Air Force Institute of Technology

AFIT Scholar

3-2-2006

Prediction of the Temporal Evolution of Solar X-Ray Flares

Aaron J. Williams

Follow this and additional works at: https://scholar.afit.edu/etd

Part of the Engineering Physics Commons

PREDICTION OF THE TEMPORAL EVOLUTION OF SOLAR X-RAY FLARES

THESIS

Aaron J. Williams, Captain, USAF

AFIT/GAP/ENP/06-21

DEPARTMENT OF THE AIR FORCE
AIR UNIVERSITY
AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the United States Government.

PREDICTION OF THE TEMPORAL EVOLUTION OF SOLAR X-RAY FLARES

THESIS

Presented to the Faculty

Department of Engineering Physics

Graduate School of Engineering and Management

Air Force Institute of Technology

Air University

Air Education and Training Command

In Partial Fulfillment of the Requirements for the

Degree of Master of Science (Applied Physics)

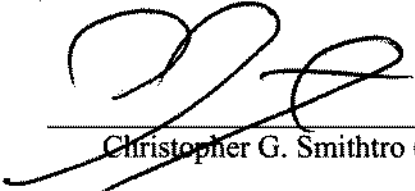Aaron J. Williams, BS

Captain, USAF

March 2006

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

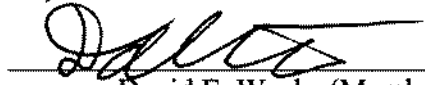PREDICTION OF THE TEMPORAL EVOLUTION OF SOLAR X-RAY FLARES

Aaron J. Williams, BS
Captain, USAF

Approved:

_____
Christopher G. Smithtro (Chairman)

2 Mar 06
_____
date

_____
David E. Weeks (Member)

2 Mar 06
_____
date

_____
Thomas A. Niday (Member)

2 Mar 06
_____
date

Abstract

A solar flare is an explosive release of stored magnetic energy on the Sun. Much of this energy is converted into x-ray photons which escape into space. As a solar flare begins, the 1-8 Å x-ray photon flux at Earth's orbit, as measured by the GOES satellite, rapidly increases. It quickly reaches a peak and slowly decays. A plot of this flux exhibits an approximate lognormal shape. A lognormal function becomes a normal, symmetric, function when the logarithm of the independent variable is taken. Once the peak flux is reached, this symmetry is used to make a prediction of the flare end time. Examining over 1300 flares, an improvement in the flare end time prediction over the current climatological method is demonstrated. Predictions of the evolution of the flux prior to reaching the peak flux are also made beginning five minutes after flare onset. An effort to predict the temporal evolution prior to the peak is made using a fourth order least squares fit to the rise-phase flux alone and the rise-phase flux plus an estimate of the decay flux. Using both methods, accurate predictions of the x-ray flux evolution are made when the rise phase averages 65% complete.

Acknowledgements

I would like to express my utmost gratitude to my thesis advisor for his guidance and insight throughout this thesis effort. I would also like to thank my thesis committee for their effort, time, and willingness. A very special thanks to my wife and kids for allowing me the time to work on this thesis and for their constant support.

Aaron J. Williams

Table of Contents

vii

List of Figures

List of Tables

PREDICTION OF THE TEMPORAL EVOLUTION OF SOLAR X-RAY FLARES

I.   Introduction

A solar flare is defined as a sudden explosion in the sun's atmosphere. Flares can last from minutes to a couple of hours [*Cravens*, 1997]. They heat the sun's atmosphere to millions of Kelvins, release approximately 10^{25} Joules of energy, and accelerate electrons, protons, and ions to speeds nearing the speed of light. These particles can cause extensive damage to humans and spacecraft in orbit around earth.   Flares also release electromagnetic radiation across the entire spectrum. X-rays, as well as extreme ultraviolet (EUV) waves, with wavelengths from 1 to 8 Å, penetrate deep into earth's atmosphere, down to 70-100 km above the surface. This region is known as the D-region of the ionosphere; here the x-rays ionize molecules to create free electrons.

During a solar flare, more x-rays than normal enter the D-region, causing an increase in ionization and therefore an increase in free electrons on the sunlit side of the earth. This increase will have an effect on the propagation of radio wave signals. A high frequency (HF) communications signal (3-30 MHz) traveling through the atmosphere encounters changing electron densities. As the signal passes through these variations, the path is bent according to Snell's Law, given by equation (1).

$$n_i \sin \theta_i = n_r \sin \theta_r \tag{1}$$

1

Where n is the refractive index and θ is the angle. The subscript i denotes the incident conditions and the subscript r the refracted conditions. Figure 1 shows an example of how Snell's Law applies for different indices and angles. A signal traveling into an area with a higher index of refraction will be bent toward the vertical. The same signal will be bent away from the vertical when moving into a region with a lower index of refraction.
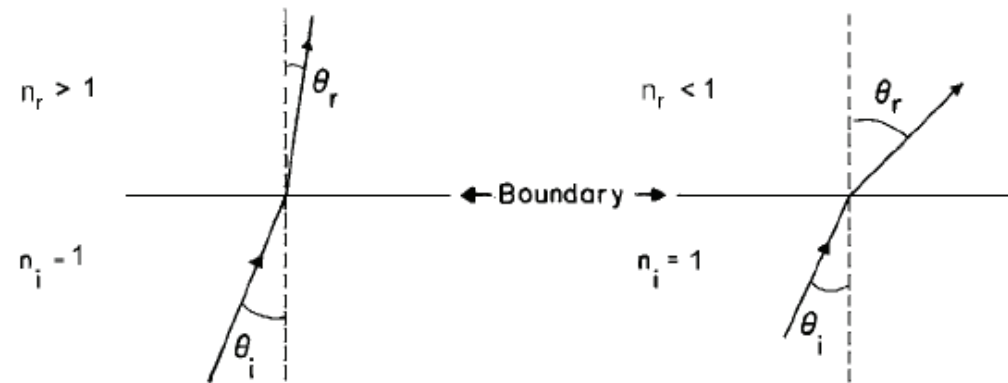
Figure 1. When traveling into an area with a higher index of refraction the signal will be bent toward the vertical. It will be bent away from the vertical when going into a lower index [Air Force Institute of Technology, PHYS 519 class notes].

As the electron density increases the index of refraction decreases. Figure 2 illustrates the path of an HF signal as it moves into the F-region ionosphere (200 – 300 km), a region of higher electron densities. The path can be bent enough that it returns to earth, allowing for long range communications. The higher electron densities allow for a lower index of refraction so the signal will be bent away from the vertical. Once the signal is bent enough so that θ reaches 90° the signal is reflected back down to earth.

2
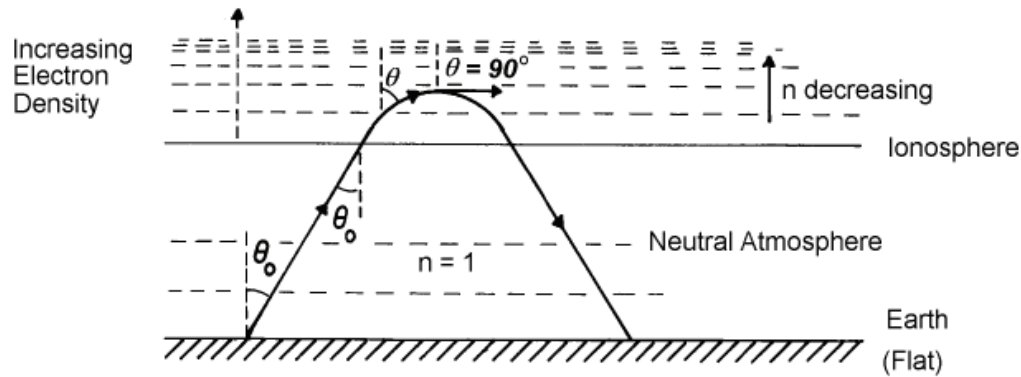
Figure 2. A decreasing index of refraction will lead to the bending of an HF signal in the ionosphere. The bending will reach a point where the angle is 90° and the signal will be reflected back to earth [Air Force Institute of Technology, PHYS 519 class notes].

During a solar flare the electron density in the D-region (below the F-region) increases due to the increased flux of x-rays. This increase in electron density will cause HF signal paths to be bent at lower altitudes than normal. Because the signal is bent at a lower altitude, it spends more time in the lower altitudes, where the collision frequency is higher. Collisions between electrons and neutrals convert the signal energy into heat and dissipate the signal. The combination of these two processes leads to the HF signal not propagating as expected and causing it to be absorbed in the ionosphere. This is known as Short Wave Fade (SWF) and can cause a significant signal degradation at frequencies between 3 and 30 MHz. The loss of communications is most noticeable on the sunlight side of the earth and begins within minutes of flare occurrence. HF communications return to normal shortly after the flare ends, when the increased flux of x-rays decreases. The electron density rapidly returns to normal as they re-combine with positive ions present in the ionosphere.

The ability to forecast the end of the SWF event would allow users of these frequencies to have an idea when communications will return to normal. The duration of

the communication outage is tied closely to the duration of the flare. So, if the duration

of the flare can be forecast, then the duration of the communication outage will be known

as well. The Space Environment Center (SEC) creates the official civilian forecasts for

space weather events to include solar flare duration. The current method of flare duration

forecasting used by the SEC is dependent on the peak x-ray flux of the flare. Once the

peak flux is reached during a flare, the SEC estimates the time for the x-ray flux, and

hence the SWF event, to return to background levels based on the value of the peak flux.

Table 1 shows the peak flux and the estimated (climatological) recovery time for the flux

to return to background levels. For example, if the peak flux is 5×10^{-5} W/m^2 then normal

background conditions are expected within 40 minutes. Based on this table, the larger the

peak flux, the longer it should take the x-ray flux to decay and for the atmosphere to

recover back to normal.

Table 1. Estimated recovery time of x-ray flux to normal background conditions.

Peak Flux (W/m^2)	Duration (Minutes)
1×10^{-5}	25
5×10^{-5}	40
1×10^{-4}	60
5×10^{-4}	120

These times relate the peak flux of a flare to the statistical average of flare duration.

These values and more information can be found at the SEC website

(http://www.sec.noaa.gov/rt_plots/dregionDoc.html). Others have devised methods to

predict the duration of flares as well. *Tobiska and Bouwer* [2005] have created a solar

flare evolution model that uses a combination of Gaussian and quadratic functions to

4

predict not only the flare duration, but the magnitude and time of the peak. *Bornmann* [2000] suggested using a lognormal function to predict the flare duration once the peak flux is reached.

When a flare begins there is a rapid increase in the x-ray flux at earth. Once this flux reaches a peak it slowly decays back toward normal background conditions. A plot of this behavior shows an apparent lognormal shape. Assuming the flare evolution is approximately lognormal, a prediction can be made for the decay phase based on the rise phase. The lognormal, discussed in detail later, has the characteristic of becoming a normal function when the logarithm of the independent variable is taken. In the case of the x-ray flux the independent variable is time, so this will be referred to as log-time from here on. The symmetry of the normal can therefore be used to forecast the duration of the flare. In log-time coordinates, the x-ray flux rise phase data can be flipped over, creating a mirror image of the data points. These flipped data points are a prediction of the decay phase. A normal curve is then fit to these data points and converted back to real time to create a forecast of the decay phase based solely on the rise phase. This forecast is created once the x-ray flux has reached its peak.

Predicting the flare evolution prior to the peak provides a greater challenge. Working in log-time so that the data is symmetric, the data available on the rise phase can be moved over to the decay portion of the flux as an estimate of what the decaying flux will be. Fitting a fourth order polynomial to these points, and converting back to real time, creates a curve that represents the x-ray flux and gives values for the peak flux magnitude and time as well as the duration of the flare. This can be done prior to reaching the peak flux.

5

Rather than move the rise phase data points over to estimate the decay, another method is used that deals strictly with the rise phase data. A fourth order polynomial can be fit to the rise data, and converted back to real time, to get the x-ray flux prediction. This method gives the resulting polynomial fit more freedom since there are no estimated decay points to constrain it as in the method above.

All three methods are used to produce predictions of flare duration throughout the evolution of the x-ray flux. X-ray flux data is available in one minute increments. Five minutes after the flare begins was chosen arbitrarily as the beginning point for predictions. Each minute a new prediction is created using the two fourth order polynomial methods. Once the peak flux is reached the lognormal fit is used. While all three methods are used, only one prediction is output per minute. Using multiple methods allows for the best possible results depending upon how much data is available as the x-ray flux evolves.

Background on both normal and lognormal functions will be given in the following section. One method of x-ray creation from solar flares as well as how the x-rays are measured and the flares classified will also be discussed. Section III shows how, for the majority of flares, the x-ray flux during a flare produces an approximate lognormal shape. Section IV will focus on the three different forecasting methods used and the error calculation between the forecast and the actual flux values. Section V will break out the results of each forecasting method. The conclusions that can be drawn from this work will be laid out in the final section.

6

II. Background

The processes that create x-rays during a solar flare can give an insight into why the x-ray flux evolution at earth takes on a lognormal shape. A lognormal shape is created by a large number of independent events that occur in a given order or simultaneously [*Aitchison and Brown,* 1957]. This chapter will explain the lognormal function, briefly explore the creation of x-rays in a solar flare and how they relate to the lognormal function, and how solar flares are classified.

2.1 The Normal Function

The normal or Gaussian is a familiar function and is defined by equation (2).

$$f(x) = A \exp\left[-\frac{(x-\mu)^2}{\sigma^2}\right] \tag{2}$$

The amplitude is given by A, μ is the mean, and σ is the standard deviation. It is symmetric about the mean; it also has a peak and a standard deviation. Figure 3 shows a normal function with a peak of 1×10^{-7}, a mean (μ) of 0.00 and a calculated standard deviation (σ) of 1.52.
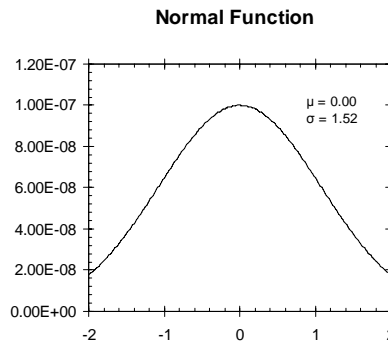
Figure 3. A normal function with a mean of 0.00 and a standard deviation of 1.52.

The normal function is created by a large number of small effects that act additively. Conversely, when a large number of small effects act multiplicatively, a lognormal distribution is formed. Figure 4 shows a physical model of the creation of a normal and lognormal distribution. In panel (a) the tip of the triangles are a distance x from the left edge and the tips of the triangles below and to the right and left are at x+a and x-a. This gives an additive effect, shifting the balls equally to the left or right, as the balls fall through the triangles to create the normal distribution. For the lognormal distribution the tips of the triangles are at x•b and x/b. This produces a multiplicative effect as the balls fall and creates a lognormal distribution.
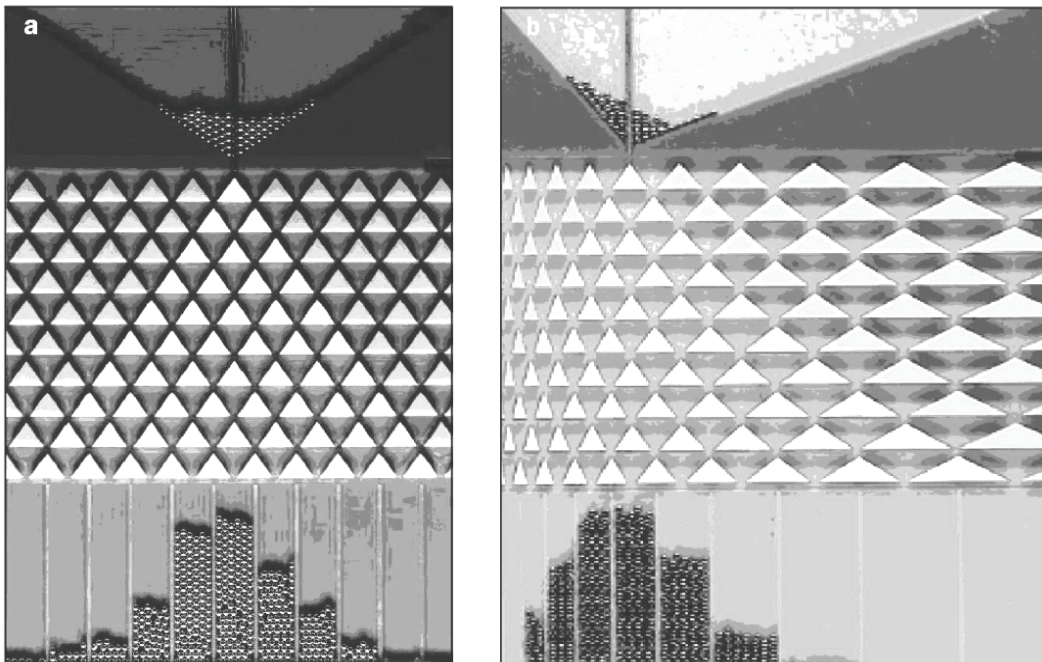
Figure 4. A physical model of the creation of both normal (a) and lognormal (b) distributions. If the tip of a triangle is a distance x from the left edge, then the distance to the tips of the triangles to the right and left are x+a and x-a for panel (a) and x • b and x/b for panel (b) where a and b are constants. As the balls fall through the triangles, in panel (a) the x+a and x-a produces an additive effect that creates the normal distribution at the bottom. In panel (b) the x • b and x/b produces a multiplicative effect that leads to the lognormal distribution at the bottom [*Limpert, et al,* 2001].

8

2.2 *The Lognormal Function*

The lognormal function is defined by equation (3), with A as the amplitude, μ as the mean, and σ as the standard deviation of the function.

$$f(x) = A \exp\left[\frac{-(\ln(x) - \ln(\mu))^2}{\sigma^2}\right] \tag{3}$$

In our case, the amplitude, A, is the value of the peak flux of the flare. This value will change based on the strength of the flare. The mean, μ, is the time of the peak flux. This value will vary depending on how quickly, or slowly, the x-ray flux increases to the peak flux value. As with a normal, the standard deviation, σ, determines the shape of the function. If the standard deviation is small, the data falls off quickly from the peak and the curve will be narrow. A large standard deviation makes the curve wider. In the case of the x-ray flux, the standard deviation will change depending on how quickly the flux rises and decays. A slow rise and decay will lead to a wider curve than a quick rise and decay.

