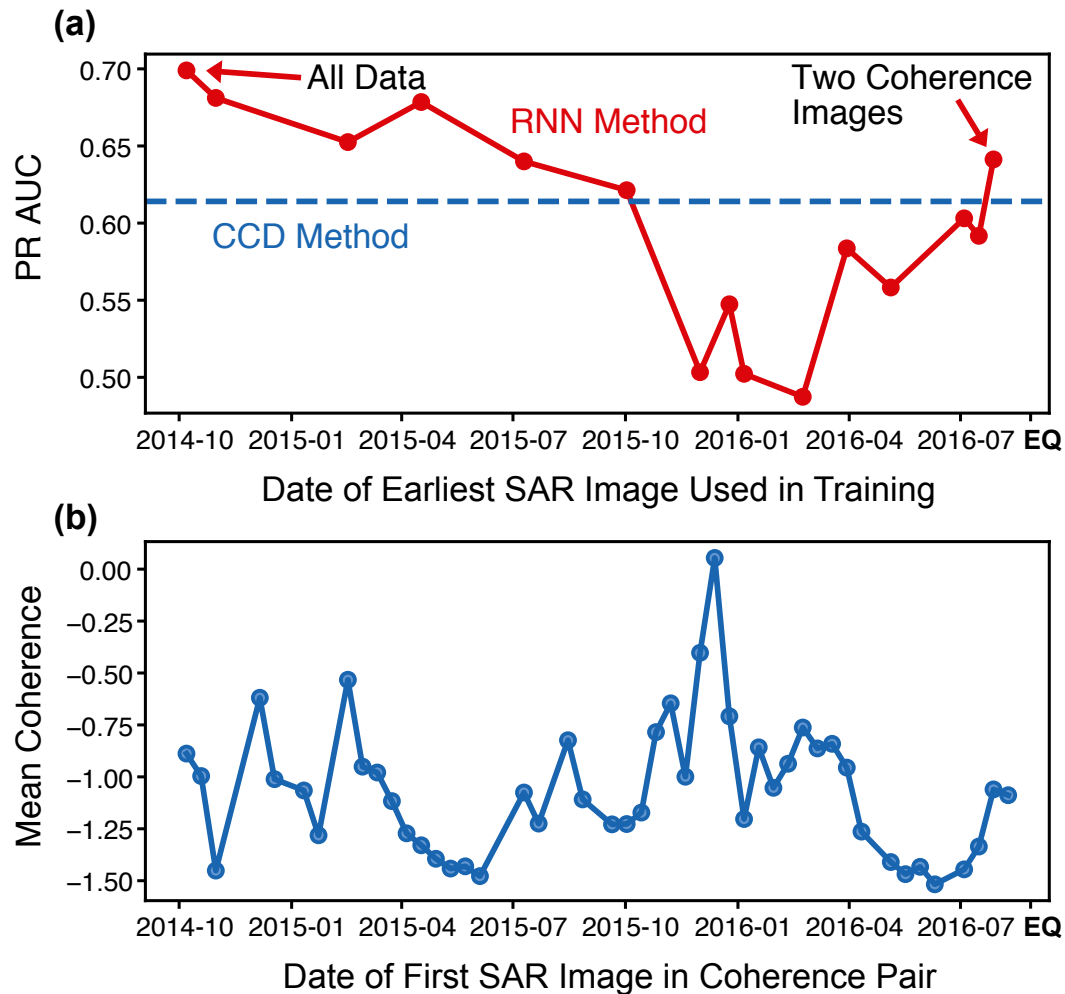


Figure B.5: Impact of varying the amount of training data. **(a)** RNN Precision-recall area under curve (PR AUC) for the Amatrice data set, with varying amounts of pre-event training and forecasting coherence data (red line). Only data acquired on or after the indicated date are used in training and forecasting. The PR AUC for the CCD method is also presented for comparison (horizontal dashed blue line). Training that uses all data is indicated at the top left. Training that only uses two pre-event coherence images is indicated on the right hand side. **(b)** Mean of the logit transform of the squared coherence against the date of the first SAR acquisition in that coherence image. Values are presented up to the final pre-event coherence image. The co-event coherence is not used in training, so its mean is not presented here. Generally SAR images in each coherence image are separated by 12 or 24 days. The date of the August 2016 central Italy earthquake is indicated by “EQ” on the right of the both plots.

