6-2022

Analysis of Aftershock Parameters for the Alaskan Subduction Zone Tectonic Region

Gabrielle M. Paris
Portland State University, parisgab@gmail.com

Analysis of Aftershock Parameters for the Alaskan Subduction Zone Tectonic Region

Gabrielle M. Paris

June 13, 2022

Undergraduate Thesis

Portland State University

Advisors

Richard C. Hugo *(Geology Department, Portland State University)*

Andrew J. Michael *(Earthquake Science Center, US Geological Survey)*

Abstract

Forecasting how many earthquakes will occur following a potentially damaging earthquake helps the public and emergency operators stay safe and make informed decisions. The U.S. Geological Survey issues aftershock forecasts following potentially damaging earthquakes, using models to predict the number of earthquakes that should occur within the next day, week, month, and year with 95% confidence to reflect the uncertainty in aftershock behavior. The USGS considers the forecast to be "successful" when the number of earthquakes observed within the forecasted time period is within the 95% confidence interval. For aftershock sequences that occur along the forearc of the Alaskan subduction zone, the observations consistently lie within this broad range of "success," however the forecasts systematically over-predict the number of aftershocks. This suggests that the parameters which drive the forecast model's rate of decay needs to change to reflect the decreased observed aftershock activity. I analyzed seventeen earthquake sequences in the Alaska forearc region and have developed a set of parameters that may improve early-sequence aftershock forecasting in the region. I stacked the 17 sequences and used maximum likelihood estimation to determine the new parameters and their uncertainty. Following the lessons of Page et al. (BSSA, 2016), I combined the inter-sequence variability and the uncertainty of the parameters from the stacked sequences to produce the total standard deviation for a new forecast model.

Introduction

The U.S. Geological Survey issues forecasts after most earthquakes above magnitude 5 in the U.S. states and territories to inform the public and emergency managers of how many subsequent earthquakes, or aftershocks, are to be expected following a potentially damaging earthquake. An aftershock forecast is a model which predicts the number of earthquakes above a minimum magnitude that should occur within the next day, week, month, and year. Aftershock activity shortly after the initial earthquake, known as the mainshock, is highly uncertain, and the early forecasts account for this uncertainty by predicting a wide range of aftershocks that could occur within a given time frame. This range is two standard deviations from the forecast median and is referred to as the forecast's 95% confidence interval. As more earthquakes happen, the uncertainty decreases. Sometimes, the forecasts overpredict or underpredict the aftershock behavior, and the USGS needs to manually update the forecasts. This was the case for the magnitude 8.2 Chignik earthquake sequence that occurred near Perryville, Alaska on 29 July, 2021. The early forecasts issued by the USGS during this event overpredicted the number of earthquakes following a magnitude of this size (Figure 1), and the USGS had to manually adjust the forecast to better describe the aftershock activity, which only improved forecasts in the short term.

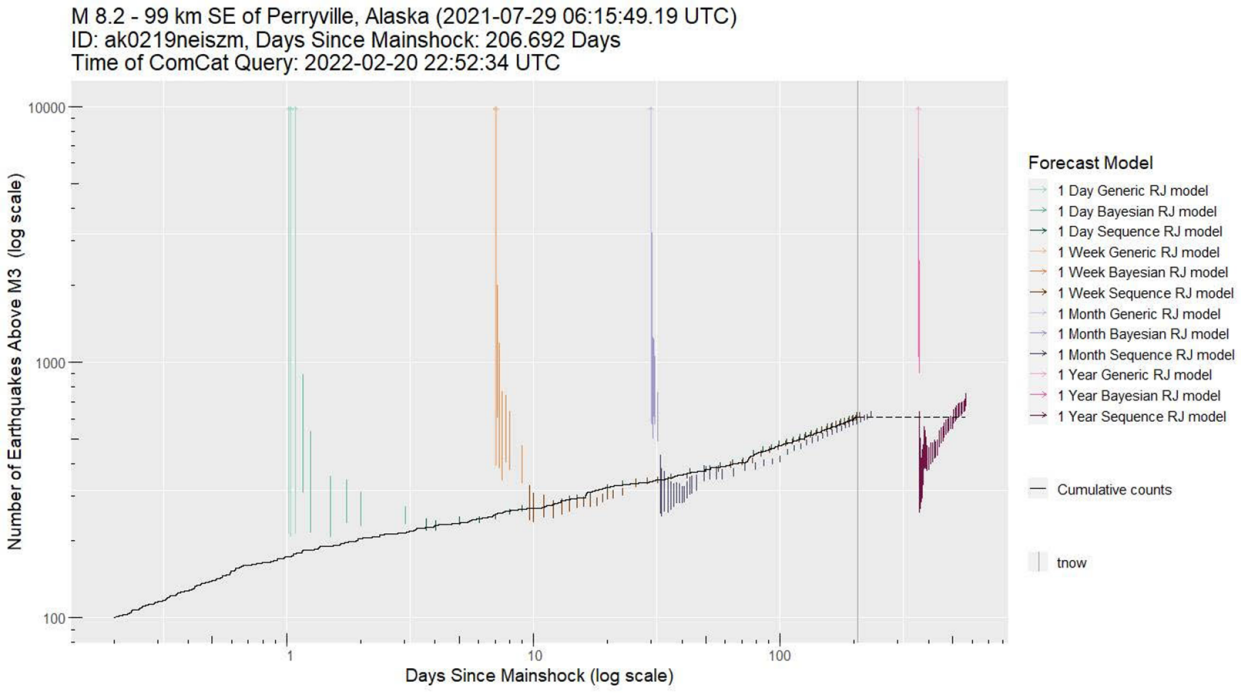


Figure 1. Aftershock forecasts issued for the 2021 M8.2 Chignik earthquake sequence. Each vertical line is a forecast, where color represents the duration of the forecast, and the shade of each color represents the aftershock model issued for a given duration. The vertical extent of the lines reflect the 95% confidence interval for the number of earthquakes that should occur by the end of the forecast duration, and the forecasts are shifted upward to account for the number of earthquakes that have occurred prior to the start of the forecast. The forecasts are shown at the end of their duration period, for example, the first forecast is issued 20 minutes after the mainshock and is plotted at a day and 20 minutes, a week and 20 minutes, and so on. A forecast that accurately predicts the number of earthquakes will be intersected by the black cumulative number line. The figure follows the design of Figure 5 from Michael et al. (2019).

Since the highest level of aftershock activity occurs shortly after the mainshock, it is critical for the forecast to reflect the early aftershock behavior so the model can be further improved. I evaluated these early forecasts to evaluate whether the initial parameters used to create the confidence interval in the forecast need to be updated. I focused on earthquake sequences that occur along the subduction zone off the coast of Southern Alaska because it is an active tectonic boundary whose aftershock behavior is not well characterized. The variation in the observations within the 95% confidence interval defined by the forecast will determine if the forecast parameters need to be updated. Specifically, if the observations cluster on one side of the forecast's median instead of randomly distributed throughout the confidence interval, the forecast is not reflecting the observed aftershock behavior.

Background

Currently, the USGS uses a Reasenberg and Jones aftershock model, which was developed in 1989 and 1994 by Reasenberg and Jones to describe the rate of earthquakes of at least some magnitude following the mainshock at a given duration period –the period of time during which an aftershock forecast is in use— after the occurrence of the mainshock (Reasenberg and Jones, 1994). This creates a model that decays with time to reflect the decreasing effect of the strain release along a fault. The model depends on the magnitude of completeness (M_c), which is the minimum magnitude above which all earthquakes are detected (Helmstetter et al., 2006; Mignan and Woessner, 2012). When evaluating the effectiveness of aftershock forecasting, it is important to consider magnitude of completeness to avoid under-sampling (when M_c is too high) or biasing the analysis (when M_c is too low) (Mignan and Woessner, 2012). Increased seismicity due to an aftershock sequence or the spatial distribution of seismic stations, for example, can affect the minimum magnitude above which an earthquake is reliably located and recorded.