A lognormal function reverts to a normal function when the logarithm of the independent variable is plotted. Figure 5 shows how the lognormal function becomes normal when plotted on a log scale. The effects of a changing mean and standard deviation can be seen. The peak remains constant for all plots. As the standard deviation changes so does the width of the curve. The top panels demonstrate this, with the larger standard deviations creating wider curves.

The lognormal function results when the effects of a large number of independently acting forces are multiplicative. In comparison, the normal function results when the forces are additive. The lognormal occurs often in nature. Some examples

from *Aitchison and Brown* [1957] are annual income size distribution, human body

weights, age of women and men at first marriage, and distribution of stars. An example

is given in Figure 6. This shows the concentration of the chemical

hydroxymethlyfurfurol (HMF) in 1573 samples of honey.
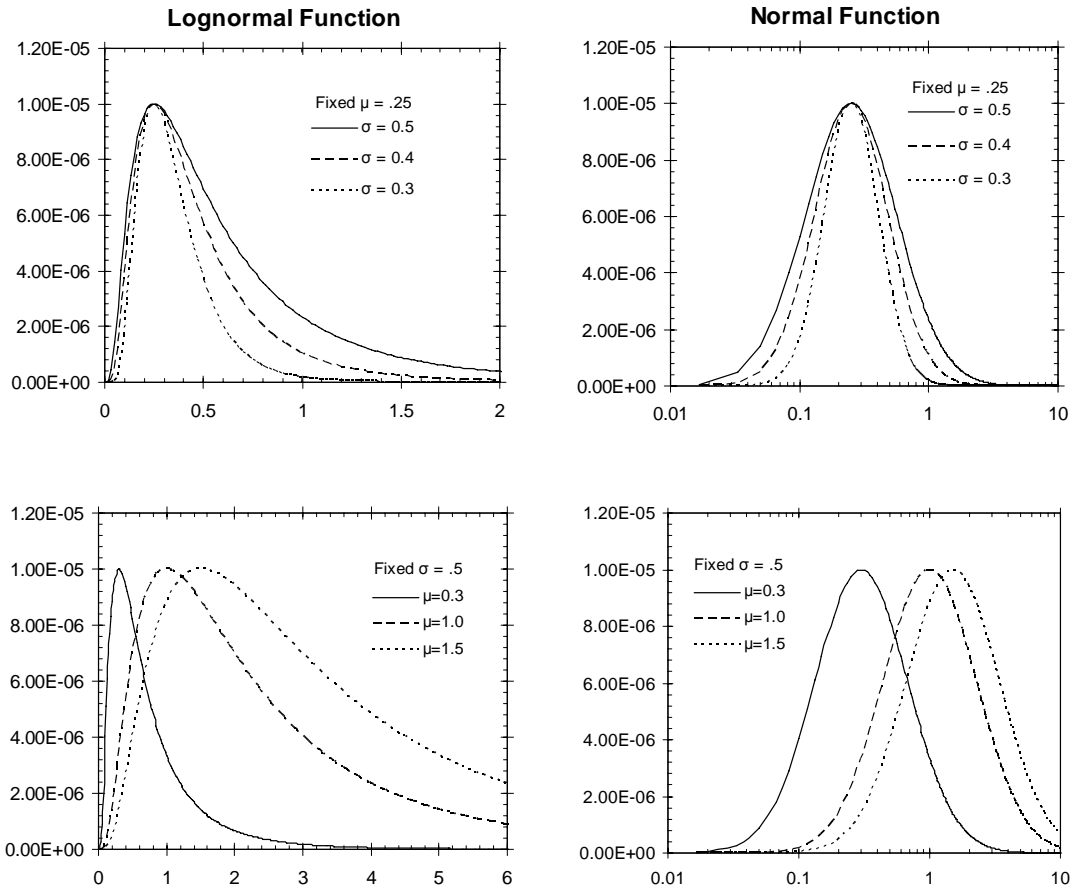
**Figure 5. The lognormal function with a fixed mean (top) and a fixed standard deviation (bottom).
The figures on the right have a log scale on the x-axis, transforming the lognormal shape to a normal.**

Because the logarithm of zero or a negative number is undefined, the independent

variable cannot take on values less than or equal to zero if a translation to a normal

function is desired.
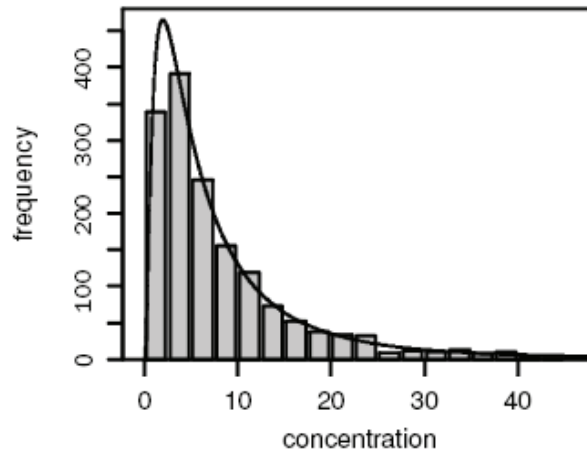
Figure 6. The concentration of the chemical HMF in 1573 samples of honey follows a lognormal distribution. The lognormal curve is the solid black line [*Limpert, et al.,* 2001].

This will not be a problem for flares since the independent variable is time and we can set the start time of the flare to zero. You can also do a coordinate translation of the independent variable without affecting the function.

2.3   Solar Flares

Lu and Hamilton [1991] argue that a solar flare is actually composed of many small events. It is possible that these small events combine multiplicatively to produce the lognormal shape found in the plot of the x-ray flux. In order to understand what those events are, a brief description of how a solar flare is formed follows.

The sun's atmosphere can be separated into three regions, the photosphere, the chromosphere, and the corona. The photosphere is the visible portion of the sun and is approximately 200 km thick, has an average temperature of 5700 K, and is where sunspots form. The chromosphere is an irregular layer above the photosphere whose temperature rises to 20,000 K. Above the chromosphere is the corona; this is the outer

11

portion of the atmosphere where temperatures can reach 2 million Kelvin [*Phillips*, 1992]. All of these layers are involved during a solar flare.

There are many models of solar flare creation. *Lang* [2000] provides an explanation for solar flares that addresses how x-rays are created. Solar flares usually occur over photospheric active regions. Magnetic loops extend from the photosphere into the lower corona. Magnetic reconnection occurs above the magnetic loop, releasing stored magnetic energy. This accelerates electrons to high speeds, which generates a burst of radio energy and hard x-rays (< 1 Å) via bremsstrahlung. Figure 7 shows this process. Some of the electrons are funneled down the loop into the chromosphere, creating hard x-rays through electron-ion bremsstrahlung. The material in the chromosphere is heated rapidly and rises up the loop where soft x-rays are created through electron-ion bremsstrahlung.
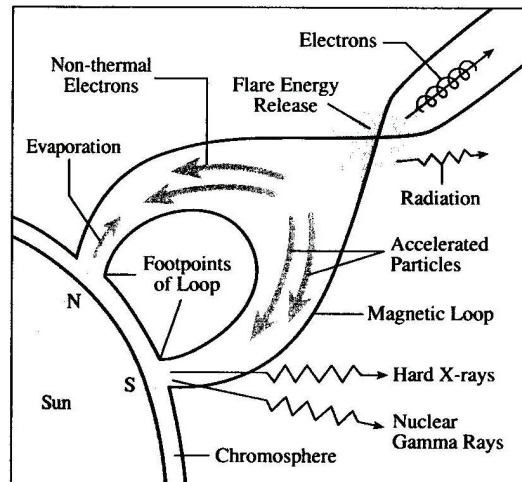
Figure 7. As magnetic reconnection occurs in the solar atmosphere, energy is released. This energy accelerates electrons and gives a quick burst of hard x-rays and radio energy [*Lang*, 2001].

Since bremsstrahlung is the method for creating x-rays, a further explanation is given. As the chromosphere is heated it releases the electrons from the atoms. These

www.manaraa.com

free electrons will eventually move near a proton in the surrounding gas. Since the

protons are much larger than the electrons they move slower so you can assume that the

proton is at rest compared to the electron. As the electron draws near the proton, it is

accelerated, thereby emitting electromagnetic radiation known as bremsstrahlung

radiation. At the electron's closest approach to the proton, the acceleration is the

greatest; the greater the acceleration, the shorter the wavelength of the emitted radiation.

This process is depicted in Figure 8. The range of bremsstrahlung emitted can be from

radio waves to x-rays. For this scenario the radiation emitted at this point is in the x-ray

wavelength. During a solar flare the x-ray wavelengths become very intense.
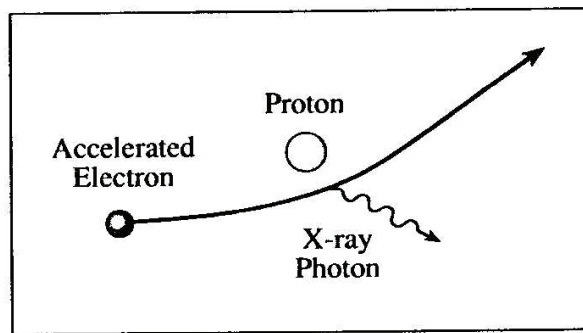
Figure 8. The energized electron's path and speed is altered by the proton. The reduction in speed produces an x-ray photon [*Lang*, 2001].

Lu and Hamilton [1991] argue that the coronal magnetic field is like a pile of

sand. When you add sand to a sandpile, the grains of sand added will cause a chain

reaction of small avalanches that will keep the sandpile slope stable. The magnetic

reconnection events that make up a solar flare correspond to these small avalanches. The

random twisting of the magnetic field by convective motions in the photosphere is the

addition of sand grains. Eventually there will be enough twisting to cause an instability,

and reconnection will occur and energy will be released. This reconnection and energy

13

release will lead to a cascade of reconnection events and energy release. This chain reaction process repeats itself until an equilibrium is reached. This is analogous to the avalanche of the sand that keeps the sandpile slope stable. The chain reaction of reconnections and energy release is the solar flare. This chain reaction of reconnections may be the large number of events that occur to create a lognormal shape of the flux.

2.4   The GOES Satellites

The GOES (Geostationary Operational Environmental Satellite) fleet of satellites are primarily terrestrial weather satellites, but they also carry an instrument package to monitor the space environment. This instrument package is called the SEM (Space Environment Monitor). The SEM is made up of a Magnetic Field Sensor, an Energetic Particle Sensor (EPS)/High Energy Proton and Alpha Detector (HEPAD), and a Solar X-Ray Sensor. The Solar X-Ray Sensor measures real-time values of the integrated solar x-ray emission in two spectral bands: 0.5-4 Å (hard x-ray) and 1-8 Å (soft x-ray). The SEM is on the GOES 8, 9, 10, 11, and GOES 12 satellites giving each the capability to monitor x-ray emissions from the sun. Currently the GOES 12 is the primary satellite used to monitor the x-ray flux from the sun. An archive of the solar x-ray flux data from the GOES series of satellites is available from the National Geophysical Data Center (NGDC) Space Physics Interactive Data Resource (SPIDR) website (http://spidr.ngdc.noaa.gov/spidr/) at 1 and 5 minute resolution.

The sun is constantly emitting x-rays; during a flare the amount of x-rays being emitted increases. The x-ray flux measured by the GOES is a disk integrated value, so even when no solar flares are present there are still x-rays being detected. These x-rays

14

constitute the background flux. When a flare occurs, a rapid rise in the flux above the background is seen. Figure 9 shows the 3-day x-ray flux plot beginning 29 Nov 2005. This plot shows both hard (0.5 to 4.0 Å) and soft (1.0 to 8.0 Å) x-rays detected by two different satellites. The soft x-rays are the line higher up on the graph and the hard x-rays are near the bottom of the graph. A non-flare background flux is clear at the beginning of the plot. This is a period where minimal solar activity is taking place. At approximately 1700 UTC on 29 Nov a flare begins, evident by the rapid rise in the flux. This rise is followed by a slow decay back toward the background flux levels. More flares are seen through the end of the plot.
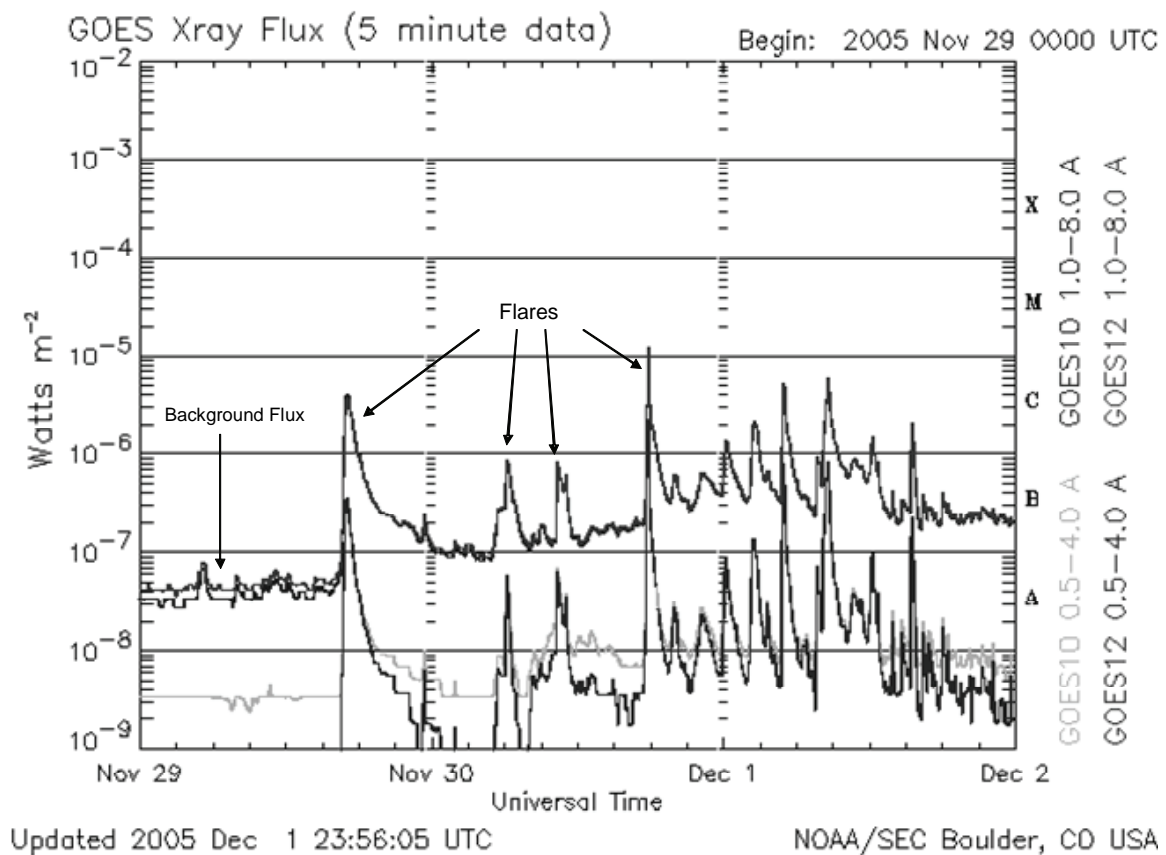
Figure 9. 3-day plot of x-ray flux. The background flux can be seen on the left of the image and a few of the flares have been identified [*NOAA/SEC*].

2.5 *X-Ray Flux and Flare Classification*

The x-ray flux that is measured is the disk integrated flux from the disk of the sun facing the satellite. When a flare occurs anywhere on the disk, the total amount of x-rays leaving the sun increases rapidly then slowly decays over time. Since this is an integrated flux, it is difficult to distinguish between two flares should they occur simultaneously. Another limitation is that the x-ray sensor reaches saturation at 1.74×10^{-3} W/m^2. So any flare exceeding this value will not be accurately depicted.

Solar flares are classified based on the peak magnitude of the x-ray flux that is observed at the satellite. Table 2 gives a breakdown of the x-ray classifications.

Table 2. X-ray Classification of Flares.

Classification	Flux (W/m^2)
A	10^{-8}
B	10^{-7}
C	10^{-6}
M	10^{-5}
X	10^{-4} and greater

A peak flux of 9×10^{-6} W/m^2 corresponds to a C9 flare and a peak flux of 1×10^{-5} W/m^2 would be a M1 flare. There are no letter identifiers beyond X, so a peak magnitude of 1.1×10^{-3} W/m^2 would be an X11. Figure 10 shows an example of a flare from 03 Nov 2004. The peak of the flare is 5×10^{-5} W/m^2, which makes this an M5 flare. Operationally, a flare is considered over when the flux returns to what is called the half-max point. The half-max point is defined as the average of the beginning and peak flux. Once the flux decays back to this half-max value it is considered operationally over. This is used because the x-ray activity on the sun doesn't always allow the flux to return back to pre-flare levels.

16

X-ray Flux

6.00E-05

5.00E-05

4.00E-05

3.00E-05

2.00E-05

1.00E-05

0.00E+00

Flux (W/m^2)

0   0.2   0.4   0.6   0.8   1

Duration (Hours)
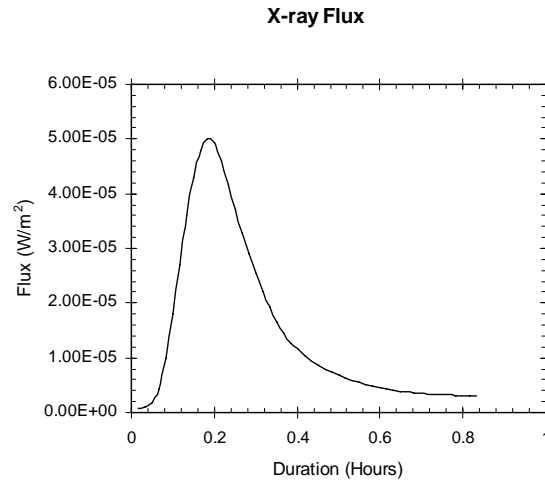
Figure 10. A peak flux of 5x10^{-5} W/m^2 gives an M5 classification for this flare from 03 Nov 2004.