*Appendix C***SOME THOUGHTS ON PRODUCTIVITY AND DOING A PHD****C.1 My Approach**

While this thesis (hopefully) contains some insights about my scientific work, it says little about how that science was done on a day-to-day basis. Tasks that take months of effort can gain barely a line of text, and there is nothing at all about the patterns of work and time management that have (and have not) worked for me. I wanted to provide a few details of how I have tried to manage my time and attention during my PhD, and the various successes and failures along the way. At the bottom, I've listed some specific tools that I use to implement the systems I describe, and given suggestions for further reading. While some life skills can only be learned through experience, I believe that time management is something that can be improved with a relatively small amount of deliberate effort. I feel it would be good if these kinds of things were discussed more frequently; it often seems to me that some people naturally arrive at systems that work for them, while many people don't realize that they might be able to get a lot more out of their days with a few tweaks.

Before going any further, I want to emphasize that everyone's approach to work can be very personal, and different methods can work well for different people. Your mileage may vary greatly with the suggestions below. More important than any particular approach, in my opinion, is regularly reviewing what you're doing, seeing what's working and what isn't, and gradually improving your systems. However, many people will not have the flexibility that I have had, perhaps because of family or teaching commitments. Furthermore, my goal here is not to say that it is your fault if you're struggling with time management. Academia can be a hostile environment for any kind of work-life balance, and while most of this appendix is focused on personal time management, working to change this broader environment is a vital activity.

A key lesson is to remember that a PhD is a marathon, not a sprint. I think a good approach is to set a minimum goal of focused, productive work a week, track your work to ensure you are meeting the goal, and then give yourself actual regular time off. By setting and tracking your goals over the long term you can see if you are regularly falling above or below where you want to be, and make adjustments

accordingly. An approach like this helps to avoid the incessant guilt of never having done enough, which, in my experience, leads to cycles of overwork and burnout.

What counts as productive work? Here I am defining it very restrictively. Whenever I take a break, check my email, chat to someone, get distracted by social media, etc, I stop the clock. I will generally not count casual academic conversations, seminars where I am not 100% engaged, and skimming through abstracts. This definition does mean that I end up not counting things that are valuable, and necessary, components of my work, but it allows me to concentrate on the component where I make the most progress—the extended periods of time where I am deeply focused on a specific problem and working to solve it.¹ I also want to emphasize that I am not saying that anything outside of this definition is necessarily “unproductive”.

Once we have some definition of “productive”, there is then the question of what minimum goal to set. In theory, graduate students at Caltech are entitled to (a laughably small) 21 days off per year. Assuming a (very ambitious) 40 hours of productive work per week, for around 48 weeks a year, for five and a half years, that translates into 10560 hours over the PhD. I would be surprised if many people get close to this. I have found that I was only able to sustain over 40 hours of genuinely productive work a week by making substantial sacrifices in many other areas of my life.² Actually getting above 40 hours of productive work required me to spend a lot more than 40 hours actually at my desk, as I fought with distractions and general fatigue. Importantly, if I had not been tracking the amount of work I was doing, I would probably have estimated that I was working 60 or more hours, giving me a false impression of how productive I was actually being.

With “productive work” defined restrictively, I believe that a target of around 25 hours a week, for 40 weeks of the year, is a reasonable and sustainable goal. This translates into 1000 hours a year, 5500 hours of work over a 5.5 year PhD, or a bit over half of the 10560 hours given above. There are a lot of other things that have to happen on top of this number (e.g., administrative tasks, teaching, community service, activism, organising, etc.). But, by setting a goal like this, it’s possible to ensure that you are making progress on your main PhD projects, as well as spending time dealing with all of the other things that you want/need to get done. By setting

¹Worth mentioning here that I consider having the freedom to do this to be an immense privilege.

