Appendices to Relationship Between M8+ Earthquake Occurrences and the Solar Polar Magnetic Fields

APPENDIX A to "Relationship Between M8+ Earthquake Occurrences and the Solar Polar Magnetic Fields": Description of Mathematical Model

The following is a description of the mathematical model used to identify periods of increased likelihood of large earthquakes (*i.e.*, earthquake windows). We begin by defining some of the notation used throughout the description. We use t as an index of days, starting with 1976-06-04 as $t = 1$. Let $t \in \{1,2,\dots,T\}$, where $T = 13741$, corresponding to 2014-01-17. Since solar polarity observations are only available every 10 days, we create a second time index, r, where $r = 1 + (\frac{t-1}{10})$ when $mod(t, 10) = 1$ and is undefined otherwise, so $r \in \{1,2,\dots,R\}$, where $R = 1375$. We use N_r to denote the magnitude of the north pole on day r and S_r to denote the magnitude of the south pole on day r in $\mu\mathrm{T}$.

Identifying Solar Maximum, Minimum, and Recovery

For each r, we first determine whether the sun is in a period of solar maximum, minimum, or recovery. Identifying these phases requires identification of the peaks and troughs of both the long-term cycles and minicycles of each of the poles. To identify peaks and troughs of the long-term cycles, we take a rolling average of the magnitude and identify the maximal (and minimal) points within 2000-day rolling windows. We define

$$A_{N,r} = \frac{\sum_{i=r-100}^{r+99} N_i}{200},$$

the rolling average of the north pole at time r,

$$M_N^1 = \left\{r: A_{N,r} = max_{i=r-100}^{r+99}\{A_{N,i}\}\right\},$$

the set of long-term local maxima for the north pole, and

$$M_N^2 = \left\{r: A_{N,r} = min_{i=r-100}^{r+99}\{A_{N,i}\}\right\},$$

the set of long-term local minima for the north pole. Long-term extremes of the south pole are defined similarly. For simplicity of notation throughout this manuscript, we have ignored the fact that indexes exceed the limits of the data ($r < 1$ or $r > R$). In the actual implementation, indexes and values that depend on them (like the denominator of A) are adjusted according to the number of available observations in a 2000-day window.

To identify mini-cycles, we fit local sine wave functions (trended, phase shifted, and scaled) with a period of 1 year at each time point and average across fitted values of several locally fitted functions. For the north pole, for each r we fit the following statistical model:

$$N_i \sim No(\beta_{0Nr} + \beta_{1Nr} i + \beta_{2Nr} \sin\left(\frac{2\pi 10 i}{365}\right) + \beta_{3Nr} \cos\left(\frac{2\pi 10 i}{365}\right), \sigma_{Nr}^2),$$

for $i \in \{r-100, r-99, \dots, r+99\}$, where $No(\cdot,\cdot)$ represents the normal distribution with specified mean and variance. Defining $\widehat{N_{rr'}}$ as the fitted value of N_r for the model at r', the smoothed mini-cycle is defined as

$$U_{Nr} = \sum_{i=r-99}^{r+100} \widehat{N_{ri}} /200.$$

The local maxima and minima of the mini-cycles are then defined as the local maxima and minima of U_{Nr} (*i.e.*, values of r for which the two adjacent points have greater (or lower, in the case of minima) magnitude).

For $r = 1$, the sun is, by definition, in the polar maximum phase. The time points $r = 2,3,\ldots,R$ are assessed sequentially with the phase being the same as the preceding point unless a transitional event happens. The transition from polar maximum to polar minimum occurs at the first time that the second of the two poles exhibits a reversal across the equator. Mathematically, if the current phase of polar maximum is one in which the north pole reached a maximum and the south pole reached a minimum, the transition to polar minimum occurs at $\max\{\min\{r': N_{r'} < 0\}, \min\{r': S_{r'} > 0\}\}$, where r' is later than the start of the polar maximum. The transition from polar minimum to polar recovery occurs at the last time that either of the poles exhibits a force directionally the same as its most recent maximum or minimum (and before the next long-term maximum or minimum). Again, if the most recent phase of polar maximum was one in which the north pole reached a maximum and the south pole reached a minimum, the transition to polar recovery occurs at $\max\{\max\{r': N_{r'} > 0\}, \max\{r': S_{r'} < 0\}\}$, where r' is later than the start of the polar minimum and earlier than the next long-term maximum or minimum. The transition from polar recovery to polar maximum always occurs 1.5 years after the transition to polar recovery.