After examining a plot of the x-ray flux in Figure 9 and Figure 10 it appears that, in general, the shape of the curve that is traced out when a solar flare occurs is approximately lognormal. This property of the x-ray flux can assist us in being able to predict how long the increased x-ray flux will last once it has started.

The advantage gained knowing that an event follows a lognormal function is that by taking the logarithm of the independent variable the function becomes a normal, or Gaussian, function, which is symmetric about the peak of the curve. Therefore, this symmetry can be used to determine the second half of the function simply by knowing the first half. This property will be used to create a forecast of the decay of solar flare x-ray flux base solely on the rise of the flux.

17

III. *Lognormal X-Ray Flux Methodology*

Knowing that the x-ray flux follows a lognormal shape is critical in being able to confidently create the forecast of the decay of the flux based on the first half of the flux. It will be shown, through looking at over 1300 flares, that the x-ray flux is approximately lognormal for the majority of flares. Once this has been established the forecast of the x-ray flux evolution can begin.

3.1   *Lognormal Plot Development*

3.1.1   *Setup*

Isolating an individual flare from 30 Aug 2004 and comparing it to a perfect lognormal shows how the x-ray flux approximates a lognormal. This is shown in Figure 11. It is also shown that when the logarithm of the independent variable is plotted, the x-ray flux becomes normal. This suggests the original data is lognormal for this flare. To convince ourselves that this is true in general, a more thorough investigation is undertaken.

Equation (3) represents a perfect lognormal. Comparing this perfect lognormal to a plot of the x-ray flux will show that the flux from a flare is approximately lognormal. As was shown earlier, a lognormal can take on different shapes as the mean and standard deviation change. To get a good comparison between the actual flux and the lognormal equation, the peak flux, mean, and standard deviation of the actual flux will be used in equation (3). Plugging these values into the equation will give a perfect lognormal to compare the actual flux data against. Therefore, if the known equation for a lognormal fits the flux it will show that the flux is indeed lognormal.

18
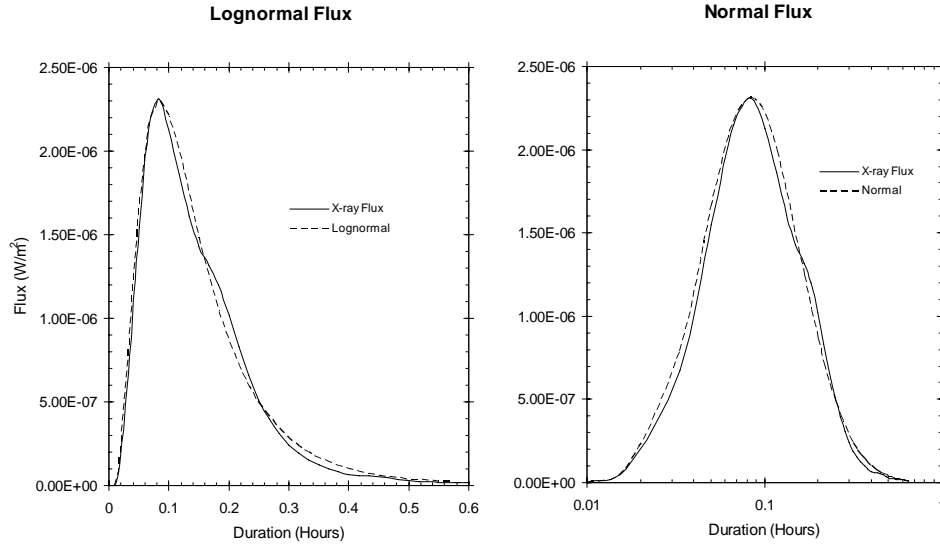
Figure 11. A C2 flare from 30 Aug 2004. The left plot shows the x-ray flux of the flare and a lognormal function. The same flare is plotted on a log scale on the right and fits well to the normal function. A perfect lognormal is superimposed for comparison.

After a flare has occurred the time, x, and the flux, f(x) are known. The peak magnitude, A, and the time of that peak, μ, are easily found without calculation. This leaves the standard deviation, σ, as the only unknown. Using equation (3) we can solve for σ by dividing by A and then taking the natural log of both sides of the equation. This gives a formula for the standard deviation.

$$\sigma = \sqrt{\frac{-(\ln(x) - \ln(\mu))^2}{\ln\left(\frac{f(x)}{A}\right)}} \qquad (4)$$

If the flare is a perfect lognormal, each data point should result in the same standard deviation. Since the flare isn't a perfect lognormal each data point gives an estimate of the standard deviation. The standard deviation for the entire flare is found by taking the average of each data point's standard deviation as seen in equation (5) where *n* is the number of data points used.

19

$$\sigma_{avg} = \frac{1}{n}\sum_{i=1}^{n}\left[\sqrt{\frac{-(\ln(x)-\ln(\mu))^2}{\ln\left(\frac{f(x)}{A}\right)}}\right] \tag{5}$$

For the 30 Aug 2004 flare, seen in Figure 11, σ_{avg} was calculated to be 0.86.

Although the flare lasted 39 minutes, only eight individual points were used to calculate

this average. The reasoning for not using all of the data points is given below. The eight

selected points produced standard deviations ranging from 0.81 to 0.98. A nonlinear fit,

using Mathematica®, gives a standard deviation of 0.88 compared to 0.86 using equation

(5). Since these two values are nearly equal, σ_{avg} is used as the standard deviation for the

entire flare.

The rise of the flux typically occurs so rapidly that, when using data with 1

minute resolution, very few data points are obtained in the rise phase compared with the

number of data points in the decay portion of the flare. For example, an M5 flare that

lasted 50 minutes had an 11 minute rise and 39 minute decay. This imbalance of data

points skews the resulting average standard deviation to the decay portion. In order to

overcome this, equally spaced points of the flare are used to calculate the average

standard deviation. Figure 12 shows a C6 flare from 27 May 2003 that has nearly three

times more points in the decay than in the rise. The decay is then thinned out so that

there are the same number of data points on both the rise and the decay. Substituting the

average standard deviation into the original lognormal equation results in the following.

$$f(x) = A\exp\left[\frac{-(\ln(x)-\ln(\mu))^2}{\sigma_{avg}^2}\right] \tag{6}$$
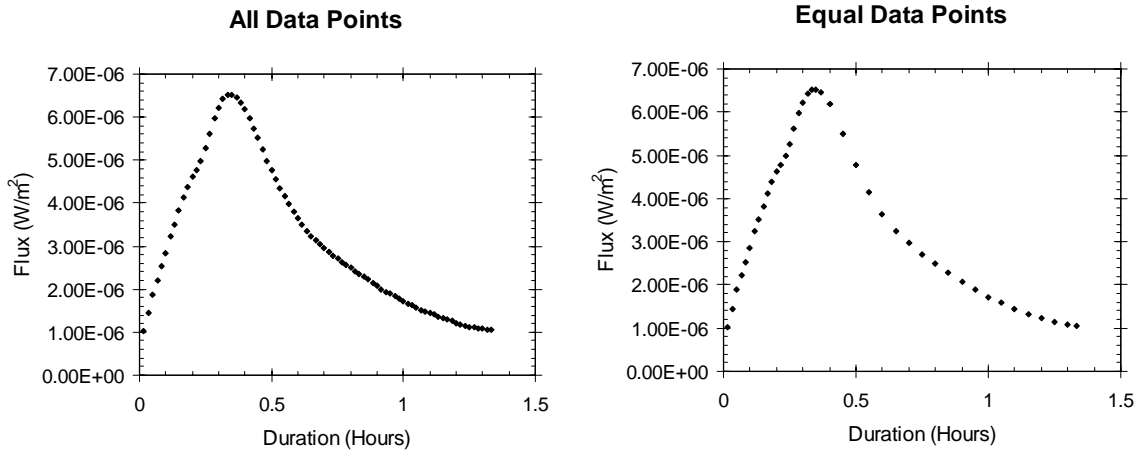
Figure 12. A C6 flare, from 27 May 2003, that had a total of 80 data points, 21 on the rise and 60 on the decay. The plot on the left shows all 80 points. The plot on the right shows the equally spaced data points used to calculate the RMSE and standard deviation.

3.1.2 Normalization.

The solar flares can vary widely in magnitude. In order to compare flares of differing magnitudes they need to be normalized to the same scale. To do this, a normalization process for the x-ray flux has been developed. This process begins with equation (6). First, divide the flux values by the peak flux to get equation (7). This normalized the flux for every flare to a value of one.

$$\frac{f(x)}{A} = \exp\left[\frac{-(\ln(x) - \ln(\mu))^2}{\sigma_{avg}^2}\right] \tag{7}$$

Next, take the natural logarithm of the left and right hand sides of equation (7) to eliminate the exponential and get equation (8). This shifts the peak flux for every flare to zero.

21

$$\ln\left(\frac{f(x)}{A}\right) = \frac{-(\ln(x) - \ln(\mu))^2}{\sigma_{avg}^2} \qquad (8)$$

Finally, multiply by the square of the standard deviation to obtain the final result seen in equation (9).

$$\sigma_{avg}^2 \ln\left(\frac{f(x)}{A}\right) = -(\ln(x) - \ln(\mu))^2 \qquad (9)$$

This has the form of a parabola, $y = -x^2$, where y is the normalized flux on the left hand side and x is $\ln(x) - \ln(\mu)$. To further simplify, if the mean is shifted to a value of one, then the logarithm of the mean will be zero. This will give every flare a normalized mean of zero and a normalized peak flux of zero. The normalization process has manipulated the original x-ray flux, which was lognormal, into flux that is parabolic. The peak of each flare will now be at zero and the rest of the flux will be on the same scale regardless of peak magnitude.

After normalization, a perfect lognormal becomes a perfect parabola. Applying this normalization to the flux data will make all the flares take on an approximate parabolic shape and they will all be on the same scale. Figure 13 shows this normalization for 16 flares ranging from C1 to X1 These flares come from 4 different days, 03 Jul 2002, 26-27 May 2003, and 26 Feb 2004, chosen because there were multiple flares for each day. A perfect parabola, the heavy dark line, has been added to the plot for comparison.
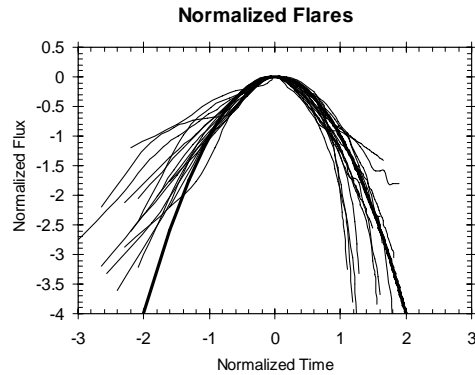
Normalized Flares

Figure 13. Sixteen normalized flares, from four different days, compared against a perfect parabola (heavy dark line). The approximate parabolic shape can bee seen.

The normalized flares follow the parabola fairly closely except for the beginning of the flare. This deviation is due to the choice of the flare onset time. Precise determination of the flare onset is difficult. The current algorithm typically retains data points that might be more appropriately removed. These points are kept in order to calculate the background flux that will be subtracted out. Figure 14 shows a much better fit to the parabola if the first two data points are removed. This is done only to show that the flares do approximate the parabolic shape. Any further calculations performed use all of the data points of the beginning of the flare.
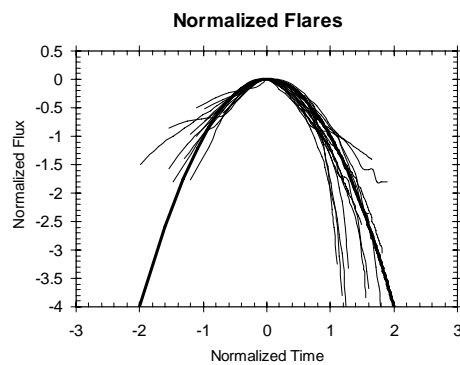
Normalized Flares

Figure 14. The same 16 normalized flares from Figure 13 with the first two data points removed. Removing the first two data points removes the beginning portion of the flare where the flux may not be rapidly increasing.

23

3.1.3   Lognormal Determination.

To determine if the flux is lognormal the flux values will be compared to a perfect lognormal. The Root Mean Square Error (RMSE) is used as a measure of the error between the flux and the perfect lognormal, the smaller the RMSE the closer the flare is to a perfect lognormal with an RMSE of zero being a perfect lognormal. The RMSE is given by equation (10).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(f-o)^2} \tag{10}$$

Where *n* is the number of comparison points, *f* is the value of the parabola for a given time, and *o* is the observed flux. The flux is normalized, through the process described in Section 3.1.2, so that every flare is on the same scale regardless of the order of magnitude of the peak flux. Since the RMSE is calculated after the normalization process, it has no units. The RMSE values that are calculated can now be compared between all classes of flares. In this normalized space the comparison is between the actual flux data and a perfect parabola.

Recall that when calculating the standard deviation the number of data points in the decay of the flux outnumbered the number of data points in the rise. This is also true for the RMSE. So, when calculating the RMSE, equally spaced points will be used as they were for the standard deviation; the same reasoning applies. Figure 12 shows an example of a flare before and after the data point filtering to get approximately the same number of points on the rise and the decay.

24

3.2 Results of Lognormal Development

The process of showing that the x-ray flux follows a lognormal function was performed on 1358 C, M, and X class flares from 1996, and 2002-2004. The RMSEs ranged from 0.014 to 4.51 with an average of 0.76 and standard deviation of 0.43. Figure 15 shows the lognormal curve plotted with the x-ray flux for a small and large RMSE. Why doesn't the flare on the right fit into the lognormal shape? The rise portion fits very well but the decay portion is where the problem is located. The x-ray flux drops off more rapidly than the lognormal function does. Shortly after the peak, the x-ray flux levels off and then continues the decay. This could have been caused by another flare beginning shortly after this flare had started. This leads to the flare not being approximated very well by a lognormal.

As a comparison for determining what values of the RMSE denote a good lognormal, a baseline RMSE has been developed. To get this value, a straight line is drawn horizontally across the parabola in a location that minimizes the RMSE.
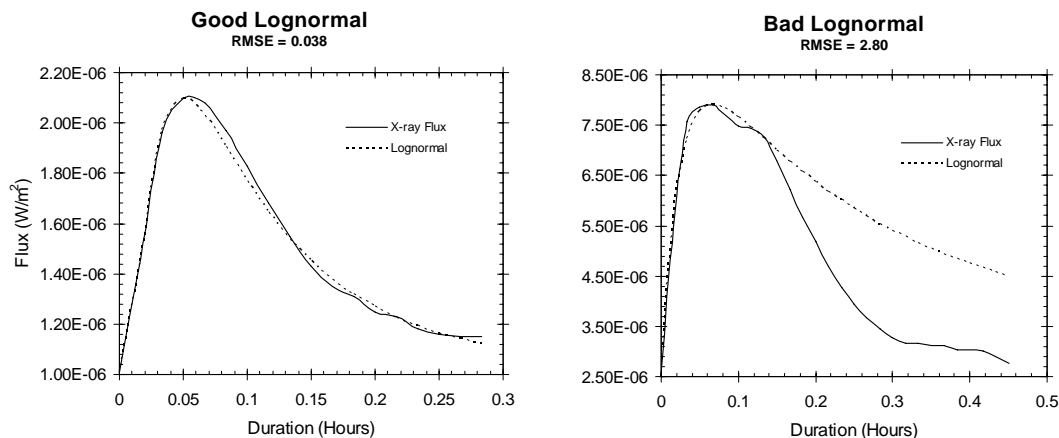
Figure 15.   The left panel shows a C2 flare from 14 Jul 2004 with an RMSE of .038. This follows the perfect lognormal shape very well. The right panel shows a C7 flare from 12 Apr 2002 with an RMSE of 2.80. This does not follow a lognormal shape.

The beginning and ending points were chosen to be plus and minus two. These values will encompass many of the flares as was seen in Figure 13. The vertical position that minimizes the RMSE is -1.19. The RMSE between the straight line and the parabola is taken to be the baseline RMSE. This baseline RMSE has a value of 1.34. This baseline is shown in Figure 16. Any RMSE larger than this is certainly a poor approximation of a lognormal.

Figure 17 shows a flare from 26 Jul 2002 that has the same RMSE as the average RMSE. The rise portion of the x-ray flux is represented well by the lognormal function. The decay portion is not represented as well but the general lognormal shape is still evident. The average RMSE is also well below the baseline RMSE of 1.34, providing more support for the idea that the x-ray flux evolution follows an approximate lognormal shape.
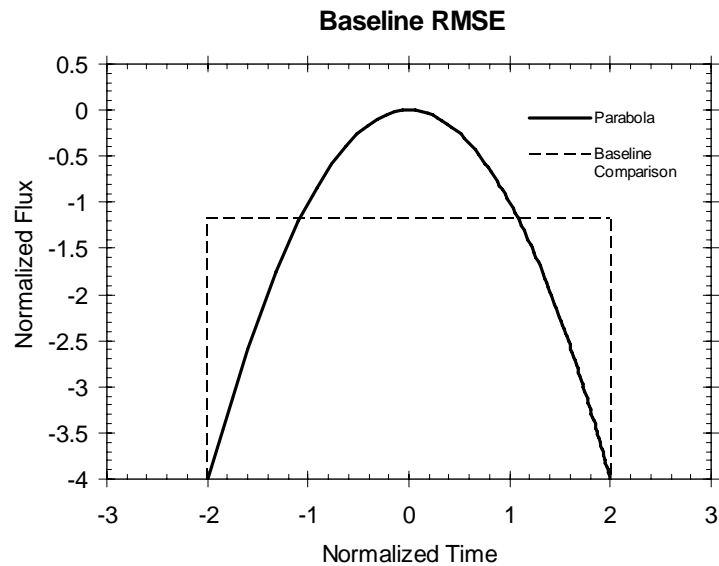
Figure 16. The baseline RMSE of 1.34 is calculated between the parabola and the baseline comparison line. Flares with an RMSE larger than this baseline RMSE do not follow a lognormal shape.