In the Reasenberg and Jones equation (Eqn. 1), the parameters that drive the aftershock model's decay rate are the productivity parameter *a*, which controls the overall expected number of aftershocks for a given mainshock magnitude *Mmain*, and the change in the rate of aftershocks with time, *p*. The parameter *b* is both the scaling of aftershock rate with mainshock magnitude and the ratio of small to large aftershocks, *c* is a unit of time representing the earliest part of the aftershock sequence before it starts to decay, *T* is time after the mainshock, and *M* is a minimum magnitude for which the model applies

$$R = 10^{(a+bMmain)}10^{(-bM)}(T+c)^{-p}. \qquad \text{Eqn. 1}$$

The Reasenberg and Jones equation can be split into three parts. The first part ($10^{(a+bMmain)}$) is the Utsu relationship for productivity, which states that there are approximately ten times as many aftershocks for a mainshock that is 1 mainshock magnitude larger, when $b = 1$ (Utsu, 1961). The second part of the equation (10^{-bM}) is the Gutenberg-Richter relationship, which describes the relationship between a given magnitude and the number of earthquakes of at least that magnitude, regardless of when and where the earthquake occurred (Gutenberg and Richter, 1944). This distribution tends to be linear on a plot of log frequency over magnitude when the earthquakes are detectable and when they are smaller than the mainshock. Båth's law is related to the Gutenberg-Richter relationship in that for a given mainshock magnitude, the average largest aftershock is between 1-1.2 times smaller, then from the largest aftershock there is an order of magnitude increase for each decrease in unit magnitude (Båth, 1965). For example, it is expected that a magnitude 6 mainshock will have a M4.8 to M5 aftershock as the largest aftershock, then 10 M4s, 100 M3s, and so on. The third part of the equation ($(T+c)^{-p}$) is the Omori law modified by Utsu in 1961, and it describes the change in aftershock frequency with the reciprocal of time (Utsu, 1961). This creates a decay curve on a plot of cumulative number of earthquakes with time for earthquakes above a time-dependent magnitude of

completeness, which is the minimum magnitude above which the information pertaining to an earthquake is reliably recorded in an earthquake catalog.

The specific values that are put into Eqn 1 are dependent on the tectonic setting. Page et al. (2016) expanded on the work of Reasenberg and Jones by better characterizing parameter uncertainty and by globally calculating aftershock parameters based on a tectonic regionalization (García et al., 2012), which defines regions of similar tectonic settings. Page et al. used a dataset of sequences from January 1, 1990, to January 1, 2015, and calculated the parameters by using maximum likelihood estimation to determine the most probable detection rate from a cumulative normal distribution. The fit to the maximum likelihood estimation becomes the value that describes the expected productivity of a sequence in a given tectonic region, and the inter-sequence variability is obtained through inversion. For onshore subduction zones (defined below in "Geologic Setting"), Page et al. calculated the productivity parameter *a* to be -2.34 and the change in productivity *p* to be 0.81.

The Page et al. aftershock parameters are referred to as 'Generic parameters' by the USGS and by this thesis, and they are used early in an aftershock sequence when the sequence's behavior is highly uncertain. Since there is high uncertainty and variability in the aftershock sequence in the early hours to days following the mainshock, the model includes a large uncertainty for the productivity parameter *a*, which in turn defines the large confidence interval seen in the early forecasts. As time progresses, the sequence's behavior becomes more certain and the model can be fine-tuned to fit the observed behavior. When the forecasts need to be manually updated, the parameters are calculated using the same methods to create the generic parameters for the aftershocks that have occurred up to the time of the forecast. The model thus transitions from Generic to a Bayesian combination, which allows the *a*-value to change while keeping the *p*-value fixed. As the sequence progresses, the model can shift to a Sequence Specific one, in which both the *a* and *p* values are allowed to change. By the time a Bayesian or Sequence Specific model is used, there have been enough aftershocks to reduce the uncertainty in the sequence's behavior.

Geologic Setting

The tectonic regions describe geologic settings with similar stresses and characteristics, such as the type of fault, the depth, and the location of the event relative to the trench (García et al., 2012). In a subduction zone, the onshore region is approximately where the forearc is, which lies between the subduction trench that is seaward of the accretionary wedge and the inland volcanic arc (Figure 2). Three areas in Southern Alaska make up the Alaskan onshore region: the Alaskan Peninsula, Kodiak Island, and the area surrounding Cook Inlet and Anchorage.

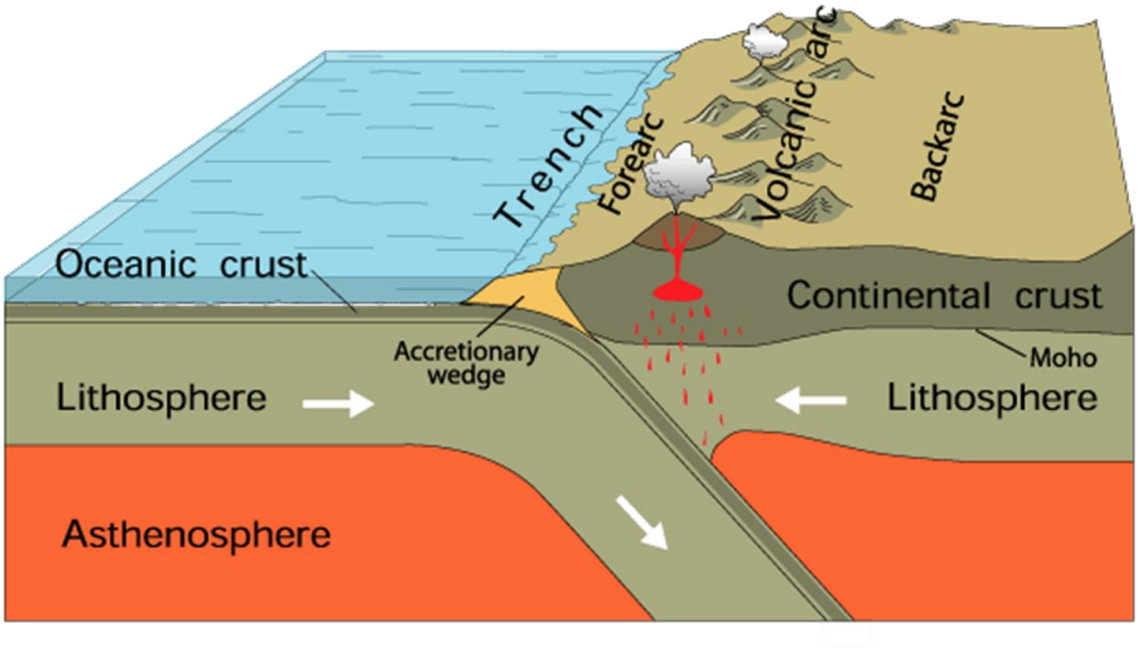


Figure 2. General components of a subduction zone. The onshore part of the subduction zone is approximately where the forearc is. It is landward of the accretionary wedge, which consists of marine sediments that were scraped off the ocean floor as the oceanic crust is diving beneath the continental crust, and seaward of the volcanic arc, which is the product of the ocean crust melting due to high temperature and pressure. Figure from U.S. Geological Survey website on subduction zones (2016).

Methods

Seventeen earthquake sequences that occurred between January 1, 2018 and February 22, 2022 were chosen for this study and their locations are displayed on the map in Figure 3. A start date of January 1, 2018 was chosen because the USGS started issuing aftershock forecasts using the Page et al. parameters in 2018. Most of the sequence mainshocks are small, between magnitudes 5 and 6, and there are four mainshocks greater than magnitude 7 with at least 1000 aftershocks: the 2018 M7.1 Anchorage earthquake, the 2020 M7.8 Simeonof earthquake, the 2020 M7.6 Sand Point aftershock to the M7.8, and the 2021 M8.2 Chignik earthquake. The M7.6 Sand Point earthquake, while classified by the USGS as a large aftershock to the M7.8 Simeonof earthquake, is considered its own aftershock sequence in this study because earthquake activity increased following this event. The only determining criterion in selecting the earthquake sequences was that the mainshocks are all classified by the USGS as onshore subduction zone earthquakes. All the earthquake sequences were retrieved from the USGS Comprehensive Catalog (see Data and Resources).