Identifying Spikes

The values of U_{Nr} and U_{Sr} represent smoothed versions of the solar cycles that preserve both long-term trends and the approximately 1-year minicycles of the poles. The next step of the algorithm uses those smoothed representations to divide the modeled time span into segments representing half-minicycles. This segmenting is performed separately for each of the poles, and, for simplicity, we only describe the method for the north pole. For each adjacent pair of local maxima and minima in U_{Nr}, we identify a segmenting time point r where the value of N_r is as close as possible to the average of the polar force at the maximum and minimum. Mathematically, if r' and r'' represent a pair of time points for which U_{Nr} is at adjacent maximums and minimums, the segmenting time point is $argmin_r |N_r - \frac{1}{2}(N_{r'} + N_{r''})|$. These segmenting points divide the span into half-minicycles, and each half-minicycle is classified as convex or concave, depending on whether the local maximum within the half-minicycle is a maximum or minimum. To simplify further notation, we introduce a new pair of indexes, q_N and q_S, indexing half-minicycles in the north and south poles, respectively. For the time period in this study, there are 79 half-minicycles of the north pole and 75 half-minicycles of the south pole, so $q_N \in \{1,2,\ldots,79\}$ and $q_S \in \{1,2,\ldots,75\}$.

Within each half-minicycle, we identify spikes using a series of rules. For simplicity, these rules are described for concave half-minicycles (in the north pole); applying them to convex half-minicycles is just a matter of reversing the sign of the magnitudes of the poles. Within a half-minicycle q_N, define the set of time points in q_N as

$$Q_{q_N} = \{r: r \geq \mathit{the\ segmenting\ point\ beginning\ } q_N \mathit{\ and\ } r \leq \mathit{the\ segmenting\ point\ ending\ } q_N\}.$$

We first identify the time points of the maximal polar force within the half-minicycle $P_{q_N} = \{r: r \in Q_{q_N} \mathit{\ and\ } r \geq N_{r'}\ \forall\, r' \in Q_{q_N}\}$ and identify these points as spikes. Starting with the minimal value in P_{q_N}, we sequentially examine each smaller value in the set Q_{q_N} and assess two conditions:

1. Is the polar force at the examined time point more than 10 greater than the smallest polar force between the examined time point and the last identified spike?

2. Is the polar force of all remaining time points between the examined time point and the end of the set Q_{q_N} smaller than the polar force at the examined time point?

If both of these conditions are satisfied, the examined time point is identified as an additional spike within the half-minicycle and we proceed to evaluate the next smaller value in the set Q_{q_N}. Once these criteria have been applied to all time points from the smallest value in P_{q_N} to the smallest value in Q_{q_N}, the process is repeated starting with the maximal value in P_{q_N} and sequentially examining each larger value in the set Q_{q_N}.

At the end of applying the above procedure, each half-minicycle has, identified within it, one or more time points at which the polar force is spiking. We use P'_{q_N} to denote the full set of spike times within the half-minicycle q_N. Note that $P_{q_N} \subseteq P'_{q_N}$. For each r in P'_{q_N}, the spike is classified as primary or secondary and as being peaked, semi-flat, or completely flat. A spike is classified as secondary if the polar force at the time of the spike is at least 10 μT less than the magnitude of the maximal force exhibited in q_N. To determine whether a spike is peaked, semi-flat, or completely flat, we examine the polar force for all adjacent time points with a polar force no more than 10 μT less than the polar force at the spike. If at least one of the adjacent time points has a force no more than 3 μT less than the polar force at the spike, it is classified as semi-flat. If any of the adjacent time points has a force exactly equal to the spike, it is classified as completely flat.

After identifying all of the spikes and classifying them as primary/secondary and peaked/semi-flat/completely flat, we define each half-minicycle as being a single, double, triple, or quadruple period. This classification is made based solely on the number of primary spikes in the half-minicycle.