Comparing the RMSE of the flares to the baseline RMSE we see that only 130 out of 1358 flares had an RMSE greater than the baseline; this is less than 10%. This leads to the conclusion that the majority of flares follow an approximate lognormal shape.
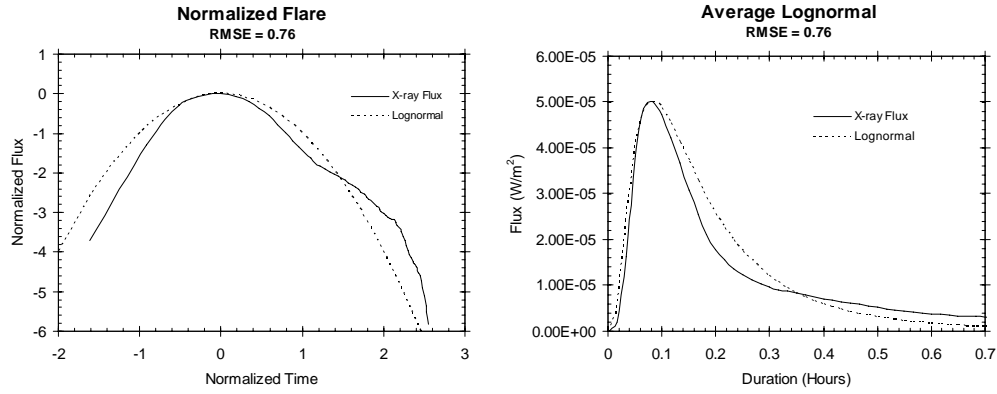
Figure 17. A flare from 26 Jul 2002 with an RMSE that is the same as the average RMSE of 0.76. The left panel shows the normalized flare. The right panel shows the same flare in real time.

IV. Forecasting Methods

A lognormal has the property that if you take the log of the time then you get a normal function, which is symmetric. That means the rise of the flux is the same as the decay of the flux. This can be used to forecast the decay of the flux once the rise is complete. *Campbell* [1996] was able to successfully do this for the Dst Index, a worldwide geomagnetic storm level index. The same technique can be applied to the x-ray flux from solar flares.

4.1 Standard Forecast Transformations

Before jumping into the forecasts it would be helpful to understand exactly what the flare detection algorithm considers a flare. To initially detect a flare, three successive flare values are analyzed. The second value must be at least 1.2 times larger than the first value and the third value must be at least 1.2 times larger than the second value. If this condition is met then a flare has begun, and these three values make up the first three points of the flare. The algorithm then records flux values as part of this flare until the end of the flare is reached. The end is determined when the flux value decreases to the flux value of the first point, or the flux is decreasing then begins to increase again. Two additional requirements are enforced on the flares in this study. The peak flux must be greater than 1.0×10^{-6} W/m^2 (C1) and there must be at least 20 minutes worth of data from beginning to end. If these qualifications are satisfied then the flux data officially becomes a flare to be analyzed.

A standard set of transformations is done for each flare before any flux forecast is created. It has already been stated that the natural logarithm of the time and the flux is

taken in order to create a symmetric curve. Each flare is also shifted so that they begin at time equal to zero. The sun constantly emits x-rays even when no flares are occurring, creating a background flux. This background flux is constantly changing with solar activity. When a flare occurs, the x-ray flux increases above the background level. In order to have all flares begin at the same point, this background is removed from the flare.

Once the flare finding algorithm detects a flare and begins saving data, the first saved point is taken to be the background x-ray flux. This flux is then subtracted from each data point. Thus, the start time and flux are both zero for all flares. Recall that the end time is set at the time of the half-max flux occurrence, which is now just one half of the peak flux.

4.2 *Dst Index Similarities*

The Dst Index is a global ring current index that monitors the magnetic storm level. As a magnetic storm begins, the horizontal component of the geomagnetic field near the equator becomes negative. This negative turning is due to the ring current that flows around the earth in the equatorial plane [*Tascione*, 1994].

As the ring current is charged with energetic particles from the solar wind the horizontal component of the geomagnetic field becomes more negative. This field slowly returns to normal as the energetic particles are lost from the ring current through a number of different mechanisms that occur simultaneously thereby creating a lognormal shape. The Dst Index follows this pattern of becoming more negative fairly quickly and

then slowly decaying back to normal; Figure 18 shows this. If the image is inverted the
shape is that of a lognormal with a quick rise and a slow decay.
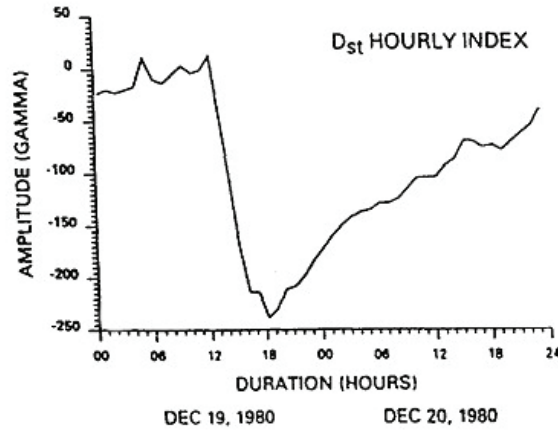
**Figure 18. The Dst index during a geomagnetic storm. A lognormal shape can be seen if the image is
inverted [*Campbell*, 1996].**

Using the inverted image and working with the logarithm of the time so that the

distribution is normal, *Campbell* [1996] took the rise of the Dst index and used the mirror

image to estimate the decay. He then fit a nine-term polynomial function to the data

points and converted back to real time. The goodness-of-fit measurement he used was

the percent difference given in equation (11).

$$100 \frac{(observed - predicted)}{observed} \qquad (11)$$

This goodness-of-fit was done for hourly values from the peak to the end. For the 10

geomagnetic storms he analyzed he was able to predict the recovery phase values to

25.7%. This is a good prediction of the recovery phase of Dst based on the rise data only.

The rise portion of the Dst index is several hours long while the decay portion can

last for days. Flares occur on a much quicker time scale, but the same logic applies

because both the Dst index and the x-ray flux follow a lognormal shape.

30

4.3 Symmetry Forecast

Now that it has been shown that a flare is approximately lognormal and that the flipping of data points has been used successfully before, the same technique can be applied to solar flare x-ray flux to forecast the decay of the flux based on the rise. By taking the natural logarithm of the time coordinate, we transform the flare into a normal (Gaussian) function. We then use the symmetry to forecast the decay portion of the flux.

As a flare progresses, the x-ray flux will increase with each minute of data. At one minute (one data point) past the peak, the flux will begin to decrease, indicating that the peak flux has occurred. At this point, using the symmetry around the peak data point, we take each data point from the rise portion of the flare and flip it over to the other side of the peak point as shown in the left panel of Figure 19. This creates a symmetric plot around the peak based solely on the rise of the flux. This constitutes a set of data points that is symmetric about the peak flux. The symmetric data is converted back to real time and used to calculate the average standard deviation using equation (5). Once this is found, the observed peak (A), the mean (μ), and the calculated average standard deviation (σ_{avg}) are plugged into equation (6). This produces a smooth lognormal curve that fits the flipped data points as demonstrated in the right panel of Figure 19.

This curve fit is the prediction of the decay of the x-ray flux based only on the rise of the flux plus the first point of the decay. Figure 20 shows how well the flipped data points can represent the decay of the flux. The curve fit to the flipped data points gives a good estimate of this flare from 3 Nov 2004. The fit is off slightly toward the end, but the majority of the flare is forecast very well.
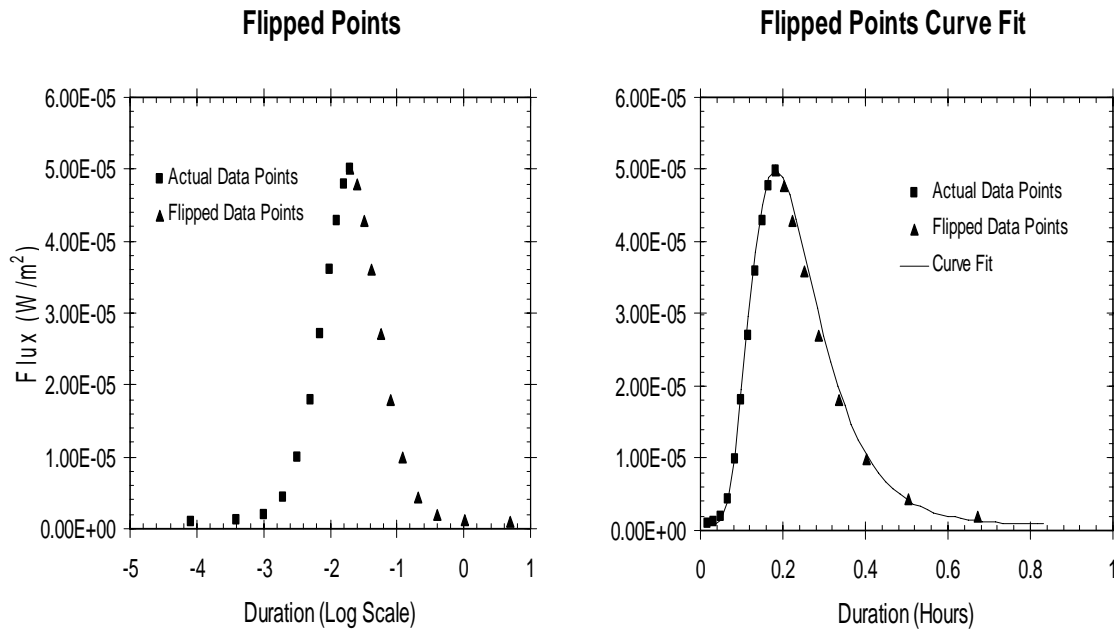
Figure 19. An M5 flare from 03 Nov 2004. The square data points are actual measurements and the triangles are those points flipped over. The panel on the left shows the data points in log time. The panel on the right shows the points converted back to real time and a curve has been fit to the points. This curve is the forecast for the evolution of the x-ray flux.

This plot gives an accurate representation of how the flare decays if there are no other influences, such as other flares erupting, during that time. A flare has a smooth decay when there are no other influences during the decay of the flux. Since the x-ray flux is the integrated disk flux, if another flare occurs anywhere on the visible disk of the sun during the decay of the flare of interest, it can increase the total flux, thereby altering the shape of the decay of the current flare significantly. A situation could also arise in which a flare occurs just as the active region rotates off of the visible solar disk. This could lead to some of the emitted x-rays being absorbed by the sun so that the x-rays reaching earth are not a complete representation of the flare.

32

Flipped Data Forecast
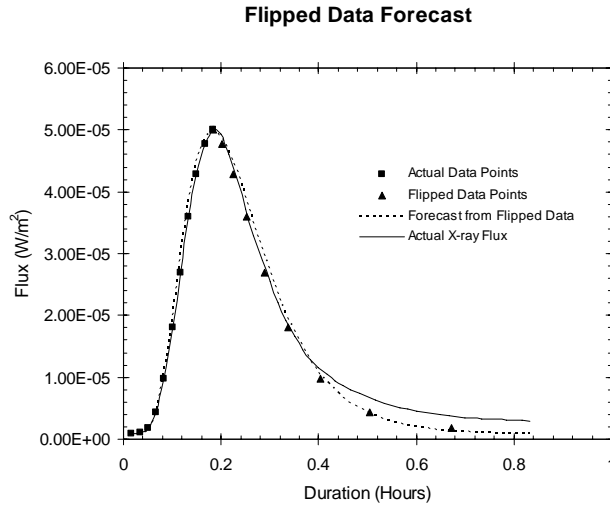
Figure 20. X-ray flux forecast generated from the flipped data points seen in Figure 19. This forecast is based on the rise portion of the flare plus one point and fits the actual x-ray flux very well.

Figure 21 shows a flare from 12 Apr 2002 that is not lognormal. Another reason for the poor forecast for this flare is that there were only four points available on the rise phase. Having only a few points on the rise combined with the non-uniform decay of the flux caused a poor forecast to be generated.

Poor Forecast
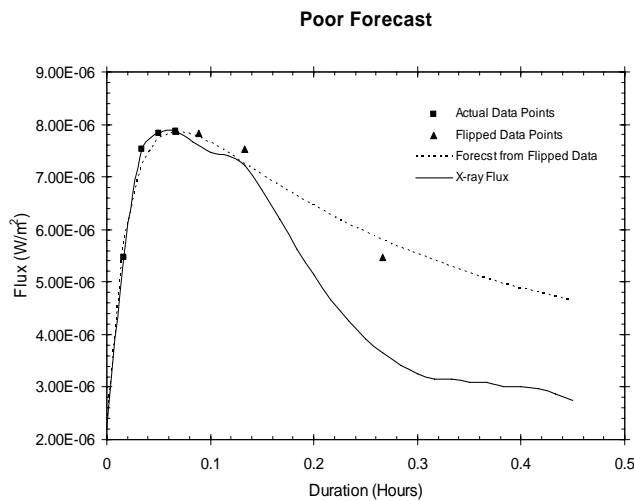
Figure 21. A C7.9 flare from 12 Apr 2002 that has only four points on the rise and has a non-uniform flux decay.

33

This method is quick, simple, and reasonably accurate for the majority of flares. The disadvantage is we must wait until the flare has reached its peak before making a forecast. Flares can be as short as 20 minutes, so waiting until the peak is reached can reduce any warning time. Being able to forecast the peak and duration of the flux as soon as the flare begins would be far more useful due to the short duration of flares.

4.4   Pre-Peak Forecasting Methods

Creating an accurate x-ray flux forecast with only a few data points on the rise portion presents a significant challenge. As described in Section 3.1.2, when the natural log of the time and the flux is taken, the flux takes on a polynomial shape, specifically a second order polynomial. Using this fact, two different forecasting techniques using polynomials to predict the flare shape were devised, the shifting method and the non-shifting method. A least squares fit is used to create a polynomial to fit the data points available. Both the shifting and non-shifting methods are used to forecast the x-ray flux evolution. During the creation and analyses of the forecasts, the end of the flares and the forecasts will be the half-max point. Since not all flares return back to the background flux values, the half-max is used as a consistent end point.

4.4.1   Polynomial Determination.

By taking the natural logarithm of both the time and the flux, the flare is transformed from a lognormal shape into the shape of a quadratic. It would make sense to use a quadratic equation for the least squares fit. Using the forecasting methods that will be discussed later, it was found that using a quadratic equation to fit the shape of the flux is not ideal. The quadratic fit does not give a peak flux forecast that is large enough

34

and the decay rate of the flux is forecast incorrectly. Figure 22 shows two flares with

forecasts for both quadratic and fourth order polynomials. These forecasts are created

every minute, beginning when five minutes of data is available and they stop when the

peak is reached. A flare that has nine data points on the rise will have five forecasts

generated; the first one at the five minute point and then one for each minute thereafter

until the peak is reached. Once the peak is reached, the symmetry forecast method is

used. The more data points available, the better the forecast will be because the

polynomial has more points to define the shape of the curve. The figure shows that the

quadratic forecasts don't decay quickly enough. Their decay time exceeds that of the

flares. The fourth order polynomial has a shape that more closely resembles that of the

flare. Details of the polynomial forecasting method will be discussed later. It is

sufficient at this point to recognize, based on Figure 22, that the fourth order polynomial

creates the better forecasts of the two polynomials.
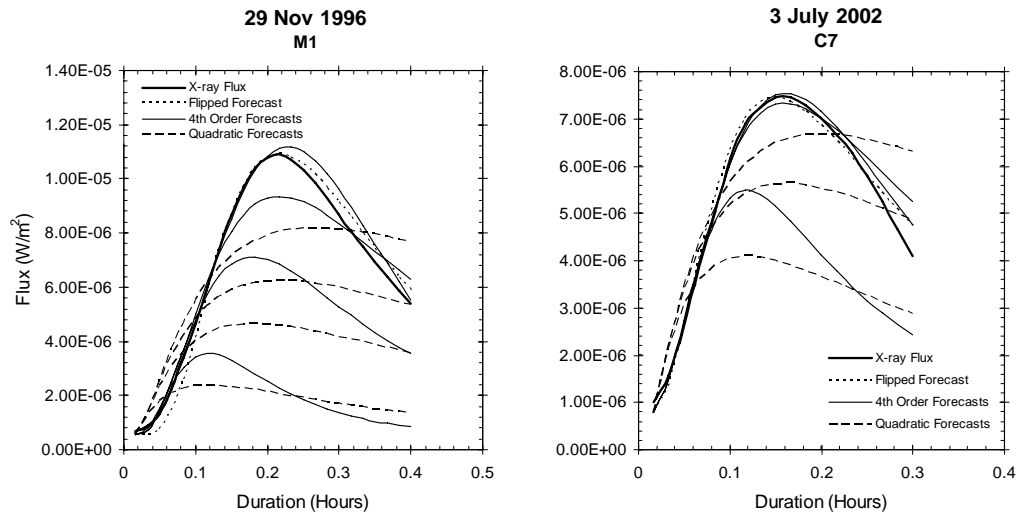
Figure 22. These plots show both the quadratic forecasts (dashed lines) and the fourth order forecasts (thin solid lines). Multiple forecasts are shown for the same flare because a forecast is generated for every minute of the rise from five minutes to the peak.

Similar results are found for the other flares as well. The forecasting technique using the fourth order polynomial will be discussed later.

A fourth order polynomial has more freedom than a quadratic to fit the end points if they don't follow the lognormal shape (in real time). Based on a visual inspection of the particular flare shown in Figure 22, the results for the fourth order polynomial were much better than for the quadratic. A fourth order polynomial removes any physical connections to the lognormal that were present with the quadratic, but a better forecast is obtained.