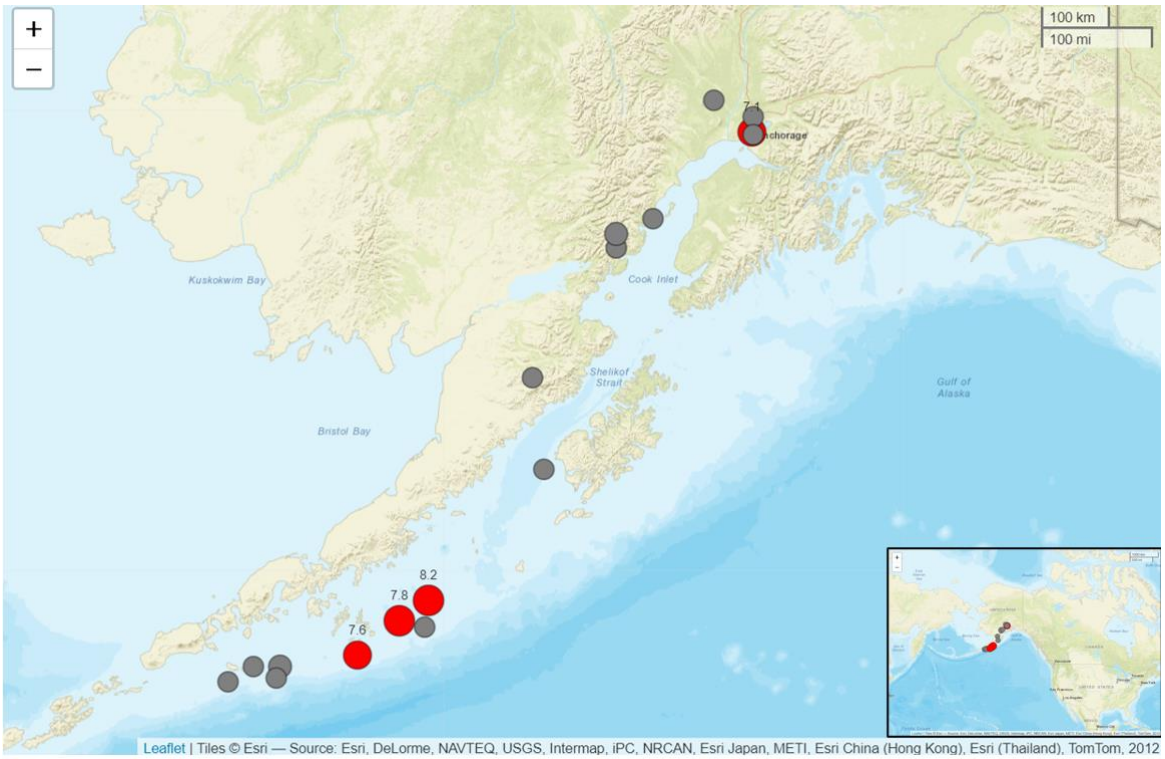


Figure 3. The spatial distribution of earthquake sequences the USGS has issued forecasts for that have occurred between January 1, 2018 and February 22, 2022 in the onshore subduction zone tectonic region. Red circles indicate sequences with a mainshock greater than M7 and contain at least 1000 aftershocks. Gray circles are sequences with a mainshock magnitude between 5 and 6.

Two tools were used to evaluate the performance of the aftershock forecasts. The first was the USGS oaftools package built in R, an open-source programming language for statistical computing (see Data and Resources for R and oaftools). The oaftools package contains functions designed to process the forecast files, get aftershock data from the Comprehensive Earthquake Catalog, and generate plots to analyze aftershock behavior (Paris and Michael, 2022). The second was the USGS Aftershock GUI, which is an interface that allows the user to change forecast parameters and analyze the impact of those changes (see Data and Resources). All other computational work was done in R.

To determine if the forecasts are accurately describing aftershock behavior, I retrieved all aftershocks for the seventeen sequences from the Comprehensive Earthquake Catalog (ComCat) and compared each sequence's rate of decay with the Generic forecasts. Since the USGS only issues Generic forecasts up to the first two hours after a mainshock, I had to simulate forecasts to thoroughly analyze the aftershock behavior. I used the USGS Aftershock GUI to create forecasts using a modified forecast schedule. The simulated schedule (Table 1) mimics the USGS

automatic forecast schedule, found on the USGS website (see Data and Resources), but with a focus on the early sequence when the Generic model is used.

Once each sequence had simulated forecasts using the Generic aftershock parameters, I divided the number of completed forecasts by the number of forecasts that accurately predicted the number of aftershocks, then plotted the 95% confidence intervals and the forecast median with the observed number of aftershocks to see where the observations were relative to the median. To better characterize the distribution of the observed number of aftershocks, I normalized the data to the forecast median by dividing the number of observed earthquakes by the median forecasted value, then calculated the percentage of observations that lie below the median. When calculating the percentage of observations that lie below the median, I excluded data at the median because these often represent a forecast that predicted zero earthquakes, and zero earthquakes occurred, which is difficult to evaluate especially in long-term forecasting.

I used two approaches to calculate new aftershock parameters. First, I calculated the most likely *a* and *p*-values for each sequence to characterize the inter-sequence variability. Second, I stacked the aftershock sequences by normalizing the times of occurrence relative to the respective mainshocks, then combined the aftershocks into one earthquake catalog with a calculated mainshock magnitude for the stacked sequence ($Mmain_{stack}$) using Equation 2

$$Mmain_{stack} = \frac{1}{b}\, log(\sum_{1}^{17} \quad 10^{bMmain}), \qquad \text{Eqn. 2}$$

where *b* is the scaling of aftershock rate with mainshock magnitude, and *Mmain* is the unstacked mainshock magnitude. For both approaches, I used the first 10 and 70 days after each mainshock, chosen to reduce contamination from background seismicity, to see how the parameters changed as the sequence transitioned from short term to long term. This method is similar to the lesson in Page et al. (2016), who calculated the most likely parameters at 10 and 100 days. I calculated the mean *a*-value, *p*-value, and respective standard deviations from the Bayesian model and the Sequence Specific model. I also calculated the most likely *a* and *p* values for the stacked catalog at different magnitudes of completeness.

Table 1. *Simulated aftershock forecast schedule.*

Time (Days since the Mainshock)	Time
0.00694	10 Minutes
0.02083	30 Minutes
0.03472	50 Minutes
0.08333	2 Hours
0.25	6 Hours
0.5	12 hours
1	1 Day
2	2 Days
5	5 Days
10	10 Days
20	20 Days
30	30 Days
60	60 Days
100	100 Days
150	150 Days
200	200 Days
250	250 Days
300	300 Days
365	1 Year

Results

Prior to analyzing the Generic parameters, I normalized all the USGS-issued forecasts to the time of each respective mainshock to understand the temporal evolution of the *a*-value (Figure 4). Most of the sequences were issued a final *a*-value lower than the Generic parameter of -2.34.