²These comments apply predominantly to the kind of work that most people in Caltech’s Seismolab are doing, which is generally desk-based. For those who have substantial lab or field components to their work I think this picture could be very different, although much of what I say could still apply to components of their work that are similar in style to the kind of work I am referring to (e.g., paper writing, data analysis, etc).

a goal around 25 hours, that hopefully preserves lots of the week for staying active and engaged in both your academic community, and the wider world.³

During my PhD, I would often finish a week which had felt immensely busy, but in reality I had spent my time running between a series of minor tasks that were not actually that important, and hadn't moved my main scientific work forward. I think it is easy to feel like you are doing lots of work because you have spent many hours at your desk, but find that you're not really moving forwards, which is more likely to push you towards burnout. By being intentional, setting achievable goals, and tracking your time, you can spend less time at your desk and do more of what you actually wanted to do.

Ok, so I've set myself some work time goals, and I'm tracking them. Next is the question of how to actually spend that time. As a very rough division, I think it's worth spending something like 60% of your time on implementation for your core projects, 30% of your time on core reading, and 10% on what I'll loosely call "broader exploration".⁴ Exactly what each of these terms mean, and the distribution between them, will change between projects, and the stages of those projects.

Let me give some illustrations of what I'm getting at. For my earthquake simulations presented in Chapter 4, I had several core tasks. These included running simulations, plotting and analyzing the outputs, and working through theoretical understandings of what was going on. These tasks could often be very time consuming, involving lengthy hunts for bugs in the code parallelization or writing long scripts to simultaneously process hundreds of simulation outputs. These are the kinds of tasks that I think should have 60% of the overall time you allocate to the project. However, unlike core reading or broader exploration, these are the tasks that determined whether or not I'd have anything to show my advisor at our next meeting, and so the short-term incentives always pushed me to work more on these. I would often spend nearly all of my time on these items, at the expense of reading and broader exploration.

My simulations work was informed by a series of core papers that dealt with computational, theoretical and experimental results. I advocate for spending about

³I received some feedback that setting a specific work hour goal could be seen as "presenteeism"—just spending a certain number of hours working for appearances, rather than actually getting anything done. Remember that the goal is to get science done, not just sit at your desk for a certain number of hours. Planning what you want to get done, and reviewing what you did each week can help you discern if you're spending the time as you want to.

⁴This division just applies to the time you allocate to your main scientific projects (i.e., 25 hours/week, or whatever you choose), and doesn't include things like compulsory classes and TAing.

30% of your weekly hours on this, but I was way under this target for most of my PhD (it was generally around 10%, rising to closer to 25% over the final year). A lack of time dedicated to reading was a major mistake—it reduced my ability to go beyond the suggestions my advisor had made and come up with my own insights into the project. I suggest keeping a running list of papers that you want to read, roughly prioritized, then set aside dedicated time to spend on these. At the end of each week, look back at how much time you've spent reading, and adjust your next week based on how close you were to your goal. For particularly important and/or complicated papers, you will want to read them in detail several times, and come back to them as you gain more experience. For less significant papers you can limit yourself to the abstract, introduction and conclusion, then come back to them if they turn out to be helpful.

Beyond the core implementation and reading, there were a series of broader topics that I didn't need to know in the moment to develop the project, but could potentially prove fruitful further down the line. Examples include learning more about parallels between friction during earthquakes, landslides and glacier motion, and reading about the preservation of pseudotachylytes. The key point is to be broadening your horizons, in a way that could inform future research directions, but not at the expense of making sustained progress on your main projects. I cheerfully admit that I was terrible at this for the majority of my PhD, spending a significant amount of time on areas that were unlikely to do much for my main projects (e.g., I went to a reading group on the scientific history of climate change, took a class on sustainability, and did some side-work on glaciology, along with a wide variety of less scientific endeavours).