If the flux followed a perfect lognormal then the least squares fit with a quadratic would be ideal, but it isn't a perfect lognormal. Because we are using very few points to fit a curve, the shape of the beginning of the flux has a huge impact on the overall fit. As was shown in Figure 13, the beginning points of the x-ray flux did not fit near the parabola leading to the conclusion that those points skew the x-ray flux away from the lognormal shape. But in Figure 14 the first two points have been removed and the x-ray flux fits more into the lognormal shape. So, trying to fit a quadratic to the first few points of the x-ray flux will be unsuccessful because those points are skewed from the perfect lognormal shape. As more data becomes available the forecast will take on a more lognormal shape but will still be skewed away from the lognormal because of the beginning points.

4.4.2   Shifting Forecast Method.

As a flare occurs, flux data is available in 1-minute intervals. As this data becomes available during a flare, a fourth order polynomial can be fit to these points. It is unlikely that a polynomial fit to just the first few points on the rise of a flare will accurately

36

determine the remainder of the flare. However, if the polynomial was fit to a few points on the rise of the flare and to points on the decay as well, there would be a point where the polynomial would have to curve back down in order to fit the estimated decay points. The so-called "shifting method" builds the decay portion of the x-ray flux using the available data from the rise portion of the flare.

Once a flare begins and there is five minutes worth of data, the shifting method forecast begins. Enough data is needed as the rise phase begins to determine if the rapid rise constitutes the beginning of a flare. It is possible to have a quick jump in the flux that lasts for several minutes that is normal background fluctuations and not a flare. So, five minutes was arbitrarily chosen to represent the dividing line between a normal rise in background flux and an actual flare. Using the fact that in log time space the flux will be approximately symmetric, these beginning points on the rise are flipped about the current data point, analogous to the method used when the entire rise was known. They are then shifted over 0.69 units in log time to simulate the end of the flare. The distance that the points are shifted is independent of the current flux. Once the points are shifted, there are data points for both the beginning of the rise and estimated data points for the end of the decay. Each time a new data point becomes available the distance that the data points are shifted is reduced by 0.06 units and the process of flipping and shifting of data points is repeated. This reduction attempts to keep the data from being shifted to far as the peak approaches. Figure 23 shows this process step-by-step for an X10 flare on 29 Oct 2003.

37

Figure 23. The shifting forecast process for the 29 Oct 2003 X10 flare. Panel (a) displays the first five data points after the natural log of the time has been taken. These points are then flipped over and shifted in panel (b). Panel (c) shows the fourth order polynomial (dotted line) that is fit to the points. This is the forecast for the flux evolution based on the first five points. The true x-ray flux is overlaid for comparison. Panels (a) through (c) are repeated to create the forecast using 10 points on the rise seen in panel (d).

There are obvious disadvantages to this method. The proper length of time to shift the data points is unknown because of the seemingly random nature of the duration of solar flares. Any deviation of flare duration from the estimated time shift will cause the forecast to be incorrect. Another problem is that this method assumes that flares are perfectly symmetric in the log time space. As was shown earlier, flares are not quite symmetric. Any other activity on the sun, besides the flare of interest, that produces x-rays will influence the shape of the flux curve. As was shown before, a separate flare

38

occurrence during the flare of interest will increase the total x-ray flux observed. This can alter the shape of the x-ray flux curve away from a lognormal shape.

4.4.3   Non-shifting Forecast Method.

Another method, referred to as the "non-shifting" forecast, has also been developed. This method employs the least squares curve fitting like the shifting method did, but with one important difference. The fourth order polynomial is fit to the flux points on the rise of the flux only. The points are not flipped and shifted over like they were in the shifting method.

Since the fit to the points is a fourth order polynomial, it is an even polynomial and will be symmetric about a vertical axis located somewhere along the x-axis. Thus it will have a peak, or peaks, even though no data points have been shifted over. This method requires that more data points are available than for the shifting method in order to give the polynomial better direction. With only five points, the fourth order polynomial has a great deal of freedom in where the peak of the curve will be. As more points are acquired, the polynomial becomes more constrained.

Therefore, once the flux has eight data points, and for each data point thereafter, a least squares fit is obtained for the known data points of the rise only. Eight data points was chosen as the least amount of points that led to reasonable results. Figure 24 demonstrates the non-shifting method for an M2 flare from 26 May 2003. This example begins with 10 data points. Panel (a) shows the first 10 data points of the flux that have been converted to log time by taking the natural logarithm of the independent variable. In panel (b) a fourth order polynomial has been fit to the points using the method of least squares. In panel (c) the data points and the curve fit have been converted back to real

39

time and the x-ray flux has been plotted to demonstrate how well the fit models the flux. This process is repeated for each minute of data. Panel (d) shows the forecast for the 13th data point.

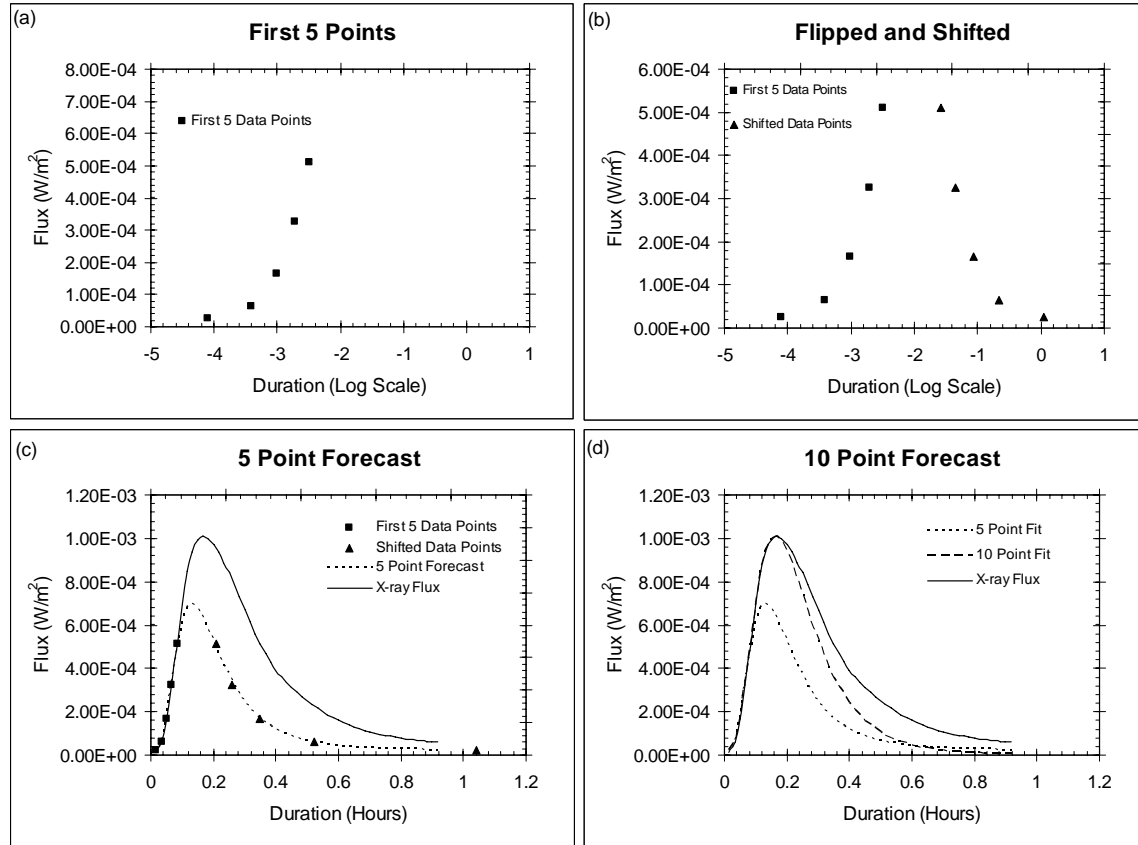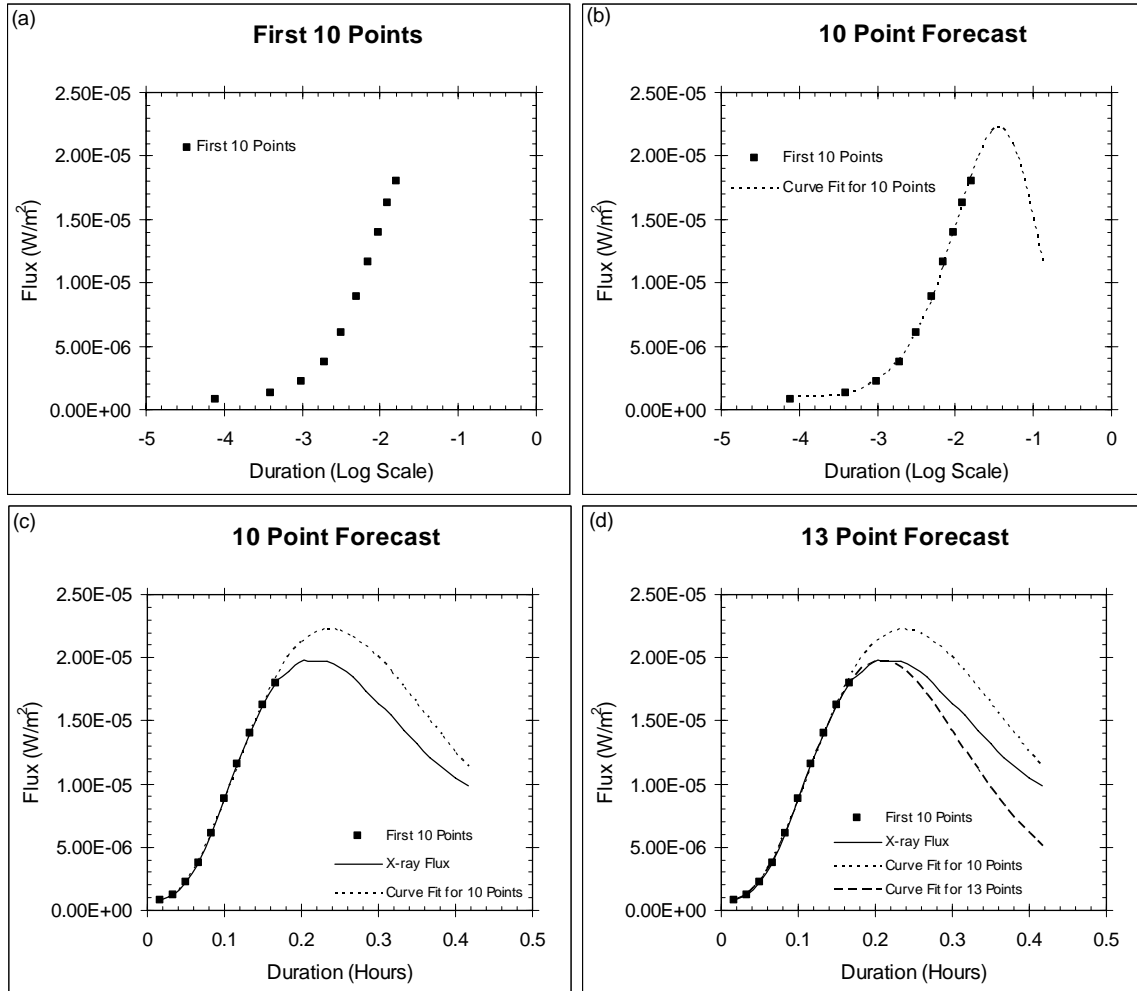**Figure 24. The non-shifting forecast process for the 26 May 2003 M2 flare. Panel (a) is the first 10 data points after the natural log of the time has been taken. A fourth order polynomial is fit to these points in panel (b). Converting back to real time, panel (c) shows the fourth order polynomial (dotted line) that is fit to the points. This is the forecast for the flux evolution based on the first 10 points. The x-ray flux is overlaid for comparison. Panels (a) through (c) are repeated to create the forecast using 13 points on the rise seen in panel (d). The peak flux for this flare is found at the 13 minute mark, so the 13 point forecast is the last forecast generated with this method.**

Each polynomial that is fit to the data points is checked to determine if it creates a flare-like shape. For example, the polynomial could fit the data points along the rise but

then just continue to rise. A check is made to ensure the fit follows the rise points but then turns back down and makes a peak like the x-ray flux does. Figure 25 shows a polynomial fit to the first 10 points of the 26 May 2003 M2 flare. The horizontal dashed line represents the value of the flux for the first data point. The fit is found to rise and then decay back down to the flux value of the beginning data point. A value called the time separation is then found from this data. It is the value of the time coordinate from point a (the beginning flux value) to point b (the point that it returned to the original flux value). This time separation value will be used to determine which forecast method, shifting or non-shifting, will be used.
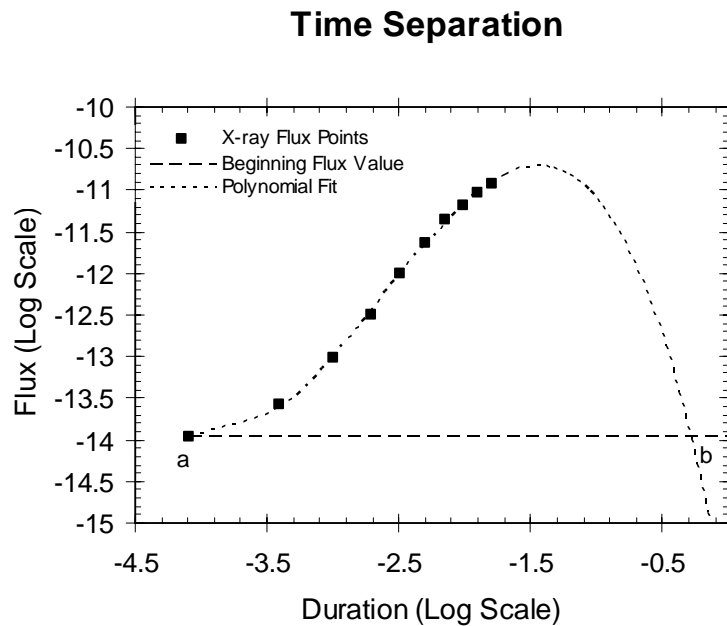
Time Separation

Figure 25. A fourth order polynomial fit (dotted line) to the first 10 points of the 26 May 2003 M2 flare. The fit shows a rise and then a decay back to the beginning flux value (dashed line). The time separation value is then taken as the distance between points a and b.

If the polynomial does not show a rise and then a decay within 60 minutes, as it did in Figure 25, it is dismissed and the forecast reverts back to the shifting method forecast.

Sixty minutes is chosen because the duration of the flare that is being forecast is not known, and 60 minutes is ample time for the majority of flares to reach the half-max end time. For flares that last longer than 60 minutes and the non-shifting forecast does not decay within 60 minutes, the shifting method will be used. This scenario and the time separation will be described in more detail in Section 4.4.4.

The main disadvantage to this method is clear. Since there are no real or estimated data points on the decay, more data points are needed on the rise to give the shape of the curve. This forces a longer wait once the flare has already begun to start producing forecasts. The arbitrary value of eight data points was the minimum number of data points chosen to start trying the non-shifting method.

4.4.4 Combined Method.

Two methods for forecasting the x-ray flux evolution have been discussed. Each of these methods have obvious disadvantages. In order to minimize these disadvantages, both methods are utilized. The method that has the shorter time separation is the method that is used for the forecast.

After a flare begins and there are five data points available the forecasts are created each minute. Recall that the non-shifting forecast does not start until there are at least eight data points. Therefore, for the first seven minutes of a flare only the shifting forecast is used. Once the eighth point is acquired both methods, shifting and non-shifting, are available for use. If the non-shifting method produces a reasonable forecast (i.e. has a rise, peak, and decay), and the calculated time separation is less than the time shift from the previous minute's forecast then the non-shifting method is used. If a reasonable forecast is produced but the time separation is larger than the previous time

42

www.manaraa.com

shift, the shifting method will be used. If a poor forecast, rise but no decay, is found with the non-shifting method then the shifting method is used. So for each forecast created the method that has the smallest time separation is used.

This procedure allows for the use of the shifting method during the early minutes of the flare, when the non-shifting method generally does not produce good results. It also eliminates the disadvantage of the shifting method, where it is not known exactly how far to shift the rise points over. The selection of the forecast based on the shortest time shift continues until the peak of the flux is reached. Once the peak is known, the forecast reverts to the symmetry forecast, as described in Section 4.3 above.

Once the forecast is created, it is necessary to convert back to the real time and flux to produce a graphical forecast. The background flux, as well as the beginning time, are also added back. The classification of flares is done without the background flux removed, so adding the flux back in allows for an accurate forecast of the peak flux value.

4.5   Forecast Error Method

After the flare ends, the RMSE is used to determine how well the forecast performed. The end of the flare is taken to be when the flux returns to the half-max value. In order to compare the forecasts for one flare against the forecasts of another flare, the data for the flare and forecast, are transformed to a normalized space to calculate the RMSE, since flare magnitude can vary over several orders of magnitude.

After the forecast is made, the flare data and forecast data are both divided by the peak flux value of the flare. This normalizes the peak flux of the flare to a value of one

43

and transforms the forecast data in like fashion. The RMSE can now be calculated to determine how well the forecast did.

The RMSE is calculated from flare start to the half-max point. If the flare does not make it to the half-max point before another flare begins, then the RMSE is calculated up to the last point of the flare before the other flare begins.

Once the peak is reached and the symmetry forecast is used, an additional statistic is generated. The symmetry forecast nearly always is the best forecast, so the timing of the half-max forecast is used as a metric of forecast performance. The half-max time difference will be negative when the forecast value occurs before the flare value; and positive when the forecast value occurs after the flare value.

RMSE values can easily be compared between different flares, but what value of RMSE determines a good, or usable, forecast? The average flux of the flare from start to end is taken as a generic forecast. If this average flux were overlaid on a plot of the flux, it would be a horizontal line through roughly the middle of the flare. The RMSE of this generic forecast, compared to the actual flux data, is used as a baseline for the shifting and non-shifting forecasts. An RMSE larger than this RMSE is assumed a bad, or unusable, forecast while an RMSE smaller than this RMSE is a good, or usable, forecast.