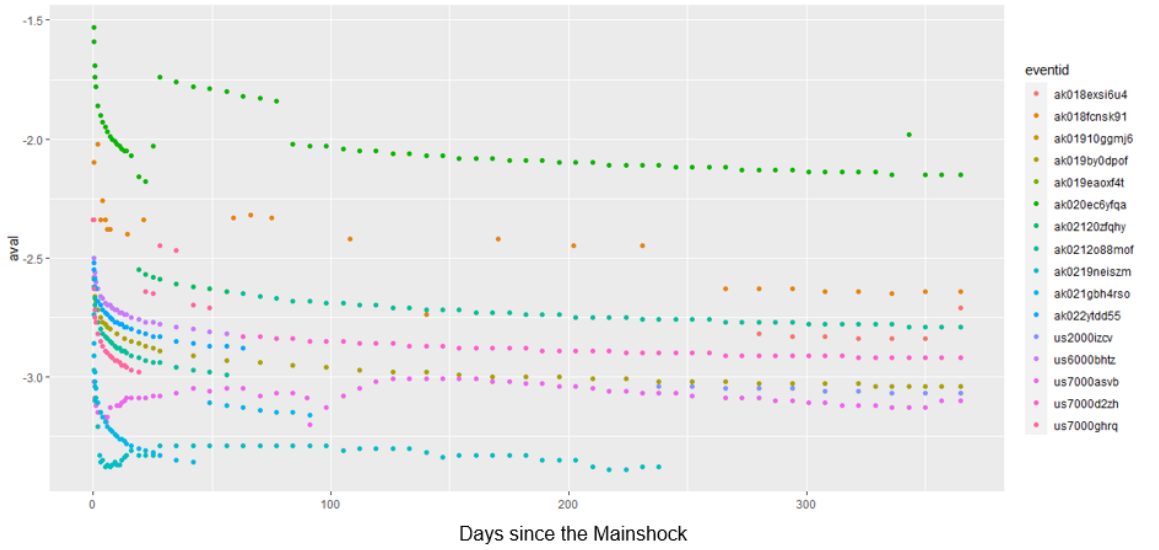


Figure 4. Manually adjusted *a*-values for the 17 earthquake sequences. The data are colored by the sequence and are plotted as days since the mainshock. Generic, Bayesian, and Sequence Specific parameters are shown on the same plot, but most of the points are Sequence Specific values. The horizontal axis is days since the mainshock, and the vertical axis is the productivity parameter *a*.

The first calculations described in my methods determined the percentage of forecasts where the observed number of aftershocks was within the forecast's 95% confidence interval. This was done to see how the forecasts were doing on a broader scale. Tables 2 and 3 show the ratio of forecasts that correctly predicted the number of aftershocks to the number of completed forecasts, for different minimum magnitudes and for different time periods, for the M7.8 Simeonof sequence and for a better characterized M5.3 sequence. For sequences with a mainshock magnitude between 5 and 6, all the forecasts reflected the observed aftershock behavior. For sequences with a mainshock magnitude greater than 7, only the high-probability, low-magnitude forecasts were failing to fit the observed aftershock behavior.

Table 2. *Table of successful forecasts for a M7.8 aftershock sequence.*

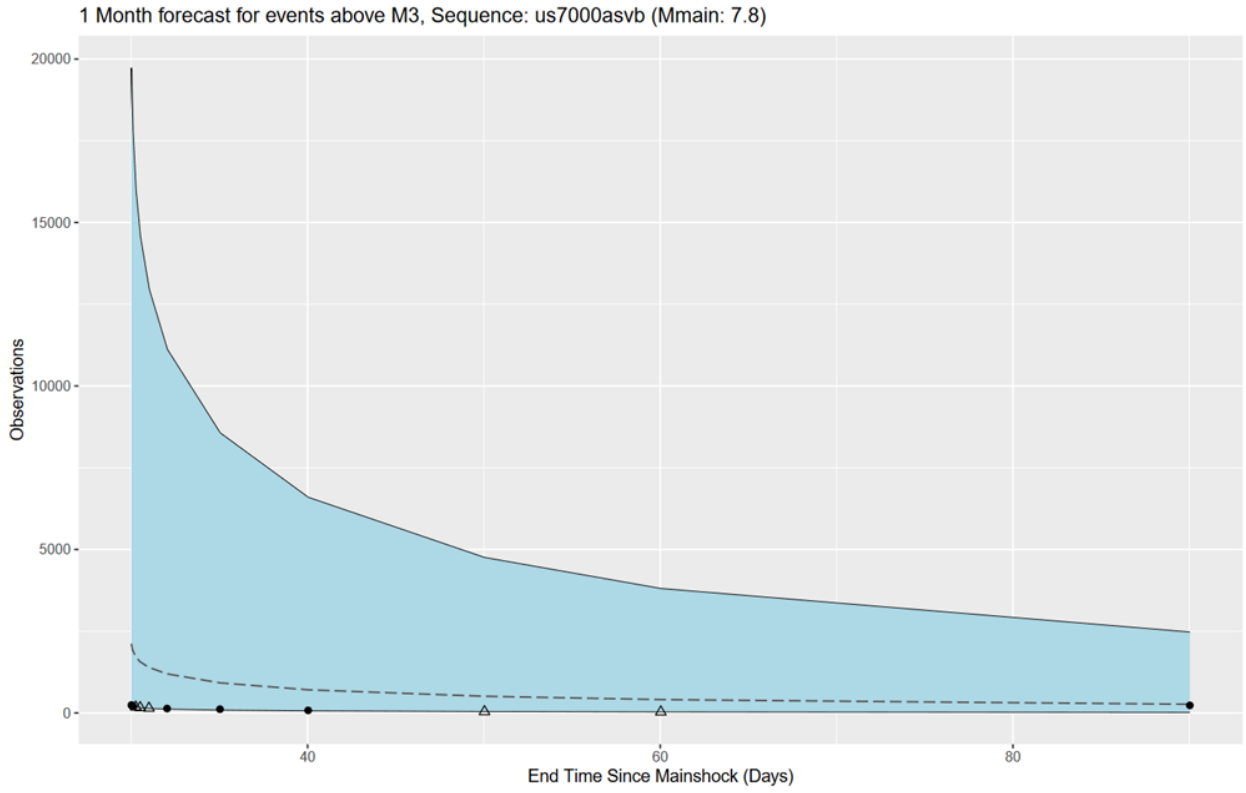
Minimum Magnitude	1 Day Forecast	1 Week Forecast	1 Month Forecast	1 Year Forecast
3	54% (7/13)	23% (3/13)	62% (8/13)	100% (13/13)
4	100% (13/13)	100% (13/13)	100% (13/13)	100% (13/13)
5	100% (13/13)	100% (13/13)	100% (13/13)	100% (13/13)
6	100% (13/13)	100% (13/13)	100% (13/13)	100% (13/13)
7	100% (13/13)	100% (13/13)	100% (13/13)	100% (13/13)

Table 3. *Table of successful forecasts for a M5.3 aftershock sequence.*

Minimum Magnitude	1 Day Forecast	1 Week Forecast	1 Month Forecast	1 Year Forecast
3	100% (19/19)	100% (19/19)	100% (19/19)	100% (19/19)
4	100% (19/19)	100% (19/19)	100% (19/19)	100% (19/19)
5	100% (19/19)	100% (19/19)	100% (19/19)	100% (19/19)
6	100% (19/19)	100% (19/19)	100% (19/19)	100% (19/19)
7	100% (19/19)	100% (19/19)	100% (19/19)	100% (19/19)

Figures 5a and 5b follow the design of Figure 2 from van der Elst et al. (2022) and show where the observed number of aftershocks lie within the 95% confidence interval for the previous two sequences. In both the M7.8 and the M5.3 sequences, the observed number of earthquakes are distributed below the forecast median. While the forecasts may be predicting aftershock behavior, the range of "success" is broad and still does not accurately represent the observed aftershock behavior.

(a)



(b)