Related to the exploration portion of your time is the question of how to pick problems to work on. I was advised to regularly go to Seismolab coffee hour to get ideas for my work. I didn't take this up as much as I should have done, but it's clear to me that picking good problems relies on regularly interacting with a broad range of scientists, along with reading and attending seminars. Being part of these interactions can really help you put your finger on the active debates in your field. Reading papers is helpful for this, of course; however, I frequently found that I could read through a paper and still be confused about what lines of research it pointed to, whereas a quick conversation with someone could highlight ideas that I had totally missed during reading. Much like reading, making time for general scientific conversations rarely felt like a priority for me. That's why it's essential to

deliberately carve out time for these kind of interactions. They don't really fit within my very restrictive definition of "productive", but as a very loose guide I think it's worth spending at least an hour a week in wide-ranging scientific conversations with your peers, along with intense periods of interaction during conferences.

After thinking about my weekly working hours, and how to allocate them, I need to decide how to spend each day. I found that when I didn't have a clear daily plan, I would naturally gravitate towards tasks that felt some combination of easy and urgent, which generally didn't lead to significant progress on my scientific work. After much experimentation, I found a couple of approaches that helped me get more done. The first was to keep a big list of tasks that I needed to do for each of my projects, sorted into high, medium and low priority. At the beginning of the day I would review this list, removing tasks that I had finished or were no longer relevant, and then select a series of tasks that I wanted to work on during the day. I would aim to select these predominantly from the list of high priority tasks, but would also add other tasks that were less important but might be easier or more satisfying to achieve. I would then plan out my day by timeboxing—planning out what I wanted to do with each hour of the day.⁵ Initially I did this using pen and paper, but later migrated to my calendar app. While it's possible to get quite granular with timeboxing, allocating specific tasks to each block of the day, I would generally just choose the overall project, and then use the allocated time to work through the list of tasks that I had already chosen.

When dividing up my day, I would allocate work slots in 1.5 hour chunks, generally intermingled with breaks, other meetings/commitments, and exercise.⁶ Depending on what else was happening in the day, I would aim to put three to six of these chunks on my calendar, and then schedule other things around these. During these sessions I would aim to do focused work, with all notifications (e.g., phone, email, etc.) muted. Without a system like this, I found that I was constantly trying to decide between several projects (both scientific and non-scientific), and what to prioritize within each of those projects. The key element here is to limit the number of times during the day at which you're having to choose what to do next. I found that I would rarely stick perfectly to my plan, but it would always provide me with a useful sense of structure so I could have extended periods of focus.

⁵Here's a basic overview: <https://hbr.org/2018/12/how-timeboxing-works-and-why-it-will-make-you-more-productive>.

⁶If you struggle with taking time off, it could be helpful to explicitly time box this as well.

The systems I've proposed here require forming specific habits (e.g., time tracking, time boxing). I spent a long time vaguely feeling like these systems would be helpful, but rapidly abandoning each attempt at actually implementing them. In the end, what has worked for me is trying to change one thing at a time, then regularly reviewing how I think the experiment has gone, and what I could change to improve it. For example, when I first started timeboxing, I would allocate two hours of focused work, followed by a five minute break, then another two hours of work. I found that I was rarely able to be productive for this long, and the system rapidly fell apart. I then dialed it back, starting by just allocating 45 minute work sessions, split up by breaks and easier admin tasks, before increasing the length of my work sessions.

Overall, the best approach for me has been to start with a system that isn't very ambitious, get that bedded in, and then gradually increase my targets. I found the book *Atomic Habits* (see reading list below) helpful in thinking about this approach. It was also very useful for me to do a weekly review, so I could reflect on how things were going and what I wanted to experiment with over the next week.⁷

Productivity systems have a tendency to become a goal in and of themselves. Having a weekly review also helped me limit the time I was spending tweaking my approach. When I had an unproductive day, I would sometimes feel the need to immediately spend time reading about new systems and tools and trying to reconfigure everything about how I was managing my time. If I know I'm going to do some kind of weekly review, I can write down new ideas/tools that I come across, then look into them during the review, rather than constantly wondering if I could do much more work if only I learnt how to use this shiny new app. When doing a weekly review, I found that the best approach was to create a Google form with a specific set of questions that I could go through each week (e.g., what did you experiment with last week, and how could it be improved? What will you experiment with next week? Are you spending your time as you want to?). This method makes the weekly review nice and straightforward.