V. Forecast Results

5.1 Symmetry Forecast Results

This forecasting method was performed on 1358 flares from the years 1996, 2002, 2003, and 2004. The accuracy of the symmetry forecast method is determined by how well the forecast predicts the time it takes for the x-ray flux to decay from the peak to the half-max value. The accuracy of this forecast is compared to climatology's ability to accurately predict the same peak to half-max value time.

Climatology of the 1358 flares examined was created for this purpose. The flares were separated into groups according to their x-ray classification. There are nine groups for C class flares (C1, C2, C3,..., C9), and nine groups for M class flares (M1, M2, M3,…, M9). X class flares don't occur as often so there were some classifications that were not observed during the given years. The breakdown for the X class flares is X1, X2, X3, X5, X8, X10, and X17. For each category of flare, the average time for the observed x-ray flux to decay from the peak to the half-max was calculated. For example, a C7 flare will take 8.4 $\pm$ 5.7 minutes to decay from the peak to the half-max value while an M2 flare will take 9.5 $\pm$ 7.5 minutes. A complete list of decay times can be found in the Appendix: Climatology.

To determine how well this climatology actually predicts the peak to half-max decay time, the RMSE between the observed decay time and the climatological decay time is calculated for each class of flare. Recall the formula for the RMSE is:

$$RMSE = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^{n} (c - o)^2} \qquad (12)$$

45

where *o* is the observed decay time, *c* is the climatological decay time, and *n* is the number of flares in the class of flares. The RMSE is also computed for the decay time predicted by the symmetry forecast. The same formula is used but instead of *f* being the climatological decay time it is now the forecast decay time. For each class of flare there will be two RMSE values, one for the climatology and one for the forecast. The method with the smaller value of the RMSE is the method that is the most accurate in predicting the decay time of the x-ray flux.

A breakdown of each flare class is given in three separate tables, Table 3 for C class flares, Table 4 for M class flares, and Table 5 for X class flares. These tables show the total number of flares in each class that were used in both the climatology and the forecast calculations. The RMSE is shown for both the climatology and the symmetry forecast. Some of the flare classes (indicated by an asterisk) have only one or two flares in them for the calculations. This data may not be representative of the particular class but is shown for completeness. For all classes of both C and M flares, the forecast RMSE is smaller than the climatology RMSE. This indicates that the symmetry forecast does a better job than climatology predicting the decay time from the peak to the half-max value for all C and M class flares. The symmetry forecast method does better than climatology for X1 and X2 flares but is slightly worse for the X3 flares. For the higher X class flares there are not enough observed flares to produce results that are statistically significant.

Table 3. RMSE of climatology and the symmetry forecast for C class flares.

X-ray Classification	Number of Flares	Climatology RMSE	Symmetry Forecast RMSE
C1	235	3.85	2.63
C2	235	5.11	2.89
C3	145	4.72	4.43
C4	122	4.37	2.98
C5	73	4.44	2.88
C6	54	6.20	3.10
C7	50	5.73	2.42
C8	43	4.03	3.42
C9	34	5.55	3.37

Table 4. M class flares have the same results as were seen for the C class flares in Table 3. The classes with an asterisk have too few flares to be statistically significant.

X-ray Classification	Number of Flares	Climatology RMSE	Symmetry Forecast RMSE
M1	159	6.18	3.18
M2	78	7.53	3.48
M3	26	4.51	2.97
M4	18	4.77	2.91
M5	18	7.14	3.30
M6	9	5.39	1.91
M7*	3	1.89	0.58
M8*	3	2.62	1.83
M9	9	5.19	2.69

Table 5. The X1 class, like the C and M classes show that the symmetry forecast is the better predictor of the decay time. The other X classes have too few flares to be statistically significant.

X-ray Classification	Number of Flares	Climatology RMSE	Symmetry Forecast RMSE
X1	26	13.43	4.67
X2	6	3.80	2.68
X3	6	7.83	8.02
X5*	2	1.00	0.71
X8*	1	0.00	0.00
X10*	1	0.00	0.00
X17*	2	2.00	2.83

It has been determined that the symmetry forecast does a better job at predicting the decay time from the peak to the half-max value than does climatology. Another

47

measure of the symmetry forecast's efficiency is how well the forecast predicts just the time of the half-max value of the flare. For example, the half-max value of a flare occurs 35 minutes after the flare onset and the forecast says the half-max will occur 37 minutes after flare onset. This forecast is two minutes off in the prediction of the time of the half-max value. The error will be positive if the prediction over forecasts the half-max time of occurrence. The error will be negative if the prediction under forecasts the half-max time of occurrence. If the forecast is for the half-max value to occur at 20 minutes after flare onset and it actually occurs at 25 minutes, then the error will be minus five minutes. An example of both a small time error and a large time error are given in Figure 26. The flare on the left is an M9 flare from 9 Nov 2004 and had a perfect forecast of the half-max time. On the right is an X1 flare from 18 Aug 2004. This flare appears to have saturated the x-ray sensor with a peak flux of 1.8×10^{-4} W/m^2, but saturation doesn't occur until 1.74×10^{-3} W/m^2. This is a bad forecast with a time error of positive 13 minutes.
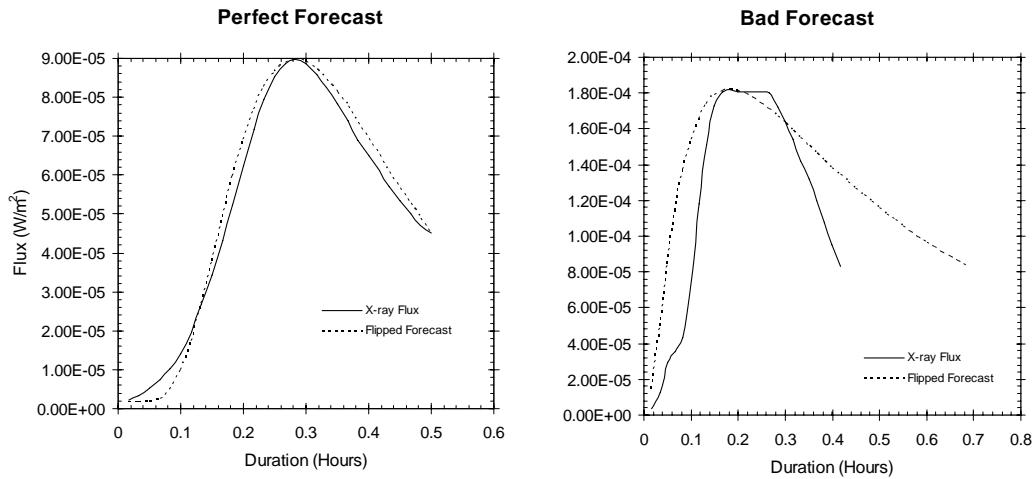
Figure 26. The M9 flare from 9 Nov 2004 gives a perfect forecast with a time error of zero minutes and is shown on the left. The X1.8 flare from 18 Aug 2004 has a time error of 13 minutes and is on the right.

This error calculation was performed on the same 1358 flares. The flares are grouped into three categories for this error calculation, C, M, and X. Table 6 shows how many flares of each classification were considered and gives the average timing error for each classification, along with the standard deviation.

Table 6. Symmetry forecast results. The average timing error for the half-max prediction was less than one minute with a standard deviation of three minutes. M class flares were predicted the best with an error of half a minute and standard deviation of three minutes.

X-ray Classification	Number of flares examined	Average time from peak to half-max (minutes)	Standard Deviation of average time (minutes)	Average half-max timing error (minutes)	Standard Deviation of timing error (minutes)
C	991	7.4	4.8	0.9	3.1
M	323	8.9	6.3	0.5	3.0
X	44	11.6	11.1	0.9	4.5
All (C,M,X)	1358	7.9	5.6	0.8	3.1

For the C class flares the forecast over-estimated the half-max time by nearly one minute with a standard deviation of three minutes. Looking at one standard deviation this gives a range of six minutes for the timing error. This range of errors is large when compared with the average peak to half-max time for C flares of only seven minutes. This gives a situation where the range of the forecast covers nearly the entire decay of the flux. The same situation is repeated for the M and X class flares as well, but not quite as extreme. This could lead to the conclusion that the symmetry forecast method is of no value, but the standard deviation of the peak to half-max time provides a large range for the half-max value to occur. With a wide range for both the half-max time and the error, the potential for both good and bad forecasts exist. The good forecasts outnumber the bad forecasts with 73.6% of the forecast half-max times falling within two minutes of the actual half-max time, and 31.6% of those getting the half-max time exactly right (with 1-

49

minute resolution data). In contrast, climatology was only able to produce a forecast half-max time within two minutes of the actual half-max time 43% of the time.

The current method of predicting the decay time from the peak to the background flux is based on climatology, as was seen in Table 1. The symmetry forecast method improves upon the climatology method for C and M class flares and the smaller X class flares. More data is needed on the higher X class flares to make a definitive statement on whether or not the symmetry forecast method is better than the current climatology method for these flares.

5.2 Combined Forecast Results

The combined method creates a forecast every minute, beginning five minutes into the flare, until the peak flux is reached. Although this produces more forecasts per flare than the symmetry method, it does not produce the quality of results that the symmetry forecast method does. Using only a few points at the beginning of a curve to create the rest of the curve is a difficult problem. This method was run on the same 1358 flares as the symmetry forecast.

Of the 1358 flares detected, 1010 successfully had forecasts created for them. The other 348 either had errors while trying to convert from a fourth order polynomial to real time. The errors occurred because the coefficients of the fitted polynomial from the least squares fitting routine were large enough to produce overflow errors when the exponential was taken while converting to real time. If the actual observed flux did not decay back down to the half-max value a forecast can still be generated, but to calculate the RMSE and the time difference at the half-max value for the symmetry forecast my

50

algorithm needs the flare to return back to the half-max flux value. So, any flare not

returning to the half-max value is left out of the forecasting routine. Of the 1010 flares

that had forecasts created, the non-shifting forecast method, as well as the shifting

forecast method, was used on 262 of them. The other 748 used only the shifting forecast

method.

Results of the forecasting are shown for the entire day of 26 Feb 2004 in Figure 27.

The dashed lines are the best forecast for that particular flare, whether if be from the

shifted or non-shifted method. The inset is a closer view of two C class flares. Both

were forecast very well.

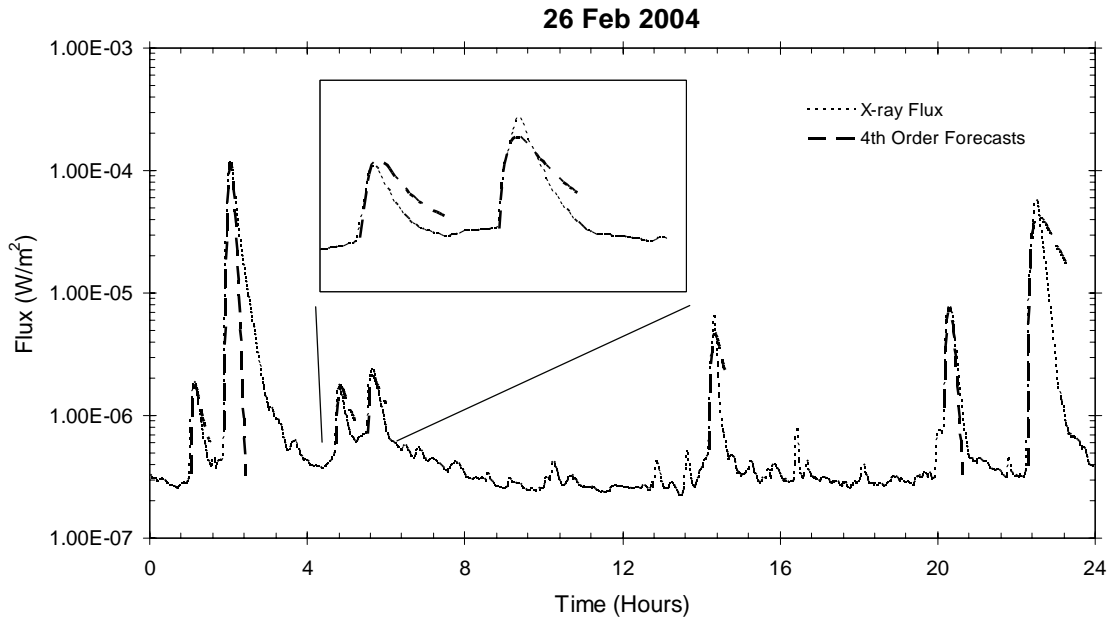**Figure 27. X-ray flux for 26 Feb 2004 with forecasts overlaid. The symmetry forecast and only the best fourth order polynomial forecast are shown. Inset is a closer view of two C class flares and their corresponding forecasts.**

Figure 27 provides a look at an entire day. Forecasting for a large number of flares

can be automated; however, to provide a more detailed analysis of the results, it is

51

necessary to focus on a few selected flares. Nine flares, three from each classification (C, M, and X), were chosen to demonstrate the combined forecasting method. Table 7 gives the date and x-ray classification of the nine selected flares. Some of the flares have the same classification so an a or b has been added to the classification to distinguish between them.

Table 7. The dates and x-ray classifications of the nine selected flares. Flares with the same classification have had an a or b added to the classification to distinguish between them.

Flare Date	Classification	Lognormal RMSE
26 Feb 2004	C7a	0.63
03 Jul 2002	C7b	0.63
20 Jul 2004	C8	0.87
29 Nov 1996	M1	0.78
26 May 2003	M2a	0.70
19 Sep 2004	M2b	0.99
10 Nov 2004	X2	0.74
23 Jul 2002	X5	1.35
19 Oct 2003	X1	0.95

The flares were chosen to represent predominantly flares that approximate a lognormal shape. Recall, from Section 3.2, that the average RMSE, when determining if the flare was lognormal, was 0.76. The chosen flare's lognormal RMSE ranges from 0.63 to 1.35, with an average of 0.85. This gives a sample of flares that were both more lognormal than average and some that may not fit the lognormal shape very well.

An RMSE for each forecast is created to determine how well the forecast fits the x-ray flux. In order to determine what value of RMSE constitutes a good forecast, a baseline forecast is created. Once the flare is over, the average x-ray flux for the flare from start to end (half-max) is computed and becomes the baseline forecast. This forecast resembles a step function. Figure 28 shows an M1 flare from 29 Nov 1996 with

52

the average x-ray flux computed and plotted. Once this average is computed, an RMSE

is calculated for the baseline forecast and is referred to as the baseline RMSE. If a

forecast RMSE is less than the baseline RMSE it is assumed a good, or usable, forecast,

but if the forecast RMSE is greater than the baseline RMSE it is a bad, or unusable,

forecast. Figure 28 also shows both a good forecast and a bad forecast for the same flare.

The baseline RMSE for this flare is 0.295. The good forecast has an RMSE of 0.093.

The bad forecast has an RMSE of 0.448.
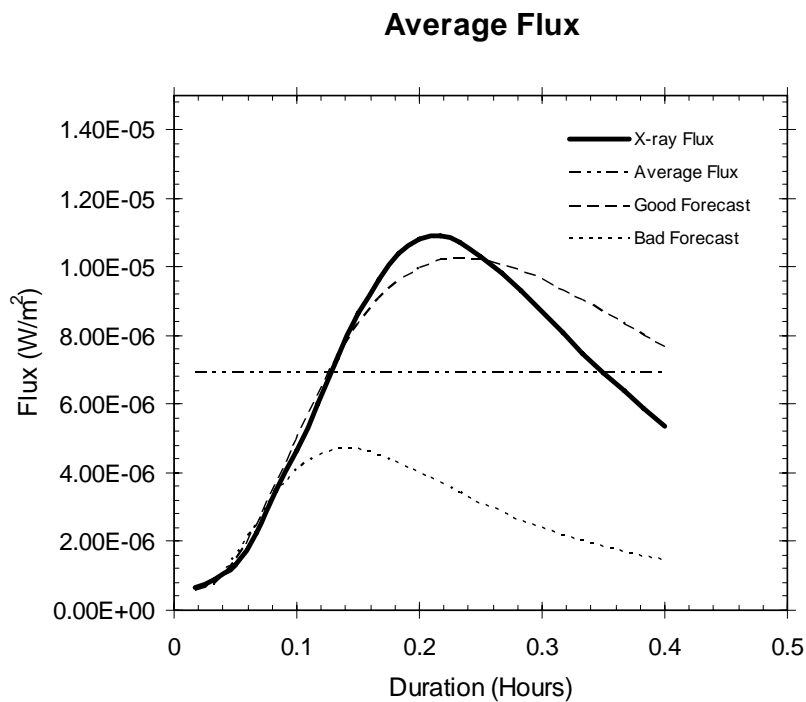
Average Flux

Figure 28. The horizontal line is the average flux measured for the flare. This line represents the "forecast" that is used to get the baseline RMSE. The dotted line is a forecast with an RMSE larger than the baseline and is considered a bad, or unusable, forecast. The dashed line is a forecast with an RMSE smaller than the baseline and is considered a good, or usable, forecast.

It is expected that the early forecasts, which only use a few data points, will have a

large RMSE. This is expected because there is not enough information, in just a few

points, to get the desired shape. As more points become available, the curve will be

53

closer to the desired shape because more information about the shape of the flux is available. As more points are used, the forecasts will improve and the RMSE will decrease. This will continue until the peak of the flux is reached and the symmetry forecast, which is typically the best forecast, is used. Table 8 shows the RMSEs for each forecast created for the C7b class flare from 3 Jul 2002. This shows how, for fewer data points, the RMSE is larger and decreases when more data points are used. It also shows that the five point forecast is a bad forecast since the RMSE is larger than the baseline RMSE. Similar behavior is seen for the other eight flares, just as was expected.