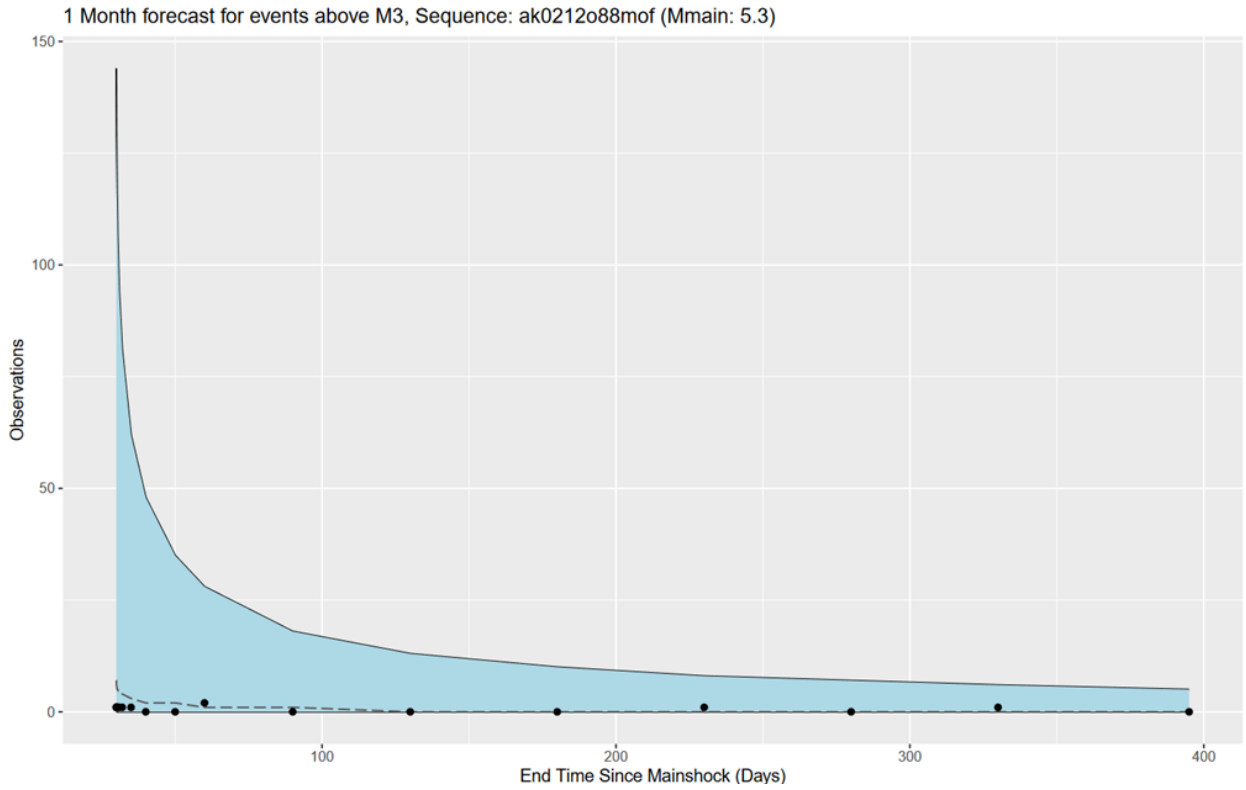


Figure 5. The observed number of earthquakes for both the M7.8 Simeonof sequence (a) and a better characterized M5.3 sequence (b) are consistently at or below the forecast median. The shaded blue region is the 95% confidence interval, with black lines capping both ends, and the dashed line is the forecast median. If the observed number of earthquakes lies within the 95% CI, the data points are circles, otherwise they are unfilled triangles.

To quantify the distance between the observed number of earthquakes and the forecast median, the data for all seventeen sequences are normalized to the median (Figure 6). I took the mean of the observations and found that most of the observations are below the forecast median, exclusive of values at the median. The percentage of observations below the median are: 73.5% for 1 Day forecasts, 80.9% for 1 Week forecasts, 86.5% for 1 Month forecasts, and 82.6% for 1 Year forecasts. The mean of the observed/median ratio in Figure 6 is approximately 0.29 for all duration periods.

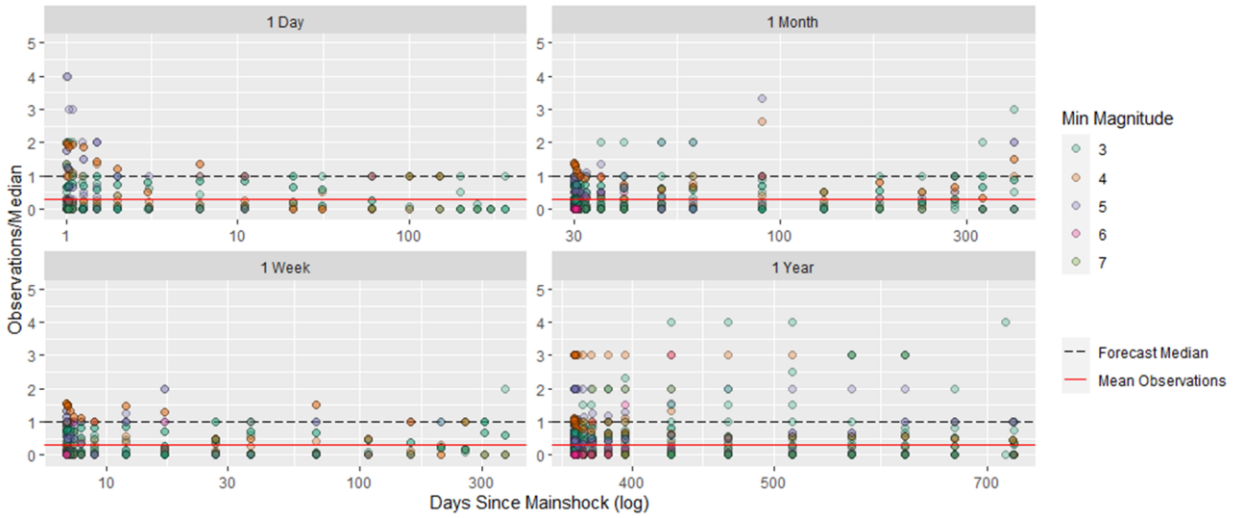
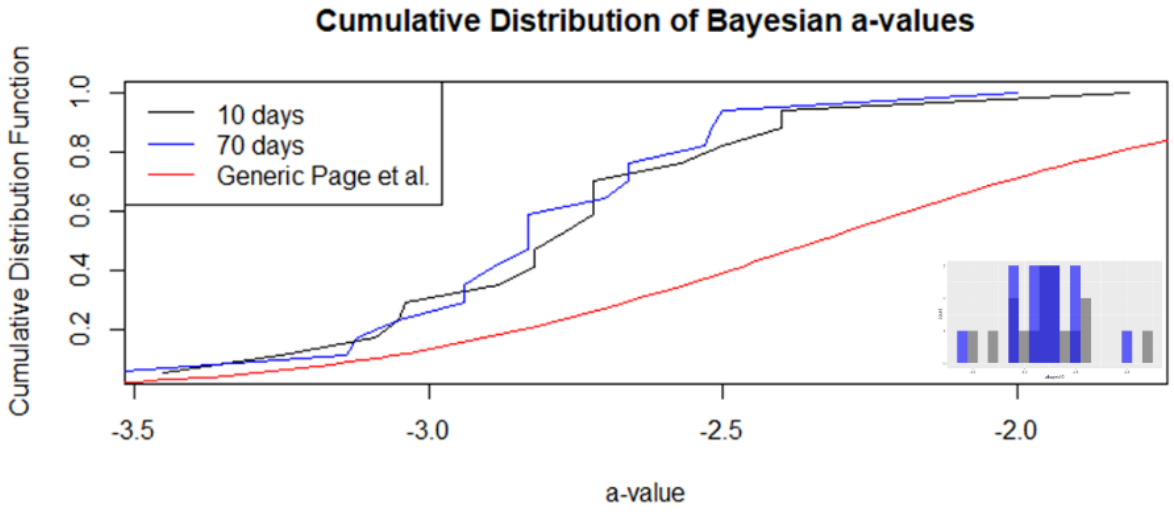