Identifying Earthquake Windows

Define a function $f(r) = 1 + 10(r-1)$ that converts the r index of 10-day periods to a corresponding day index in $\{1, 2, \ldots, T\}$. Also, define the average polarity at time r as $B_r = (N_r + S_r)/2$. Then, a day t is in an earthquake window if any of the following conditions are met:

1. There exists a spike at some time r' in some set P'_{q_N} such that (a) r' is in a period of solar maximum, (b) the half-minicycle q_N is a single or double period, (c) the spike at time r' is peaked and primary, (d) $N_{r'}$ has the same sign as $N_{min(Q_{qn})}$ and $N_{max(Q_{qn})}$, (e) $\left|N_{\min(P_{qn})} - N_{\min(P_{qn-1})}\right| \geq 40\mu\mathrm{T}$, (f) $\left|N_{\min(P_{qn})} - N_{\min(P_{qn+1})}\right| \geq 40\mu\mathrm{T}$, and (g) $|f(r') - t| \leq 20$.
2. There exists a spike at some time r' in some set P'_{q_N} such that (a) r' is in a period of solar maximum, (b) $|N_{r'}| > 175\mu\mathrm{T}$, (c) the half-minicycle q_N is a single period, (d) the spike at time r' is peaked and primary, (e) $N_{r'}$ has the same sign as $N_{min(Q_{qn})}$ and $N_{max(Q_{qn})}$, (f) $\left|N_{\min(P_{qn})} - N_{\min(P_{qn-1})}\right| \geq 40\mu\mathrm{T}$, (g) $\left|N_{\min(P_{qn})} - N_{\min(P_{qn+1})}\right| \geq 40\mu\mathrm{T}$, and (h) $|f(r') - t| \leq 25$.
3. There exists a spike at some time r' in some set P'_{q_N} such that (a) r' is in a period of solar maximum, (b) the half-minicycle $q_N - 1$ is a single period with some r'' such that $|N_{r''}| > 175\mu\mathrm{T}$, (c) the half-minicycle $q_N + 1$ is a single period with some r''' such that $|N_{r'''}| > 175\mu\mathrm{T}$, (d) the half-minicycle q_N is a single period, (e) $N_{r'}$ has the same sign as $N_{min(Q_{qn})}$ and

$N_{max(Q_{q_n})}$, (f) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n-1})}\right| \geq 40\mu\text{T}$, (g) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n+1})}\right| \geq 40\mu\text{T}$, and (h) $|f(r') - t| \leq 25$.

