FOREST FIRES AND OIL FIELDS AS PERCOLATION PHENOMENA.

William J. Reed*

JUNE, 1999.

Abstract

A probability distribution derived from percolation theory is fitted to large datasets on the sizes of forest fires and of oil-fields, providing excellent fits. The results lend support to modelling forest fires and oil fields as percolation phenomena, as well as suggesting a new size distribution model which may be useful for estimating oil reserves and for forest management under the 'natural disturbance' paradigm.

Forest fires and oil fields are often cited as examples of phenomena exhibiting percolation behaviour *(1,2)*. Whether a percolation model can be anything more than a "toy" model for such complex phenomena is open to question, although there have been some studies *(3,4,5)* which suggest that percolation modelling can at least provide qualitative insights into the growth and spread of fires. The purpose of this note is to present the results of a statistical analysis of empirical size distributions of forest fires and oil fields, which indicate consistency with percolation theory. Furthermore the results

*Department of Mathematics and Statistics, University of Victoria, PO Box 3045, Victoria B.C.,Canada V8W 3P4. e-mail reed@math.uvic.ca. Research supported by NSERC grant OGP 7252.

suggest a simple parametric form for the size distributions which could lead to better estimates of oil reserves, and be of use for ecosystem based forest management under the ‘natural disturbance’ paradigm *(6)*.

If forest fires and oil fields can be adequately represented by a site percolation then the final extent of individual forest fires and of oil fields can be considered as percolation clusters and their frequency-size distributions should reflect cluster size distributions (cluster numbers) in two and three dimensions. The exact distribution of cluster size is not known. However there are some partial results available. Below the percolation threshold all clusters are finite and their size distribution exhibits a crossover phenomenon. Below a crossover level s_0, the probability mass function (p.m.f) for cluster size S behaves like $s^{-\tau}$ *(7)*, while above it behaves asymptotically like $s^{-\theta}\exp(-cs)$ $(s \to \infty)$ *(1)*. Its behaviour at and immediately above the crossover level s_0 is not known, although s_0 is known to be related to the pair connected correlation length ξ and the fractal dimension of clusters, d_f, by $s_0 = \xi^{d_f}$ *(7)*.

Areas burnt by forest fires and volumes of oil fields are best represented as continuous variates, with distributions described by a probability density function (p.d.f). The form used for fitting to the data assumes that the density behaves like $x^{-\theta}\exp(-cx)$ for *all* sizes, x, above the crossover. It can be written

$$f(x) = \begin{cases} p\,\frac{1-\tau}{x_0^{1-\tau}}\,x^{-\tau}, & \text{for } 0 < x < x_0 \\ (1-p)\,\frac{x_0^{\theta-1}}{E_\theta(cx_0)}\,x^{-\theta}e^{-cx}, & \text{for } x \geq x_0 \end{cases} \qquad (1)$$

where $E_\theta(y)$ is the exponential integral $\int_1^\infty \frac{e^{-yt}}{t^\theta}\ dt$. Note that $f(x)$ may not be continuous at the crossover size x_0, and that it integrates to p over $0 < x < x_0$, and to $(1-p)$ over $x \geq x_0$. This distribution has five parameters (τ, θ, c, x_0 and p) which can be estimated by maximum likelihood (ML) *(8)*.

This model was fitted to three sets each of data on fire size and oil-field size. The fire data comprised records of fire size (a) on U.S Forest Service land in the Sierra Nevada (2536 fires recored between 1910-1992) *(9)*; (b) in Nez Perce National Forest, Idaho (1795 fires between 1900-1994) *(10)* and (c) in Clearwater National Forest, Idaho (884 fires between 1900-1994) *(10)*. Fig.1 displays the results of the fits. The Q-Q plots in the lower panels are extraordinarily close to the 45° line indicating superb model fits. The ML estimates of τ were comparable for the three zones (0.664, 0.607 and 0.847) as were those of θ (1.44, 1.55 and 1.54) and c (8.08e-5, 6.55e-5 and 1.61e-5).