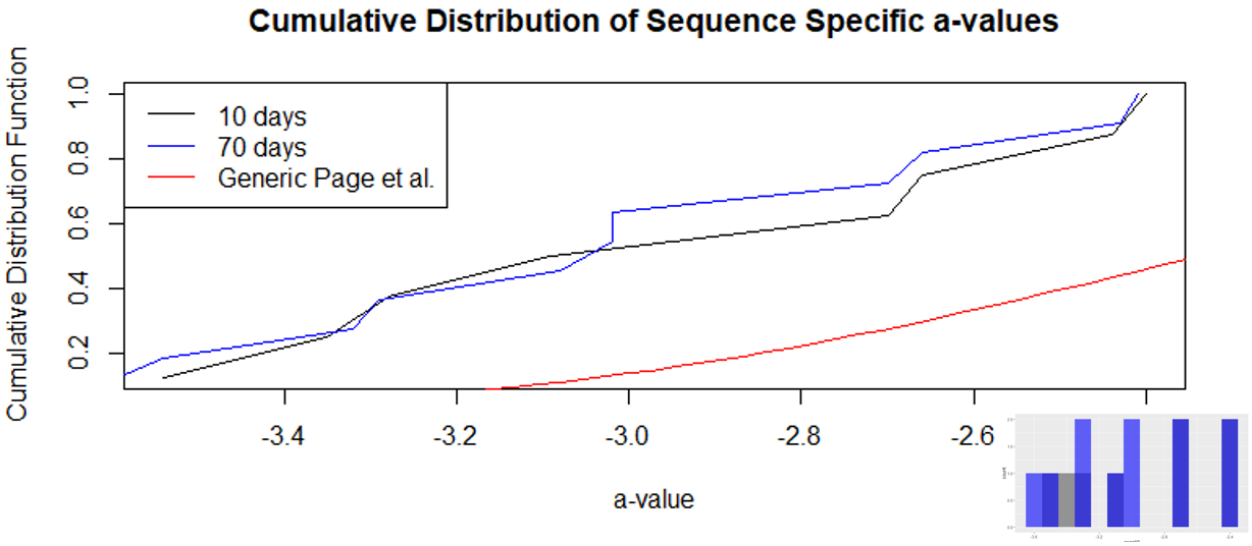
Figure 6. Normalized aftershock productivity for the 17 aftershock sequences. The observed number of aftershocks are normalized to the forecast median (dashed line) and are colored according to the minimum magnitude the forecast was issued for. The data are grouped by forecast duration period to avoid overlapping. The red line is the mean of the data, calculated as roughly 0.29 for each duration period.

A cumulative distribution plot shows the distribution of data, and a steeper slope indicates the data's mode, analogous to the peak of a histogram. I compared the inter-sequence variability for all seventeen sequences by estimating the most likely short-term and long-term *a* and *p* values. In Figures 7a and 7b, the lines representing 10 day and 70 day most likely parameters overlap, suggesting that the decay across all 17 aftershock sequences behaved similarly. The 10-day and 70-day distributions are also distinct from the Generic *a*-value distribution, represented in Figures 7a and 7b as a red line. In Figure 7b, sequences with a likely *a*-value of -4.5 were excluded because these sequences have almost no recorded aftershocks. Therefore, it is more difficult to identify a mean likely *a*-value because of the lack of an identifiable inflection point in the distribution. In Figure 7c, the most likely *p*-values are horizontally offset but still exhibit a similar distribution, which indicates that the aftershock rate changes as expected between short term and long-term forecasting. The Generic *p*-value is represented as a vertical line in Fig. 7c because there is no uncertainty in the *p*-value for the Generic aftershock model.

(a)


Cumulative Distribution of Bayesian a-values
Cumulative Distribution Function
0.2
0.4
0.6
0.8
1.0
10 days
70 days
Generic Page et al.
-3.5
-3.0
-2.5
-2.0
a-value

(b)


Cumulative Distribution of Sequence Specific a-values
Cumulative Distribution Function
0.2
0.4
0.6
0.8
1.0
10 days
70 days
Generic Page et al.
-3.4
-3.2
-3.0
-2.8
-2.6
a-value

(c)

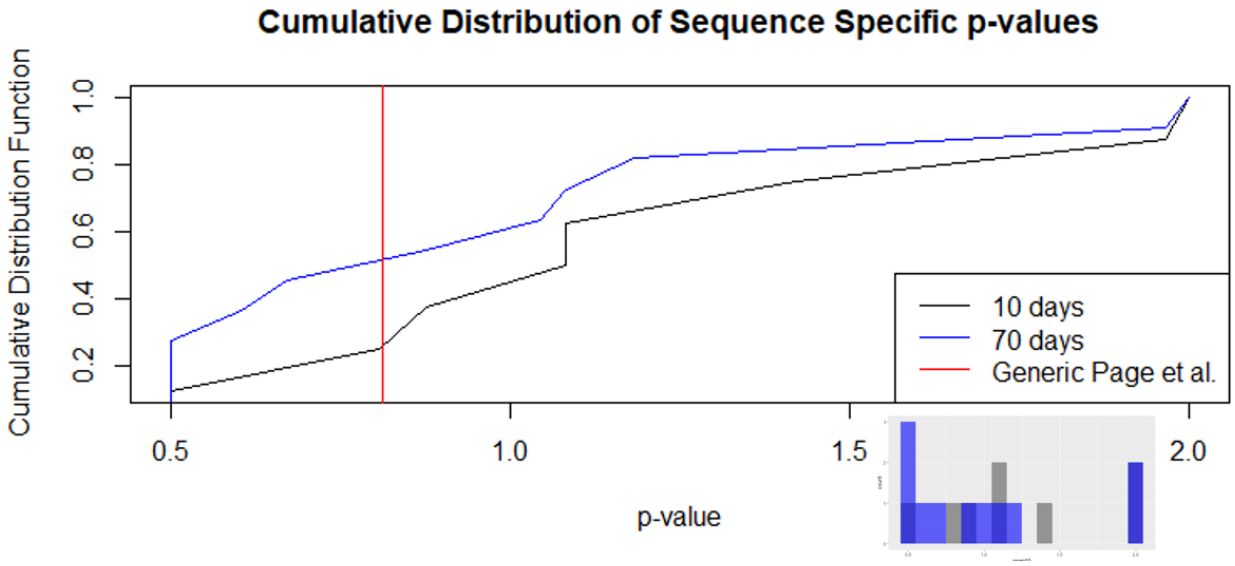


Figure 7. The distribution of most likely *a*-values for the 17 sequences. The 10-day and 70-day distributions are similar to each other , and the two are distinct from the Generic *a*-value distribution from Page et al. (2016). The *p*-value showed some variability between short-term and long-term forecasting because the decay rate is higher at 10 days than at 70 days, so the 70-day line is shifted to lower values to account for that. In the Bayesian model (a), the *p*-value was fixed to 0.81 and the *a*-values were determined using maximum likelihood estimation, whereas in the Sequence Specific model (b,c), both *a* and *p* were allowed to change. A histogram of the same data is included in the lower right corner of the plot for reference, colored by the same non-Generic parameters used in the cumulative distribution plot.

To account for the inter-sequence variability, I stacked the first 70 days of all seventeen aftershock sequences and calculated the most likely *a* and *p* values for the stacked catalog. I also computed the Reasenberg and Jones aftershock model for the stacked catalog, which is shown in Figure 8. The models that use the most likely Bayesian and Sequence Specific parameters better fit the observed aftershock behavior than the Generic Page et al. parameters.

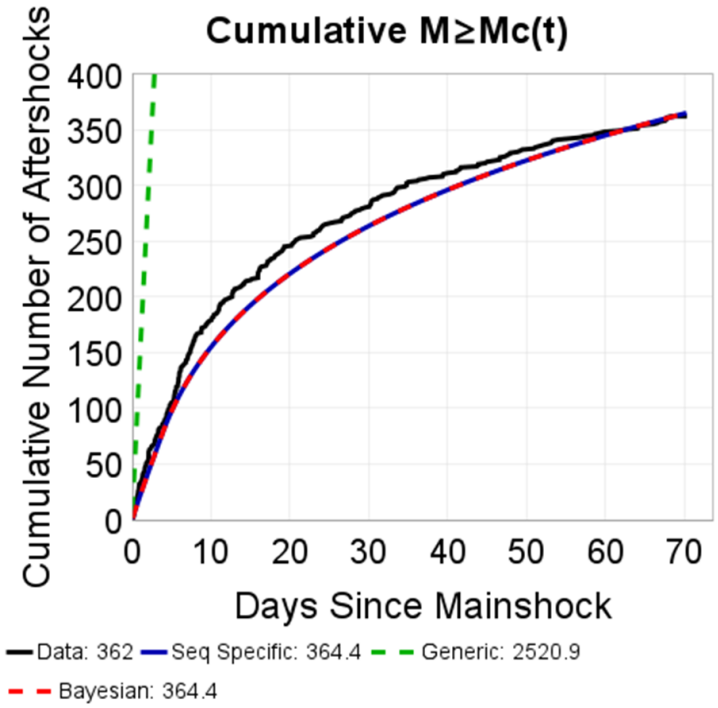


Figure 8. Cumulative aftershock activity of the stacked aftershock sequencces. The black line is the cumulative number of aftershocks, the green dashed line is the Generic model, and the red dashed and solid blue lines are the Bayesian and Sequence Specific models, respectively.