The point of all of this is not to try and squeeze every ounce of productivity from every second of the day, or to make a beautiful daily plan, but rather to help find some amount of balance between life's competing priorities; something that I think

⁷For some general advice on how to do a weekly review, see here: <https://todoist.com/productivity-methods/weekly-review>.

most PhD students struggle with.⁸ It's really important to make sure you've got time for family, friends, hobbies, sleep, eating well, and exercise. I also believe that academics have a responsibility to share their science outside of the lab, and work to ensure that the academic environment is welcoming to all, both of which can be time consuming activities that often fall on people who are already marginalised in academia. Overall, you don't want to have your identity so wrapped up in your research that you need your work to be going well in order to feel good about yourself. When things aren't going so well (which will definitely happen), you will feel much better if you have a rich life outside of the lab. You also shouldn't hesitate to look for mental health support if you're having a tough time, although unfortunately stigma can make this a challenge.⁹

Finally, I should caveat that it is a lot easier to give advice than to take it. Most of my learning has come from error, with a few successes along the way. Hopefully some of this will be useful to you!

C.2 A Few Extra Pieces of Scattered Advice

- If you're writing code that only you are going to use, think about the end user as yourself in six months. When you come back to this after forgetting the details, will you be able to quickly understand what you were doing and rerun the code? You will not regret taking the time to add in clear comments/docstrings.
- Figure out the format your thesis will have to be in, and write your papers using that format. It's a lot easier to take a bunch of papers in $\text{\LaTeX}$ and put them into a thesis template in Overleaf than converting from Word.
- When you're making figures, try and make as much of each figure using code, and keep modifications from a vector graphics editor to a minimum. This advice applies mostly when there's a high chance that you'll have to remake the figures during paper revisions.
- Spend a bit of time deliberately developing communication skills, for example by taking a specific scientific communications class. I found that these were

⁸I'm not spending much time on the broader systemic problems in academia in this appendix, but here's a related polemic: <https://www.benkuhn.net/grad/>.

⁹Mental health in grad school is a really significant issue. If you're having a hard time, you're definitely not the only one!

skills that could be rapidly developed with minimal effort, but at the same time neglected by many people in the community.

- Figuring out a good desk setup will help you avoid neck and back trouble and save you a world of pain. Caltech offers (or at least did offer) workplace ergonomic consultations¹⁰ that can really help. I found that removing the arms from my desk chair helped me slide further under the desk so I wasn't stretching forwards, and also stopped me constantly leaning to one side.
- Blocking distracting apps on my phone late at night helped me avoid endless late night scrolling. I also deleted the Twitter and Facebook apps from my phone, just using the website versions when I wanted to access them on my phone. The fact that the tech companies really want you to install the apps should tell you something about how much more addictive the apps are compared to the websites.
- Your brain is bad at remembering long lists of things. If you need to remember to do something, write it down.
- Relatedly, always have some method of taking notes with you when meeting with your advisor. By the end of the meeting, try to have agreed on a clear set of tasks to do next, and a rough prioritisation.
- If you come up with systems that work for you, share them! The more people having discussions like this the better.

C.3 Some Specific Tools That I Have Found Helpful

- toggl (<https://toggl.com/>). A cross-platform time tracking app with a free option.
- Freedom (<https://freedom.to/>). An app for restricting access to distracting websites.
- Microsoft OneNote (<https://www.onenote.com/download>) and Notion (<https://www.notion.so/>). Two note-taking apps that I've used for keeping track of projects. There are lots of these, with different features, e.g., Obsidian, (<https://obsidian.md>), which is good for building connections between topics.

¹⁰See here <https://safety.caltech.edu/root-pages/ergonomics>. You can also consult: <https://www.mayoclinic.org/healthy-lifestyle/adult-health/in-depth/office-ergonomics/art-20046169>.

I think it's best to start simple, then gradually experiment with different features.