4. There exists a spike at some time r' in some set P'_{q_N} such that (a) r' is in a period of solar maximum, (b) $|S_{r''}| > 175\mu\text{T}$, where $r'' = \max\{r: r < r' \; and \; r \in P'_{q_S} \; for \; some \; q_S\}$, (c) $|S_{r'''}| > 175\mu\text{T}$, where $r''' = \min\{r: r > r' \; and \; r \in P'_{q_S} \; for \; some \; q_S\}$, (d) the half-minicycle q_N is a single period, (e) $N_{r'}$ has the same sign as $N_{min(Q_{q_n})}$ and $N_{max(Q_{q_n})}$, (f) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n-1})}\right| \geq 40\mu\text{T}$, (g) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n+1})}\right| \geq 40\mu\text{T}$, and (h) $|f(r') - t| \leq 25$.
5. There exists a spike at some time r' in some set P'_{q_N} such that (a) r' is in a period of solar maximum, (b) the half-minicycle q_N is a triple period, (c) the spike at time r' is peaked and primary, (d) $N_{r'}$ has the same sign as $N_{min(Q_{q_n})}$ and $N_{max(Q_{q_n})}$, (e) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n-1})}\right| \geq 40\mu\text{T}$, (f) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n+1})}\right| \geq 40\mu\text{T}$, and (g) $|f(r') - t| \leq 15$.
6. There exists a spike at some time r' in some set P'_{q_N} such that (a) r' is in a period of solar maximum, (b) the half-minicycle q_N is a quadruple period, (c) the spike at time r' is peaked and primary, (d) $N_{r'}$ has the same sign as $N_{min(Q_{q_n})}$ and $N_{max(Q_{q_n})}$, (e) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n-1})}\right| \geq 40\mu\text{T}$, (f) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n+1})}\right| \geq 40\mu\text{T}$, and (g) $|f(r') - t| \leq 10$.
7. There exists a spike at some time r' in some set P'_{q_N} such that (a) r' is semi-flat, (b) $N_{r'}$ has the same sign as $N_{min(Q_{q_n})}$ and $N_{max(Q_{q_n})}$, (c) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n-1})}\right| \geq 40\mu\text{T}$, (d) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n+1})}\right| \geq 40\mu\text{T}$, and (e) $|f(r') - t| \leq 10$.
8. There exists a spike at some time r' in some set P'_{q_N} such that (a) r' is completely flat, (b) $N_{r'}$ has the same sign as $N_{min(Q_{q_n})}$ and $N_{max(Q_{q_n})}$, (c) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n-1})}\right| \geq 40\mu\text{T}$, (d) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n+1})}\right| \geq 40\mu\text{T}$, and (e) $|f(r') - t| \leq 15$.
9. There exists a spike at some time r' in some set P'_{q_N} such that (a) r' is secondary, (b) $N_{r'}$ has the same sign as $N_{min(Q_{q_n})}$ and $N_{max(Q_{q_n})}$, (c) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n-1})}\right| \geq 40\mu\text{T}$, (d) $\left|N_{\min(P_{q_n})} - N_{\min(P_{q_n+1})}\right| \geq 40\mu\text{T}$, and (e) $|f(r') - t| \leq 10$.
10. There exists some time r' such that (a) $sign(N_{r'}) = -sign(N_{r'-1})$, (b) there exists some $r'' \in [\max\{r: r < r' \; and \; sign(N_r) = -sign(N_{r-1})\}, r' - 1]$ such that $|N_{r''}| > 15\mu\text{T}$, (c) there exists some $r''' \in [r' + 1, \min\{r: r > r' \; and \; sign(N_r) = -sign(N_{r+1})\}]$ such that $|N_{r'''}| > 15\mu\text{T}$, and (d) $|f(r') - t| \leq 15$.
11. There exists some time r' such that (a) $r' - 1$ is during solar maximum, (b) $sign(N_{r'}) = -sign(N_{r'-1})$, (c) there is no $r'' < r'$ in the same period of solar maximum as $r' - 1$ for which $sign(N_{r''}) = -sign(N_{r''-1})$, and (d) $|f(r') - t| \leq 15$.
12. There exists some time r' such that (a) solar minimum ends at r' and (b) $|f(r') - t| \leq 15$.
13. There exists some time r' such that (a) $sign(B_{r'}) = -sign(B_{r'-1})$, (b) either the phase is solar maximum at time r' or the following two conditions are satisfied

a. there exists some $r'' \in [\max\{r: r < r' \; and \; sign(B_r) = -sign(B_{r-1})\}, r' - 1]$ such that $|B_{r''}| > 50\mu\text{T}$

b. there exists some $r''' \in [r' + 1, \min\{r: r > r' \; and \; sign(B_r) = -sign(B_{r+1})\}]$ such that $|B_{r'''}| > 50\mu\text{T}$,

(c) $\min\{r: r > r' \; and \; sign(B_r) = -sign(B_{r+1})\} - r' \leq 9$, and (d) $|f(r') - t| \leq 15$.

For conditions 1 – 11, identical conditions using the south pole also apply.

The above conditions determine whether each time point falls into an earthquake window, but the occurrence of earthquakes can reduce the likelihood of a subsequent earthquake, requiring shifting of windows that occur soon after. Each time a large earthquake (8+ magnitude) or series of slightly smaller earthquakes (4 7.5+ magnitude quakes in a 100 day window) occurs, earthquake windows occurring in the next 100 days are delayed. For earthquake windows starting less than 30 days after the occurrence of one of these earthquake events, the window is pushed back by the minimum of either 20 days or half the width of the interval for the condition above that created the earthquake window. For earthquake windows starting between 31 and 100 days after the occurrence of one of these earthquake events, the window is pushed back by 10 days.