Table 8. RMSE values for each forecast for the C7b class flare from 3 Jul 2002. Positive numbers in the difference column indicate a good, usable, forecast while negative numbers indicate a bad forecast.

Data points used	RMSE	Baseline RMSE	Difference
5	0.265	0.264	-0.001
6	0.106	0.264	0.158
7	0.054	0.264	0.210
8	0.033	0.264	0.231
9	0.032	0.264	0.232

Figure 29 shows the C7b, 3 Jul 2002, forecast graphically. The forecast for five data points is not a very good forecast, as was indicated by the RMSE in Table 8. The forecasts for eight and nine points are good and are nearly identical, both in Table 8 and Figure 29. For this particular flare the shifting forecast was used for points 5, 6, and 7 while the non-shifting forecast was used for points 8 and 9.

The forecasts for an M flare follow a similar pattern. An example is shown in Figure 30 with an M1 flare from 29 Nov 1996. This flare utilizes both the shifting and non-shifting forecast methods. Again it is clear that the forecast for five data points is a bad forecast and as more data points become available the forecasts get better. In this

54

case the shifting forecast method is handling the flare very well. When the non-shifting

forecast, seen with 13 data points, is used, the decay rate of the flux is now more in-line

with the observed x-ray flux decay.
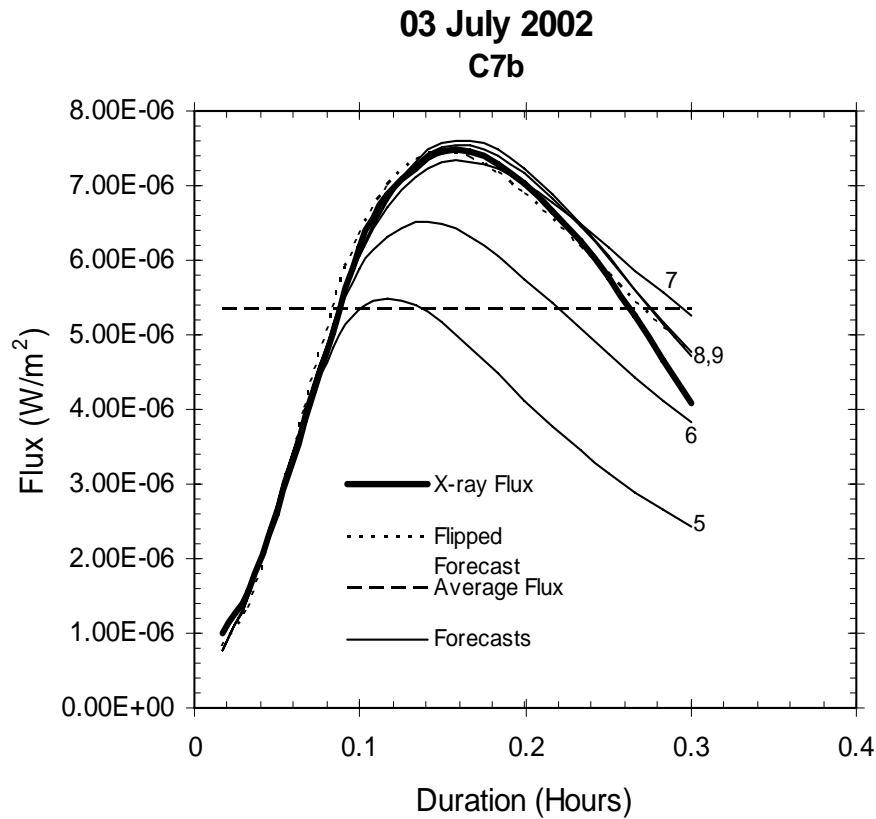
Figure 29. Forecasts for the C7b flare. The numbers on the right are how many data points were used to create that particular forecast. The forecasts for eight and nine points are nearly identical. This graphical look at the forecast corresponds to the numerical forecast in Table 8.

Although the forecast decay begins a little later than the actual flare, such that the

forecast will predict the end of the flare slightly later than the actual end of the flare, this

still produces an excellent forecast for the decay portion of the x-ray flux. This

demonstrates how, after enough data points are available, the non-shifting method can

produce a very accurate forecast.

55

29 Nov 1996
M1
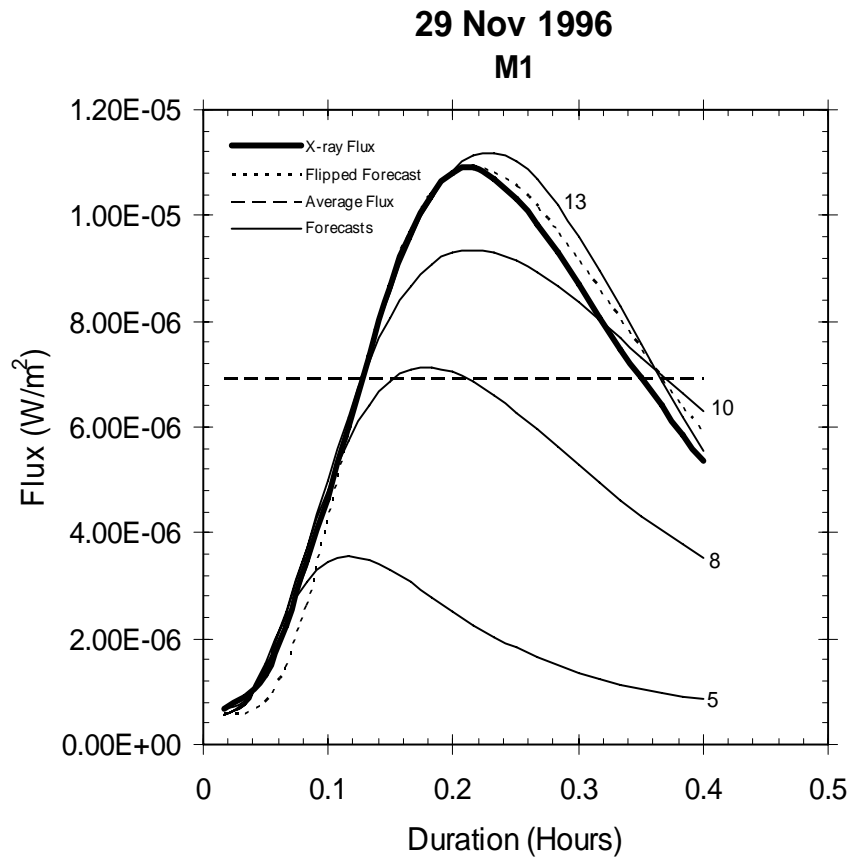
Figure 30. Selected forecasts for an M1 flare. The forecasts for 5, 8, 10, and 13 data points are shown. The 13th data point forecast uses the non-shifting forecast method.

As was mentioned before, the non-shifting forecast method isn't always used. Figure 31 shows an X1 flare from 19 Oct 2003. This is an example of when the shifting method is used exclusively. In this case the shifted forecasts handle the decay of the flux fairly well. As data points are added and the peak flux is approached, the forecast for the peak becomes better, but the decay portion becomes slightly worse. The forecast flux decay rate is a little too slow so at point 19 the end of the flare will be over-forecast. The forecasts for the remaining six selected flares are given in Figure 32. Each thin solid line is a forecast for a given number of data points.

56

These plots make it easy to see the progression of the forecasts as the flare steps through its lifetime. The forecasts are shown without labels to reduce clutter. It is clear, as with the previous three flares shown, that as more data points become available the forecast improves.

19 October 2003
X1
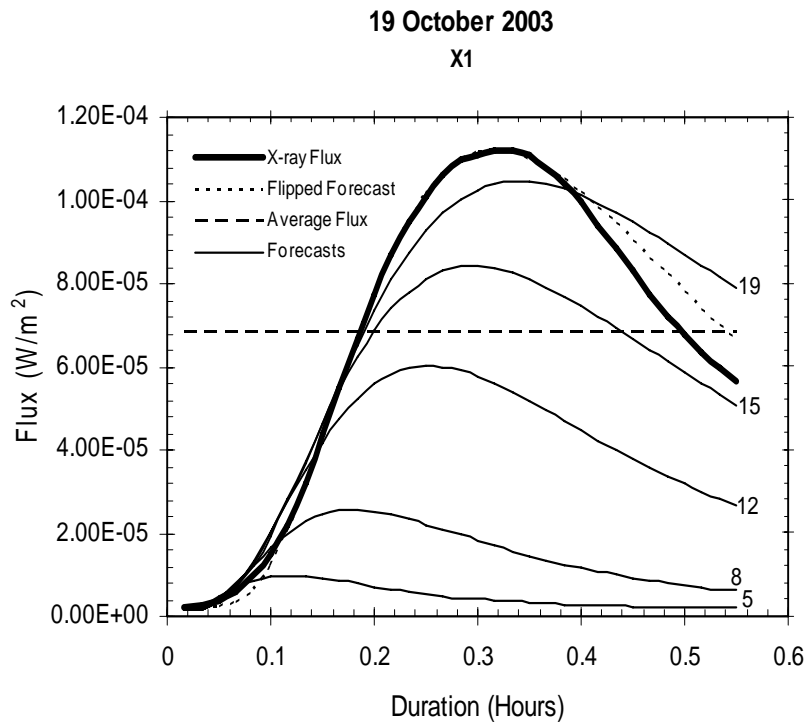
Figure 31. An X1 flare with selected forecasts. All forecasts are from the shifted forecast method. The non-shifted method was not used.

Two distinct situations arise from these forecasts. The first scenario is one in which only the shifting method is used for the forecasts, the second scenario is when both the shifting and non-shifting forecasts are used. Figure 32 demonstrates both of these. The shifting only scenario is shown by the plots in the left hand column while the right hand column shows flares that used the shifting and non-shifting scenario.

57
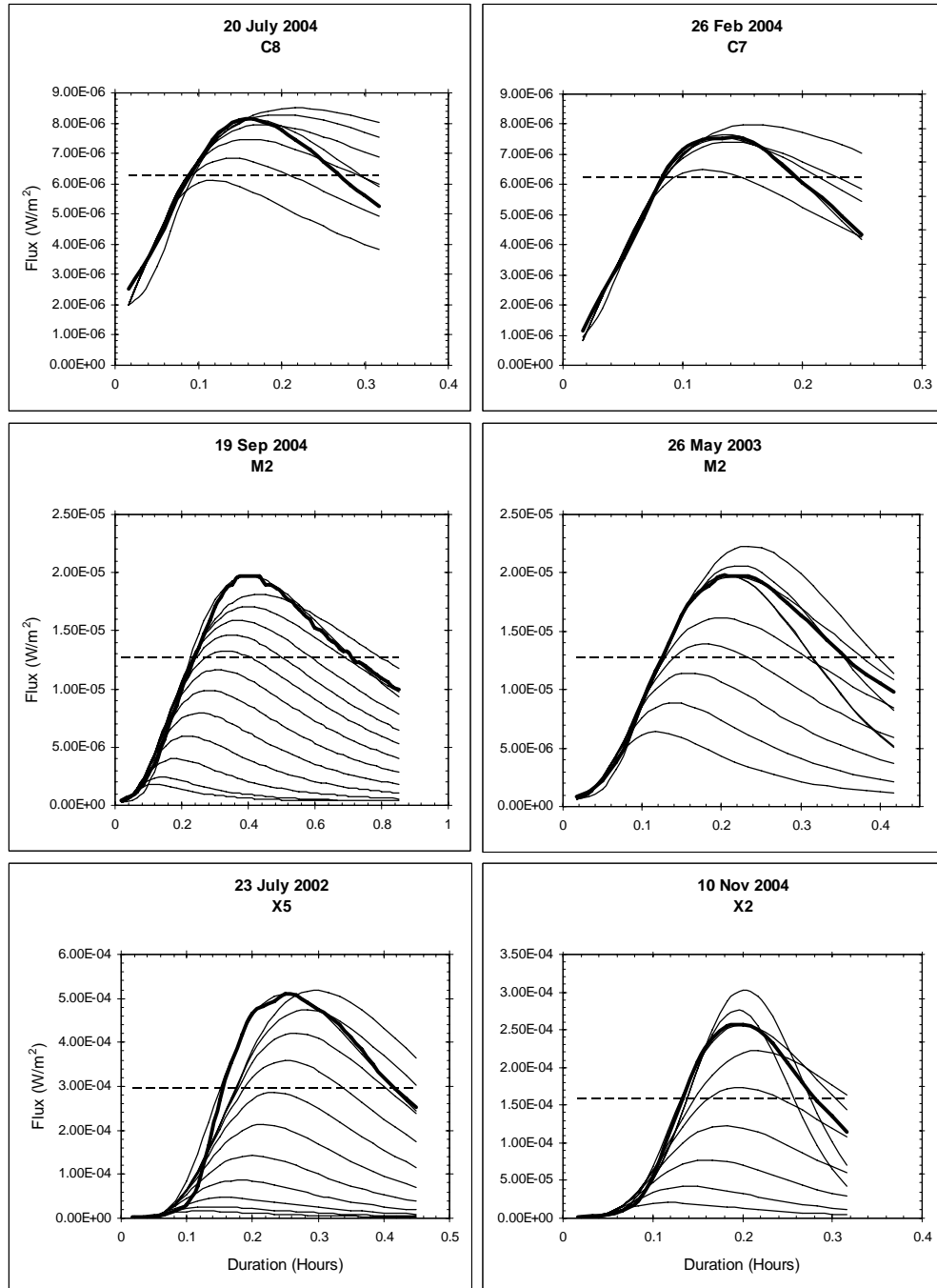
Figure 32. Automated forecasts for six flares. The thick dark line is the x-ray flux, the thin solid lines are the shifted or non-shifted forecasts, and the dashed line is the average flux. The forecasts improve as more data becomes available with the best forecast occurring once the peak has been reached or near to the peak. The left column uses the shifting forecast method exclusively while the right column utilizes the non-shifting method as well.

Some general characteristics of the two scenarios can be seen from these plots. One characteristic is that the shifting-only forecasts show a steady improvement as more data points are added to the rise. If the flare length is close to the length of the time shift then this steady improvement continues until the peak flux point is reached. If the flare length is not close to the time shift then a steady improvement occurs up to a point and then the forecast accuracy begins to decrease because the decay doesn't match up well with the actual flux. An example of this can be seen on the C8 flare in Figure 32. The forecast peaks get closer to the peak of the flare but the forecast decay doesn't line up well with the flare decay. This is due to shifting the data points over too far based on the flare. Figure 33 shows a more detailed look at the C8 flare and how the data points for the last forecast generated have been shifted over too far so that the forecast does not line up with the x-ray flux. A similar situation occurs with a different flare and can be seen in Figure 34.

Shifted Data Points
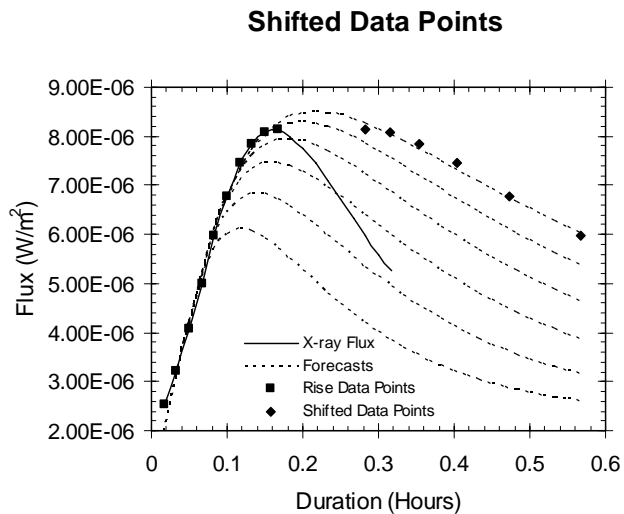
Figure 33. The C8 flare from 20 Jul 2004. Due to the use of a set time shift for the shifting forecast, the data points are sometimes shifted over too far. This creates a forecast that doesn't line up well with the x-ray flux.

59

For the shifting and non-shifting scenario, the steady improvement is seen during the shifting forecasts with an abrupt change in overall forecast shape when the non-shifting method is used, creating a forecast that fits better to the actual flux. The forecast is better because it eliminates the dependence on the estimate of flare length that was being used for the shifting method. This is most easily seen in the plot for the X2 flare in Figure 32.

Other characteristics can be seen by looking at the RMSE for each forecast made for each of the nine flares. Table 9 shows the nine selected flares identified by their x-ray classifications in the left most column. The value beneath the classifications is the average RMSE for that particular flare. Along the top are the numbers 5 to 24. These are the number of data points on the rise that were used to create the forecast. The table skips from 19 to 24 to accommodate the M2b flare which had 24 points on the rise. The values in the table are the RMSEs for each forecast. The asterisk next to the RMSE indicates that the non-shifting forecast method is used for that particular forecast. Forecasts are made until the peak is reached. Empty blocks indicate that the peak of the flare has occurred so no forecast was made. Table 9 shows how each of the nine selected flares has a decreasing RMSE as more data points become available. In general the RMSE decreases, but does take an occasional turn upward. This upward turn can be seen on the M2a flare, between data points 11 and 12. This is also seen on the X2 flare between points 11 and 12. Both of these upward changes in RMSE occur while the non-shifting method is being used, but not all flares that use the non-shifting method exhibit this behavior.

60

Table 9. Forecast RMSE values for the nine selected flares. The values 5-24 along the top are the number of flux data points used when creating the forecast with the RMSE for that forecast listed below. The value with the classification is the average RMSE. The asterisk (*) indicates when the non-shifting forecast was used. The a and b added to the flare classifications is to distinguish between the flares with the same classification.