I averaged the likely *a* and *p* values with their respective uncertainties for values obtained from a Bayesian model and a Sequence Specific model at 10 days and 70 days, and the results are in Table 4. The new *a*-values are lower than the Generic Page et al. *a*-value, while the *p*-values are closer to the Generic *p*-value. To test for robustness, I evaluated the *a* and *p* values at different magnitudes of completeness. The *a*-value increases as magnitude of completeness increases, and the uncertainty of the *a*-value range is within ± 0.1, indicating that the results discussed in this paper are not due to catalog incompleteness. Furthermore, the *p*-value range has an uncertainty of ± 0.05, indicating that the aftershock rate in the onshore region is consistent as magnitude of completeness changes.

Table 4. *Estimated means of a and p values for the stacked aftershock catalog.*

Aftershock Model	Magnitude of Completeness (M_c)	a-value	p-value
Bayesian	3.5	-3.15 ± 0.028	0.81 ± 0.000
Sequence Specific	3.5	-3.09 ± 0.044	0.89 ± 0.054
Bayesian	4.0	-3.04 ± 0.037	0.81 ± 0.000
Sequence Specific	4.0	-3.03 ± 0.048	0.84 ± 0.064
Bayesian	4.5	-2.94 ± 0.054	0.81 ± 0.00
Sequence Specific	4.5	-2.94 ± 0.062	0.79 ± 0.078

Discussion

The decrease in *a*-value across all sequences to a value below the Generic parameter in Figure 4 suggests that the productivity and decay rate seen in the M8.2 Chignik earthquake sequence was not specific to that event. Additionally, the consistent decrease in the *a*-value indicates that the seventeen aftershock sequences have a similar productivity and decay at a similar rate.

The distribution in Figures 5a and 5b are made up of the forecast median and the 95% confidence interval. These are not normal distributions because it is more likely that fewer earthquakes will be observed, but since the minimum number of earthquakes observed cannot go below zero, the distribution is bounded on the lower side. By performing this analysis on a magnitude 7.8 and a 5.3, as well as normalizing the data to the median, I could rule out the possibility that the size of the mainshock affected the decay rate of the aftershock sequence.

If the mean observed/median ratio in Figure 6 was closer to the forecast median, then the existing forecast parameters may still be working, but the mean is far from the median and therefore the parameters need to be adjusted. The randomness of the minimum magnitudes across all forecast duration periods in Figure 6 indicates that the over-prediction is not specific to one forecasted minimum magnitude, nor is there a distinction between short-term forecasting and long-term forecasting. Another way of looking at the distribution of observations relative to the median is with a histogram (Figure 9). It is easier to see that not only do most of the observed number of earthquakes lie beneath the forecast median, but most of the observations are close to zero. This abundance near zero is expected when considering that it is more likely for zero earthquakes to occur than one or more earthquakes. The abundance of earthquakes near zero in

Figure 9 could also be due to sequences with smaller mainshock magnitudes reaching a probability of zero earthquakes earlier.

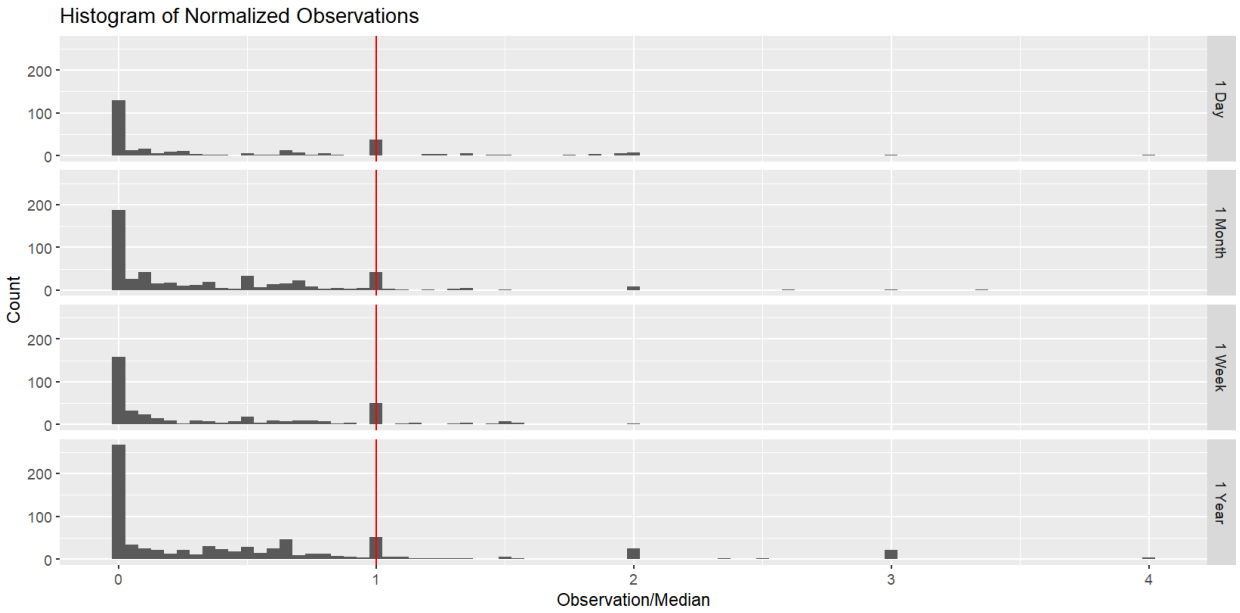


Figure 9. Distribution of forecasts relative to the forecast median. There is an increased number of observations at the median, but this could be due to the inclusion of low-probability forecasts, when the predicted number of earthquakes is zero and there are zero observations. The data are grouped by forecast duration period, and the red vertical line is the forecast median.

The overlap between the 10-day and 70-day parameters in Figures 7a and 7b suggest that the decay across all 17 aftershock sequences behaved similarly and in a consistent manner. The offset between 10-day and 70-day *p*-value distributions could reflect the Omori decay of the sequence. In other words, the aftershock rate is high in the early sequence compared to 70 days after the mainshock, resulting in a series of *p*-values that shifts the distribution to lower *p*-values.

This work is important for maintaining operational awareness and providing the public with reliable information after potentially damaging earthquakes. It is also necessary to provide reliable forecasting because the public can lose confidence in the USGS if the aftershock forecasting is excessively over- or under-predicting the number of aftershocks following a potentially damaging earthquake.

As mentioned in the previous section, the results from my analysis are not due to catalog incompleteness or inconsistent aftershock decay rates. These results are anomalous, however, when compared to Page et al.'s maximum likelihood solutions for the onshore region. In Figure 10, the estimated most likely *a*-values from Table 4 plot within the generic stable ocean region (SOR-GENERIC) and the active non-subduction zone hotspot region (ANSR-HOTSPOT). The two regions have significantly lower productivity parameters compared to the onshore

subduction zone (SZ-ONSHORE), which has an Omori *a*-value of -2.34. This means that the Alaskan subduction zone is less productive than other subduction zones included in the estimation of the Page et al. aftershock parameters. Sequences in the Alaska onshore region also return to background seismicity at a slightly faster rate than sequences in other onshore regions globally. To test the accuracy and robustness of my estimations, my methods should be repeated on the same aftershock sequences with new parameters. The parameters that might best represent the variation in Table 4 are an *a*-value of -3.03 ± 0.5 and a *p*-value of 0.82 ± 0.6. These parameters were estimated by taking the mean *a* and *p* values, and the uncertainty accounts for the inter-sequence variability.