- Todoist (<https://todoist.com/>). A todo-list app. I used this for keeping track of a wide range of administrative tasks, but generally kept the specific tasks for my scientific projects in my note-taking apps.
- Google scholar (<https://scholar.google.com/>). I set up alerts for several key phrases so I would get emails when new papers were published in relevant areas. I also found it very helpful to look at papers that had cited the paper I was currently reading so I could find more recent advances in the field. There are more sophisticated versions of this (e.g., ResearchRabbit, <https://www.researchrabbit.ai/>), although I haven't experimented with them.
- A reference manager, e.g., Zotero (<https://www.zotero.org/>). You definitely want to avoid keeping all of your papers as pdfs stored in random locations on your hard drive. Save all papers that you read in your reference manager, then use it to generate a Bibtex file for future use in L^AT_EX.¹¹
- Overleaf (<https://www.overleaf.com/>). An online L^AT_EX editor.
- Timeout (<https://www.dejal.com/timeout/>). An app that gets you to take regular breaks from your computer so you can stand up and stretch.
- Headspace (<https://www.headspace.com/>). A guided meditation app.
- A cheap Bluetooth thermometer (e.g., <https://us.govee.com/collections/home-improvement/products/govee-bluetooth-hygrometer-thermometer-h5075>). I found this very helpful for figuring out comfortable working and sleeping temperatures.

C.4 Recommended Reading

Much of my advice from above has been adapted from a variety of “business productivity” books. They can be rather limited in their world view, and make appeals to research of dubious quality. I could spend longer discussing their limitations¹²,

¹¹If you don't know what these things mean, you can start here: [https://www.overleaf.com/learn/latex/Free_online_introduction_to_LaTeX_\(part_1\)](https://www.overleaf.com/learn/latex/Free_online_introduction_to_LaTeX_(part_1)).

¹²For some criticism, see here: <https://www.newyorker.com/culture/office-space/the-frustration-with-productivity-culture>. For some spicy criticism, see here: <https://jacobin.com/2018/03/four-hour-workweek-tim-ferriss-work>.

but these books contain specific tips that have helped me in my day-to-day work, and I think are worth consulting:

- *Deep Work: Rules for Focused Success in a Distracted World* (Cal Newport). Basically, we have lots of distractions in our lives, and these stop us from doing extended period of focused work.
- *Indistractable: How to Control Your Attention and Choose Your Life* (Nir Eyal). Pairs well with *Deep Work*, with some specific advice about how to remove distractions.
- *Getting Things Done: The Art of Stress-Free Productivity* (David Allen). A system for managing your todo-lists. Quite an intricate system, and the book is a bit out of date, but still helpful. The key concept is that your brain is bad at storing tasks, so you should write everything down, then regularly sort through these tasks so you have clear priorities.¹³
- *The Effective Executive: The Definitive Guide to Getting the Right Things Done* (Peter Drucker). Originally published in the 1960s, but has some useful suggestions about tracking your time.
- *Atomic Habits: An Easy & Proven Way to Build Good Habits & Break Bad Ones* (James Clear). A good book if you want to work on building the kind of habits you need to do the things suggested in the other books.
- *Houston, We Have a Narrative: Why Science Needs Story* (Randy Olson). Not a productivity book, but a great resource for thinking about how to communicate your science.
- While I've got you here, I also want to recommend: *I've Got the Light of Freedom: The Organizing Tradition and the Mississippi Freedom Struggle* (Charles M. Payne). This book probably won't help you manage your weekly schedule, but it's a truly masterful book that changed how I see American history.

¹³For a review of different productivity systems, see: <https://todoist.com/productivity-methods>. My basic approach is a simplified version of the *Getting Things Done* method, roughly what is described here: <https://todoist.com/productivity-methods/systemist>.

Is the truth depressing? Some may find it so. But I find it liberating, and consoling. When I believed that my existence was such a further fact, I seemed imprisoned in myself. My life seemed like a glass tunnel, through which I was moving faster every year, and at the end of which there was darkness. When I changed my view, the walls of my glass tunnel disappeared. I now live in the open air.

—Derek Parfit (1942-2017), *Reasons and Persons* (Chapter 13, p. 281)