	Forecast RMSE															
	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	24
C7a .253	.101	.083	.163	.014*												
C7b .264	.265	.106	.054	.033*	.032*											
C8 .199	.212	.108	.052	.086	.127	.156										
M1 .295	.536	.448	.346	.240	.139	.072	.093	.051*	.044*							
M2a .302	.565	.469	.355	.236	.123	.109*	.031*	.110*	.108*							
M2b .293	.704	.666	.649	.625	.597	.564	.528	.489	.448	.408	.367	.328	.290	.254	.221	.059
X2 .361	.604	.551	.455	.324	.182	.112	.094*	.119*								
X5 .359	.667	.656	.633	.588	.516	.421	.312	.203	.115	.088	.129					
X1 .331	.663	.644	.615	.576	.526	.467	.402	.335	.268	.205	.150	.106	.080	.077	.089	

It was stated that the RMSE should decrease as more data points became available, giving the smallest RMSE once the peak data point is reached. We can see that on five of the nine flares, the smallest RMSE occurs before the peak point. These flares are C8, M2a, X2, X5, and X1. There are two scenarios where this occurs.

The first scenario involves flares that used the shifting forecast method only. These are the C8, X5, and X1 flares in Table 9. Figure 34 shows the X1 flare of 19 Oct 2003 and selected forecasts. The left panel shows how each successive forecast gets closer to fitting the flux and then, for the top forecast, the decay moves beyond the observed decay and the forecast accuracy decreases. The right panel shows the same forecasts along with the data points that the polynomial for the top forecast was fit to. This illustrates how the shifting method can have the data points shifted over too far.

61

This is due to not knowing how long the flare is going to last and having to use a set time shifting value regardless of the actual length of the flare.
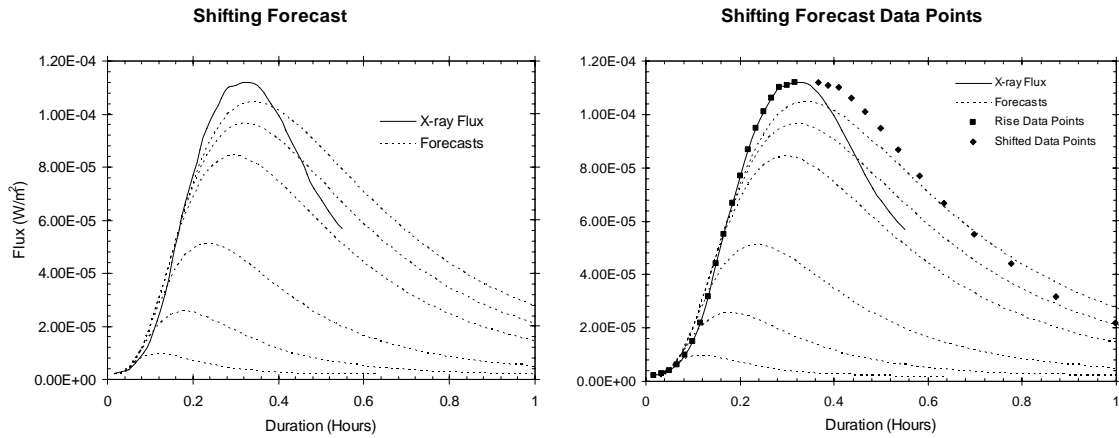
Figure 34. The X1 flare from 19 Oct 2003 that uses only the shifting forecast method. The left panel shows selected forecasts as dotted lines. The forecasts increase in quality as more data points are used. The bottom forecast uses five data points and the top forecast uses 19 data points. The 19 data point forecast decay does not accurately predict the decay of the flux.

The other scenario involves forecasts that use both the shifting method and the non-shifting method to create forecasts. The X2 and M2a flares in Table 9 fit into this scenario. The shifting method is used for the data points during the early minutes of the flare and then the forecasts switches to the non-shifting method toward the end of the rise phase. Figure 35 shows an X2 flare from 10 Nov 2004 with both the shifted forecasts and the non-shifted forecasts. The forecast for data point 10 is the last shifted forecast. The next forecast is for data point 11 and it is a non-shifted forecast that overshoots the peak and has a quick decay. This gives a poor peak forecast but a decent decay forecast. The next forecast, data point 12, does not overshoot the peak as much but the decay doesn't match up with the flux decay as well. This gives a good peak forecast but a poor decay forecast. The forecast that predicts the decay the best has the smaller RMSE and

thus the better overall forecast. In this case the forecast for data point 11 has the smallest

RMSE.

Shifting and Non-Shifting Forecasts
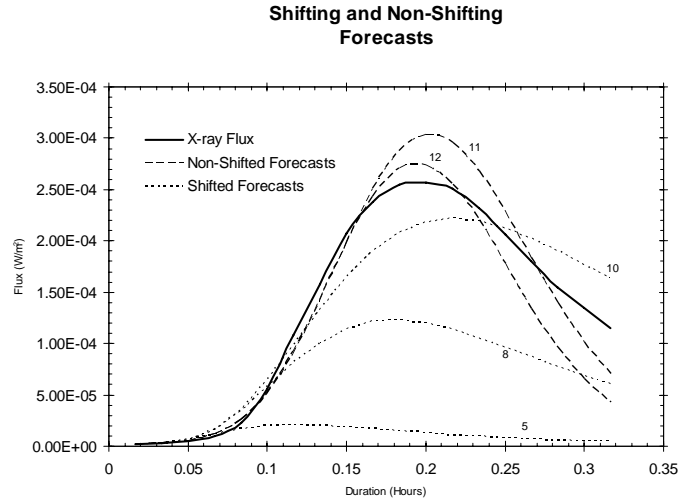
Figure 35. Forecasts for both the shifted and non-shifted methods for the X2 flare from 10 Nov 2004. The first non-shifted forecast (point 11) overshoots the peak of the x-ray flux more than the non-shifted forecast for point 12.

Whether the RMSE decreases all the way to the peak point, or there are some

small variations toward the end, the forecasts produced have moderate success predicting

the x-ray flux evolution as the flux nears the peak. A decreasing RMSE shows that the

forecasts are getting better as more data points become available. Another measure of

how well this method works is to determine how long after the flare onset before a good,

usable, forecast is created. The forecast from the symmetry forecast method couldn't be

generated until the peak flux was reached thereby using 100% of the rise portion of the x-

ray flux. The combined, shifting and non-shifting, method was devised in order to create

a forecast prior to reaching the peak flux. Using the results for the nine selected flares

found in Table 9, it can be determined how long it takes, after flare onset, before a usable

forecast is generated. Recall that a usable forecast is one in which the RMSE is smaller

than the baseline RMSE. Using the C8 flare (the third entry in Table 9) as an example, it can be seen that there are a total of 10 data points to the peak flux and that data point 6 is the first forecast that has an RMSE smaller than the baseline RMSE. This makes 60% of the rise known before a usable forecast is created. In contrast, the symmetry method uses 100% of the rise to get a forecast. So, the smaller amount of the rise of the flux used to get a usable forecast the better. The best scenario comes from the C8 flare where only 60% of the rise was known before creating a usable forecast. The worst case is the X5 flare that uses 73.3% of the rise before creating a usable forecast. On average 65.8% of the rise is known before a usable forecast is made. This is a noticeable improvement on the symmetry forecast which uses 100% of the rise.

Once the peak is reached the combined forecasting method is no longer used. The forecast then becomes the symmetry forecast method. Recall that this method takes all the data points up to the peak and flips them over to make a symmetric curve. A function is then fit to those points and becomes the forecast. This method is used because it has a greater accuracy than the other methods already used. Table 10 shows each of the nine selected flares with the best RMSE from the each of the methods that have been discussed. Six of the symmetry forecasts gave better results than the other forecast methods. There were three symmetry forecasts that were worse than the other forecast methods. Two of those were only worse by .001 and .003, which is a small margin. Based on these results, using the symmetry forecast when the peak is reached is the best option.

Table 10. A comparison between the best forecast from the shifting, non-shifting forecast, and the symmetry forecast methods. A dash indicates that this method was not used.

Flares	Best Shifting Forecast RMSE	Best Non-shifting Forecast RMSE	Symmetry Forecast RMSE
C7a	.163	.014	.061
C7b	.054	.032	.031
C8	.052	--	.055
M1	.072	.044	.038
M2a	.123	.031	.032
M2b	.059	--	.025
X2	.112	.094	.055
X5	.088	--	.043
X1	.077	--	.042

While not perfect, the combined forecasting method provides a decent forecast for each minute of the rise of the x-ray flux. In general, for each data point added to the rise phase it generates a forecast that was better than the previous forecast. It is also able to generate usable forecasts prior to reaching the peak. On average, this method generates the first usable forecast after less than two-thirds of the rise is available.

ال, المنارة للاستشارات

www.manaraa.com

VI. Conclusions and Future Work

6.1   Conclusions

The evolution of the x-ray flux can be forecast with some accuracy. The most accurate forecasting method is the symmetry forecast method. Using the fact that the x-ray flux follows an approximately lognormal shape, and therefore a normal shape when the natural logarithm of the time is taken, an accurate forecast of the decay portion of the flux is made. This method is successful because all of the data up to the peak of the flux is used. Only flares that have some sort of anomaly in the decay of the flux, which makes them not follow a lognormal shape, will be poorly forecast. One cause of anomalies in the decay is multiple flares occurring at the same time, thereby increasing the total x-ray flux, while the flare of interest is in the decay phase of its life cycle. These anomalies appear to be either the exception rather than the rule, or their impact is minimal based on the results of the symmetry forecast accuracy.

Climatology is the method currently in place to forecast the end time of a flare. Once the peak of the flux is reached, the average time to the end of the flare, based on climatology, is used as the forecast for the flare end time. Comparing the results of the symmetry forecast to the climatology forecasting method shows that the symmetry forecast does offer a significant improvement over climatology. An accurate forecast of the half-max time to within 2 minutes was made using the symmetry method for nearly three-quarters of all flares examined. Climatology was much worse with less than half producing the same kind of accuracy.

66

The symmetry forecast produces great results but it can't be implemented until the peak flux is reached. To overcome the need to wait for the peak flux, an algorithm was created to forecast the flux evolution prior to reaching the peak flux. Two methods, shifting and non-shifting, were employed in an effort to provide the best forecast possible. The shifting method took advantage of the symmetry of the flux in log time, while the non-shifting method had the advantage of not needing to know the duration of the flare, which is the disadvantage of the shifting method. The disadvantage of the non-shifting method is that it tries to create a curve based on very few data points, leading to inaccurate forecasts when just the first few data points of the rise phase are used. This leads to waiting until more data is available to get a usable forecast.

The combination of the two methods leads to some promising results. Usable forecasts can be generated when just under two-thirds of the rise phase is known. These forecasts have reasonable accuracy and continue to improve as more data points are added to the rise of the x-ray flux. While this gives a forecast prior to the peak flux being reached, the accuracy is not necessarily as good as the symmetry forecast method. Further refinement of this method could lead to a greater improvement on both timeliness and accuracy.

Overall, in terms of an accurate forecast, the symmetry forecast outperforms both the shifting and non-shifting forecast methods. The climatology method currently in use to forecast the duration of a flare is based on the average decay time for a flare of a given magnitude. The actual duration of a flare can stray from the average thereby reducing the accuracy of the forecast. It has been shown that the symmetry forecast method outperforms climatology and works on the majority of flares due to their lognormal

67

shape. Using the symmetry method could improve the current flare duration forecasts now in use.

Developing metrics and statistics for solar flare x-ray temporal evolution prior to reaching the peak of the flux is complicated due to the lack of current forecast methods. By not having other forecast methods available, there is no way to determine if this method is an improvement in forecasting skill. *Tobiska and Bouwer* [2005] have developed a method to forecast the x-ray flux evolution but are still developing the metrics for the prediction skill score.

6.2   Future Work

Throughout the data analysis phase it became clear that multiple improvements could be made to the computer code used to find the flares and produce the forecasts. Areas of improvement include the method used to scan through years of data to find the flares, the subtraction of the background flux, the determination of the length of the time shift, and the baseline RMSE comparison.

The current method of flare identification looks for three consecutive data points where the flux value for each successive data point is at least 1.2 times greater than the previous flux value. With this method it is possible for entire flares to be skipped over because the flux doesn't increase fast enough. Finding a better way to determine if a flare has started will reduce the number of flares skipped over.

The background is currently defined as the first flux value of the flare. *Tobiska and Bouwer* [2005] use a background index that is calculated from the previous hour of x-ray flux data. A better background separation would allow for a better representation of the x-ray flux from just the flare.

68

The length of time that the data points are shifted in the shifting forecast method is based on an average of a large number of flares. With a wide variance in flare lengths, it would be a great improvement to the shifting forecast method to develop a method that could estimate the length of the flare based on characteristics of the rise phase. With an accurate method to determine the shifting length, the forecast improvement could be great enough to eliminate the need for the non-shifting forecast method.

Nine flares were subjected to a thorough analysis. This is not enough to identify any trends or tendencies of the forecast algorithm. Applying an analysis to more flares should reveal a more accurate representation of the ability of this algorithm to provide early stage forecasts.

A list of the average time from the peak to the half-max value and the standard deviation for each flare classification. Asterisks indicate classifications with three or fewer observations. Omitted classifications indicate that no flares of that class were observed during 1996, 2002, 2003, and 2004.

Table A. Climatology of the average time from the peak x-ray flux to the half-max flux. Data taken from 1996, 2001, 2002, 2003, and 2004.

X-ray Classification	Number of flares	Average peak to half-max time (minutes)	Standard Deviation (minutes)
C1	235	6.7	3.9
C2	235	7.3	5.1
C3	145	7.9	4.7
C4	122	7.3	4.4
C5	73	7.1	4.4
C6	54	8.7	6.2
C7	50	8.4	5.7
C8	43	8.0	4.0
C9	34	8.4	5.6
M1	159	8.7	6.2
M2	78	9.5	7.5
M3	26	8.7	4.5
M4	18	8.2	4.8
M5	18	8.6	7.1
M6	9	8.2	5.4
M7*	3	5.7	1.9
M8*	3	9.3	2.6
M9	9	10.4	5.2
X1	26	11.4	13.4
X2	6	7.2	3.8
X3	6	15.0	7.8
X5*	2	13.0	1.0
X8*	1	14.0	0.0
X10*	1	11.0	0.0
X17*	2	15.0	2.0
All C	991	7.4	4.8
All M	323	8.9	6.3
All X	44	11.6	11.1

Bibliography

Aitchison, J., and J.A.C. Brown, *The Lognormal Distribution, with Special Reference to its uses in Economics,* Cambridge University Press, 1957.

Bornmann, Patricia, L, "Prediction of X-Ray (and Particle Event) Evolution using Temporal Lognormal Fits," National Oceanic and Atmospheric Administration Space Environment Center Seminar, 20 Jun 2000.

Campbell, Wallace H, "Geomagnetic storms, the Dst ring-current myth and lognormal distributions," *Journal of Atmospheric and Terrestrial Physics,* 58(10), 1171-1187, 1996.

Cravens, Thomas E, *Physics of Solar System Plasmas*, Cambridge University Press, 1997.

Department of Defense, *Performance Validation of DoD Ionospheric Models.* Joint Air Force Institute of Technology-Naval Research Laboratory Project Research Plan. 24 Jul 1997.

Golub, Leon and Jay M. Pasachoff, *The Solar Corona,* Cambridge University Press, 1997.

Lang, Kenneth R, *The Sun from Space*, Springer-Verlag Berlin Heidelberg, 2000.

Lu, Edward T., and Russell J. Hamilton, "Avalanches and the Distribution of Solar Flares," *The Astrophysical Journal* Vol. 380, 20 Oct 1991: L89-L92.

Phillips, Kenneth J. H, *Guide to the Sun*, Cambridge University Press, 1992.

Smithtro, Christopher G., Lecture notes, PHYS 519, The Space Environment. Graduate School of Engineering and Management, Air Force Institute of Technology, Wright-Patterson AFB OH, Fall Quarter 2004.

Space Environment Information System, "GOES 12 Query Form," 20 Oct 2005. *http://www.spenvis.oma.be/spenvis/help/models/databases/goes_12.html*

Tascione, Thomas F, *Introduction to the Space Environment*, Krieger Publishing Company, 1994.

Tobiska, W. Kent and S. Dave Bouwer, A065: "Solar Flare Evolution Model for Operational Users," Proceedings of the Ionospheric Effects Symposium, 2005.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to the Department of Defense, Executive Services and Communications Directorate (0704-0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE *(DD-MM-YYYY)* 06-03-2006	2. REPORT TYPE Master's Thesis	3. DATES COVERED *(From - To)* Sep 2004 - Mar 2006

4. TITLE AND SUBTITLE

Prediction of the Temporal Evolution of Solar X-ray Flares

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Williams, Aaron, J., Captain, USAF

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Air Force Institute of Technology
Graduate School of Engineering and Management (AFIT/EN)
2950 Hobson Way
WPAFB OH 45433-7765

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/GAP/ENP/06-21

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFWA/DN
Attn: Lt Col Trey Cade
106 Peacekeeper Dr
OFFUT NE 68113   DSN: 271-6177

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

13. SUPPLEMENTARY NOTES

14. ABSTRACT
A solar flare is an explosive release of stored magnetic energy on the Sun. Much of this energy is converted into x-ray photons which escape into space. As a solar flare begins, the 1-8 L x-ray photon flux at Earth's orbit, as measured by the GOES satellite, rapidly increases. It quickly reaches a peak and slowly decays. A plot of this flux exhibits an approximate lognormal shape. A lognormal function becomes a normal, symmetric, function when the logarithm of the independent variable is taken. Once the peak flux is reached, this symmetry is used to make a prediction of the flare end time. Examining over 1300 flares, an improvement in the flare end time prediction over the current climatological method is demonstrated. Predictions of the evolution of the flux prior to reaching the peak flux are also made beginning five minutes after flare onset. An effort to predict the temporal evolution prior to the peak is made using a fourth order least squares fit to the rise-phase flux alone and the rise-phase flux plus an estimate of the decay flux. Using both methods, accurate predictions of the x-ray flux evolution are made when the rise phase averages 65% complete.

15. SUBJECT TERMS
Solar Flare, X-ray, Lognormal, Evolution, Prediction, Forecast

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Christopher G. Smithtro, Maj, USAF (ENP)
U	U	U	UU	83	19b. TELEPHONE NUMBER *(Include area code)* (937) 255-3636, ext 4505

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18