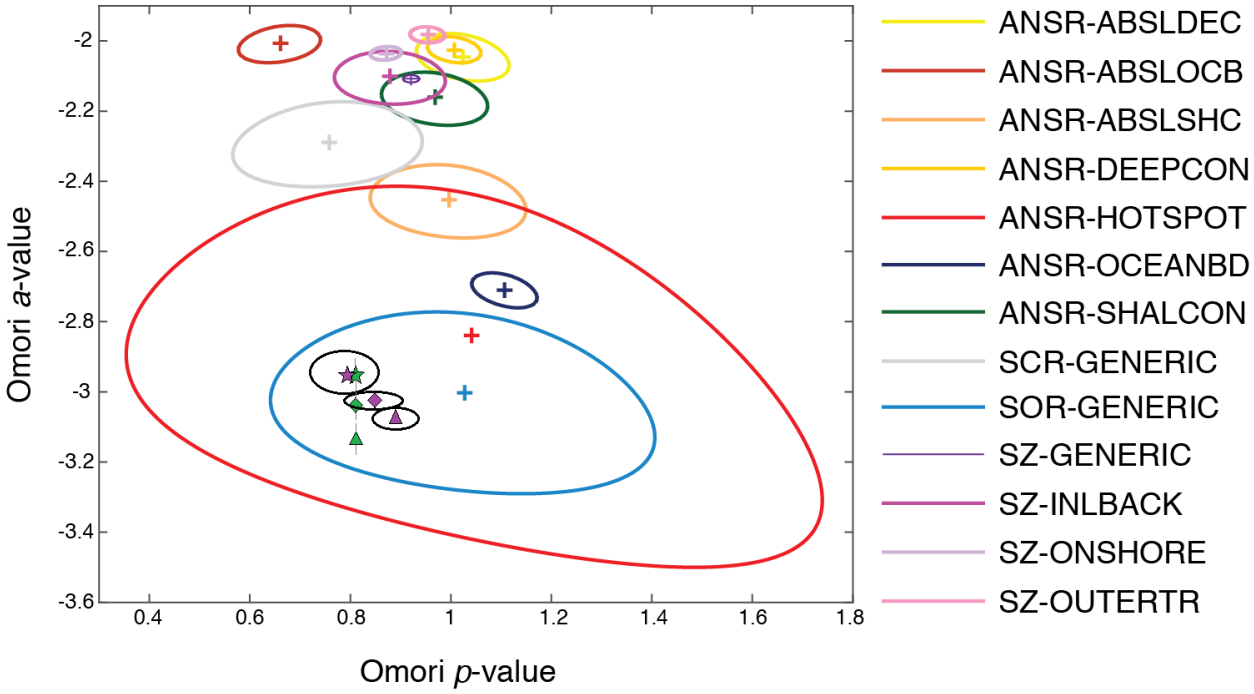


Figure 10. Aftershock parameters derived by Page et al. (2016) and this study. The proposed aftershock parameters using a Bayesian model (green) and a Sequence Specific model (purple) with respective uncertainties have a lower *a*-value compared to the maximum likelihood solution for the onshore subduction zone *a*-value (light purple cross, near the top of the plot). Triangles indicate parameters derived using a M_c of 3.5, diamonds derived with a M_c of 4, and stars derived using a M_c of 4.5. The figure is amended from Page et al. (2016) and shows Omori 95% confidence limits for each tectonic region defined by García et al. (2012), with the maximum likelihood solution as a cross. The uncertainty for the proposed *a*-value using a Bayesian model is a line because the aftershock rate *p* is fixed for the model.

Conclusion

Based on the analysis, the focus for accurate aftershock forecasting is on the forecast median and not necessarily on the confidence interval because a "successful" forecast can be

hidden inside a large confidence interval that does not reflect the true behavior of an aftershock sequence. A well-characterized aftershock forecast would see observations more randomly distributed on both sides of the forecast median, but the aftershock forecasts issued for earthquake sequences in Southern Alaska are systematically predicting more earthquakes than observed, with at least 77% of the observations below the forecasted median. The most likely *a* and *p* values for a stacked catalog are lower than the parameters defined by Page et al. (2016) for the onshore region. Therefore, the Alaskan onshore region is less productive than forecasted, and the aftershock sequences in the Alaskan onshore region return to background seismicity faster than other subduction zones. The new parameters I propose (*a*: -3.03, *p*: 0.82) for the Alaskan onshore subduction zone region provide an adequate starting place for further testing to refine the aftershock parameters in this region.

Data and Resources

The earthquake catalog and aftershock forecasts can be obtained from the Comprehensive Catalog (ComCat) at https://earthquake.usgs.gov/earthquakes/search/ (last accessed June 2022), and background information on the forecasts forecast parameters can be found at https://earthquake.usgs.gov/data/oaf/ (last accessed June 2022). The R Statistical Programming language was downloaded from https://www.r-project.org/ (last accessed June 2022). The USGS Aftershock GUI was downloaded from https://github.com/mbarall/oaf-gui-releases (last accessed June 2022). A public release of the oaftools package and the Operational Aftershock Viewer will be available on https://code.usgs.gov (last accessed June 2022) following the publication of the associated paper in Seismological Research Letters. The aftershock forecast schedule can be found at https://earthquake.usgs.gov/data/oaf/schedule.php (last accessed June 2022).

Acknowledgements

This thesis work was substantially improved by valuable discussion from Michael Barall, Jeanne Hardebeck, Ruth Harris, Andrea Llenos, Sara McBride, Andrew Michael, Morgan Page, Max Schneider, Nicholas van der Elst, and Anne Wein. I also thank Prof. Richard Hugo and Prof. Ashley Streig for the resources to do this analytical work.

References

Båth, M. (1965). Lateral inhomogeneities in the upper mantle, *Tectonophysics*, **2**, 483-514,

Garcia, D., Wald, D. J., & Hearne, M. G. (2012). A global earthquake discrimination scheme to optimize ground-motion prediction equation selection. *Bulletin of the Seismological Society of America*, *102*(1), 185–203. https://doi.org/10.1785/0120110124

Gutenberg, B. and Richter, C.F. (1944) Frequency of Earthquakes in California. *Bulletin of the Seismological Society of America*, 34, 185-188.

Helmstetter, A., Y. Y. Kagan, and D. D. Jackson (2006). Comparison of short-term and time-dependent earthquake forecast models for southern California, *Bull. Seism. Soc. Am.*, **96**, 90-106.

Michael, A. J., S. K. McBride, J. L. Hardebeck, M. Barall, E. Martinez, M. T. Page, N. van der Elst, E. H. Field, K. R. Milner, and A. M. Wein (2019). Statistical Seismology and Communication of the USGS Operational Aftershock Forecasts for the 30 November 2018 Mw 7.1 Anchorage, Alaska, Earthquake, *Seismological Research Letters*, **91**, 153-173, doi:10.1785/0220190196.

Mignan, A., J. Woessner (2012), Estimating the magnitude of completeness for earthquake catalogs, Community Online Resource for Statistical Seismicity Analysis, doi:10.5078/corssa-00180805. Available at http://www.corssa.org.

Page, M. T., van der Elst, N., Hardebeck, J., Felzer, K., & Michael, A. J. (2016). Three ingredients for improved global aftershock forecasts: Tectonic Region, time-dependent catalog incompleteness, and intersequence variability. *Bulletin of the Seismological Society of America*, *106*(5), 2290–2301. https://doi.org/10.1785/0120160073

Paris, G. M., & Michael, A. J. (2022). An Interactive Viewer to Improve Operational Aftershock Forecasts, *Seismological Research Letters*, 26 pp., submitted.

Reasenberg, P. A., & Jones, L. M. (1989). Earthquake hazard after a mainshock in California. *Science*, *243*(4895), 1173–1176. https://doi.org/10.1126/science.243.4895.1173

USGS Communications and Publishing. (2016). *Subduction fault zone diagram*. Subduction Fault Zone Diagram | U.S. Geological Survey. Retrieved June 9, 2022, from https://www.usgs.gov/media/images/subduction-fault-zone-diagram

Utsu, T. (1961) A Statistical Study on the Occurrence of Aftershocks. *The Geophysical Magazine*, 30, 521-605.

van der Elst, N. J., Hardebeck, J. L., Michael, A. J., McBride, S. K., & Vanacore, E. (2022). Prospective and retrospective evaluation of the U.S. Geological Survey Public Aftershock Forecast for the 2019–2021 Southwest Puerto Rico earthquake and aftershocks. *Seismological Research Letters*, *93*(2A), 620–640. https://doi.org/10.1785/0220210222