Note added January 14, 2015: Some of Riley's statistics are ambiguously reported in his paper. Since the author has ignored all requests for clarification, I had to use my best guess to resolve the ambiguities. New evidence since the present paper was posted has caused me to change my best guess. This changes most of the probability estimates reported below.

This new evidence together with revised probability estimates is reported in a reanalysis of Riley's paper in file www.math.umb.edu/∼sp/2ndlook.pdf. It is discussed and can be accessed from the January 14, 2015 entry of the "papers" page on my website www.math.umb.edu/∼ sp . I have not changed the estimates reported below in order to maintain as a historical document the "Comment" below which was submitted to *Space Weather*. Later it was resubmitted with the last two paragraphs of the first section omitted at the request of the Chief Editor.

Comment on "On the probability of occurrence of extreme space weather events", by P. Riley[1]

1 Introduction

The recent paper [1] under review attempts to estimate the probability that a solar storm as bad as the Carrington event of 1859 will occur in the next decade. A National Academy of Science report [3] estimates the damage from such a storm in the *trillions* of dollar. The paper settles on an estimate of 12% probability, which has been widely reported in the popular press in contexts suggesting that it is scientifically reliable.

A careful study of [1], has uncovered so many questionable assumptions and mistakes that in my opinion, this 12% estimate should be cited only by those who have carefully read the paper, independently checked its calculations, and are prepared to vouch for it.

It is unfortunate that I have not had the opportunity to reconcile the differences with the author, who has not responded to any of my inquiries. That is the way that anomalies should be resolved, and any that arise from genuine misakes should be corrected in an erratum. The lack of response is the reason for this Comment which I should have preferred not to submit.

Space limitations prevent full documentation here of the anomalies. Complete documentation is posted on the August 25 entry of the "papers" page of www.math.umb.edu/∼sp, where further discussion may be added. The URL of the document itself is www.math.umb.edu/∼sp/analysis.pdf .

[1]P. Riley, *Space Weather* **10** (2012), S02012

2 Incomplete definition of "power law distribution"

Many of the paper's ambiguities and apparent errors can be traced to an incomplete definition of "power law probability distribution". The paper's equation (1) incorrectly states that:

> "Here we outline the basic tools we will employ to compute the probability of occurrence of an extreme space weather event. A set of events, x, is said to follow a power law distribution if the probability of occurrence, $p(x)$, obeys the following relationship:
>
> $$p(x) = Cx^{-\alpha}, \tag{1}$$
>
> where the exponent α, is some fixed value and C is a constant determined from where the power law intercepts the y axis."

This should read something like:

> A random variable X taking on positive real values is said to have a *power law distribution* with exponent α if its probability density function (pdf) $p(x)$ is of the form
>
> $$p(x) = \begin{cases} Cx^{-\alpha} & x_{min} < x < \infty \\ 0 & \text{otherwise,} \end{cases} \tag{1}$$
>
> where $\alpha > 1$, $C > 0$, and $x_{min} > 0$ are constants.

The paper's statement that "C is a constant determined from where the power law intercepts the y axis" is wrong—for $\alpha > 0$, the graph of $x \mapsto Cx^{-\alpha}$ is asymptotic to the y axis but never intercepts it. The inclusion of x_{min} in the definition for $\alpha > 1$ is essential; otherwise $\int_0^\infty p(x)dx$ will diverge at the lower limit.

This is important because the analyses of the various data sets never specify which x_{min} is being used to produce the probability estimates, though this can sometimes be guessed from the figures. There is internal evidence that in at least one case, an inappropriate x_{min} may have been used.

The omission of this important piece of information and vagueness in the exposition elsewhere makes some detective work necessary to check the paper's calculations. I have done the work and found that the resulting probability estimates often differ significantly (sometimes by an order of magnitude) from those reported in the paper.

3 Results and discussion

The aim of the paper is to obtain estimates for the probability of a geomagnetic storm worse than the Carrington event in the next decade. For brevity, when probabilities are mentioned below, they will always refer to this probability unless otherwise specified.

Separate probability estimates are obtained from three data sets under various assumptions. These vary from 1.5% (p. 8, paragraph [39]) to 85% (p. 6, paragraph [31]). The three data sets are Coronal Mass Ejections, Geomagnetic Storms, and Nitrate Records (from ice core samples).

Coronal Mass Ejections

The paper first assumes that the Coronal Mass Ejection (CME) data have a power law probability density function for CME speeds at least 700 km/sec because the Figure 4(a) histogram of speeds looks more or less that way. This results in a probability estimate of 85%, which the author regards as "not credible".

There is evidence that the paper's analysis may have used an incorrect $x_{min} = 100$ instead of the correct $x_{min} = 700$. However, a correct analysis would have resulted in an even *greater* probability, probably over 99%. The point is that the "not credible" estimate was not due to a possible mistake, but instead shows that the power law assumption applied to observed speeds (all less than 3500 km/sec) was incapable of credibly predictioning the probability of speeds greater than the assumed 5000 km/sec of the Carrington event. Details can be found in [2].

The paper then restricts the data set to speeds over 2000, and repeats the analysis, obtaining a probability of 12%. I carried out an independent analysis using the original data from the website from which the paper obtained its data and obtained a comparable probability estimate of 14%.

The paper does not report that this 12% estimate derives from the restricted data set containing only about 20 observations. (It later expresses skepticism about the results of the nitrate deposit analysis because that data set contains only 70 observations.) It does not discuss why the result of an analysis based on speeds over 2000 should be more reliable than the "not credible" result obtained for speeds over 700.

Geomagnetic Storms

Based on the paper's reported data, I repeated the arithmetic and obtained a probability estimate of 78% compared to the paper's 12%. Of course, the 78% estimate is not credible. If it is not due to errors in the data presented in the paper, I would regard it as evidence that the paper's assumption that the data follow a power law is unrealistic.

Nitrate deposits in ice core samples

Again I repeated the arithmetic and obtained probability between 18% and 25% compared to the paper's 3%. My probability estimates were a range rather than a single probability because of uncertainty about x_{min}, which the paper never reports.

4 Conclusions

The paper's analysis is based on an unjustified assumption that the data follow a power law. In one instance this led to a probability estimate which the paper characterizes as "not credible". There is enough evidence of probable mistakes to justify a recommendation that the paper's probability estimates be independently verified before being cited or used for policy purposes.

References

[1] P. Riley, "On the probability of occurrence of extreme space weather events", *Space Weather* **10** (2012), S02012

[2] www.math.umb.edu/∼sp/analysis.pdf

[3] National Research Council. Severe Space Weather Events–Understanding Societal and Economic Impacts: A Workshop Report. Washington, DC: The National Academies Press, 2008. available at http://books.nap.edu/catalog.php?record_id=12507