The oil-field data comprised volumes of discovered oil fields, *(11)*, in (d) West Siberian Basin (634 fields), *(12)* the oil province ranked largest in the world *(13)* ; (e) Alberta Basin (361 fields, ranked 19) *(13)* and (f) Denver Basin (742 fields, ranked 123) *(14)*. The data for (e) and (f) came grouped into classes of equal width on the log scale. Fig 2 displays the results of the fits. Again the fit is excellent. The kink in the Q-Q plot for W. Siberian Basin is artificial, being due to the fact that the volume of many fields in the dataset was rounded to 2.00 million barrels. Also the observed quantiles in the Q-Q plots for the grouped data (Alberta and Denver) are really upper bounds for the quantiles, being in fact the upper cell boundaries. The ML estimates for τ (0.101, 0.456 and 0.583) showed some varaibility as did those of c (2.50e-4, 2.02e-3 and 5.04e-3), while those for θ were more similar (1.31, 1.20 and 1.53).

The exceptionally good fits of the model to these large sets of data offers support for the application of percolation theory for modelling forest fires and oil fields. In both cases the distribution (1) fitted the data better than the more familiar ones which have been proposed (*e.g.* lognormal, generalized Pareto *etc.*). Also it is similar to a multifractal model which has been proposed for mineral deposits *(15)*. Since the assumed form of the underlying field-size distribution is critical in the estimation of total oil and gas reserves, the use of this new distribution could lead to improved estimates and better decisions concerning the development and production of hydrocarbon reserves as well as long-tem energy policy.

References and Notes.

1. D. Stauffer and A. Aharony. *Introduction to Percolation Theory.* (Taylor and Francis. London, 1992).

2. A. Bunde and S. Havlin. In *Fractals and Disordered Systems.* A. Bunde and S. Havlin Eds. (Springer-Verlag, Berlin, 1991), pp.51-93.

3. T. Beer and I. G. Enting. *Mathl. Comput. Modelling* **13** 77 (1990).

4. T. Beer *Combust. Sci. and Tech.* **72** 297 (1990)

5. B. D. Malamud, G. Morein and D. L. Turcotte. *Science* **281** 1840 (1998).

6. M. L. Hunter. *Biol. Conserv.* **65** 115 (1993).

7 A. Aharony. In *Fractals and Disordered Systems.* A. Bunde and S. Havlin Eds. (Springer-Verlag, Berlin, 1991), pp.151-173.

8. It is not necessary to maximize simultaneously over 5 parameters. The following procedure involves numerical maximization over one and over two parameters. First fix the level x_0. The MLE $\hat{p}$ of p is simply the proportion of observations less than x_0. Also for the given x_0 the MLE $\hat{\tau}$ can be easily be found numerically using only the observations less than x_0; while $\hat{\theta}$ and $\hat{c}$ can be found using only the remaining observations. Thus a profile (or concentrated) log-likelihood depending only on x_0 can be found. Maximizing this with respect to x_0 provides the MLE $\hat{x_0}$, and consequently the MLEs of the other paarmeters.

9. *Sierra Nevada Ecosystem Project Final Report to Congress* (Centers for Water and Wildlife Resources, University of California, Davis, 1996). Vol II, pp.1119-1138.

10. K. S. McKelvey, USDA Forest Service, Missoula, MT. *pers. comm.*

11. The density (1) may hold even if the probability of discovery is size-dependent (*e.g.* M. Power *Math. Geol.* **24** 929). If the underlying size distribution is of the form (1) and the probability of discovery is proportional to volume or some power of it, then the distribution of *discovered* fields will also be of the form (1).

12. *Oil and Gas resources of the West Siberian Basin, Russia* (Energy Information Administration, U. S. Dept. of Energy, Washington, D.C., 1997) DOE/EIA-0617, pp 157-172.

13. *Ranking of the World's Oil and Gas Provinces by Known Petroleum Volumes* (U. S. Dept. of the Interior Geological Survey, 1997) Open-File Report 97-463, available at *http://energy.cr.usgs.gov:8080/energy/WorldEnergy/OF97-463.*

14. S. Javad Seyedghasemipour and B. B. Bhattacharyya *Math. Geol.* **22** 557, (1990).

15. F. P. Agterberg *Inter. Geol. Rev.* **37** 1 (1995).

16. I thank Kevin McKelvey, USDA Forest Service, Missoula, MT. for providing forest fire data, and for stimulating my interest in fire-size modelling.

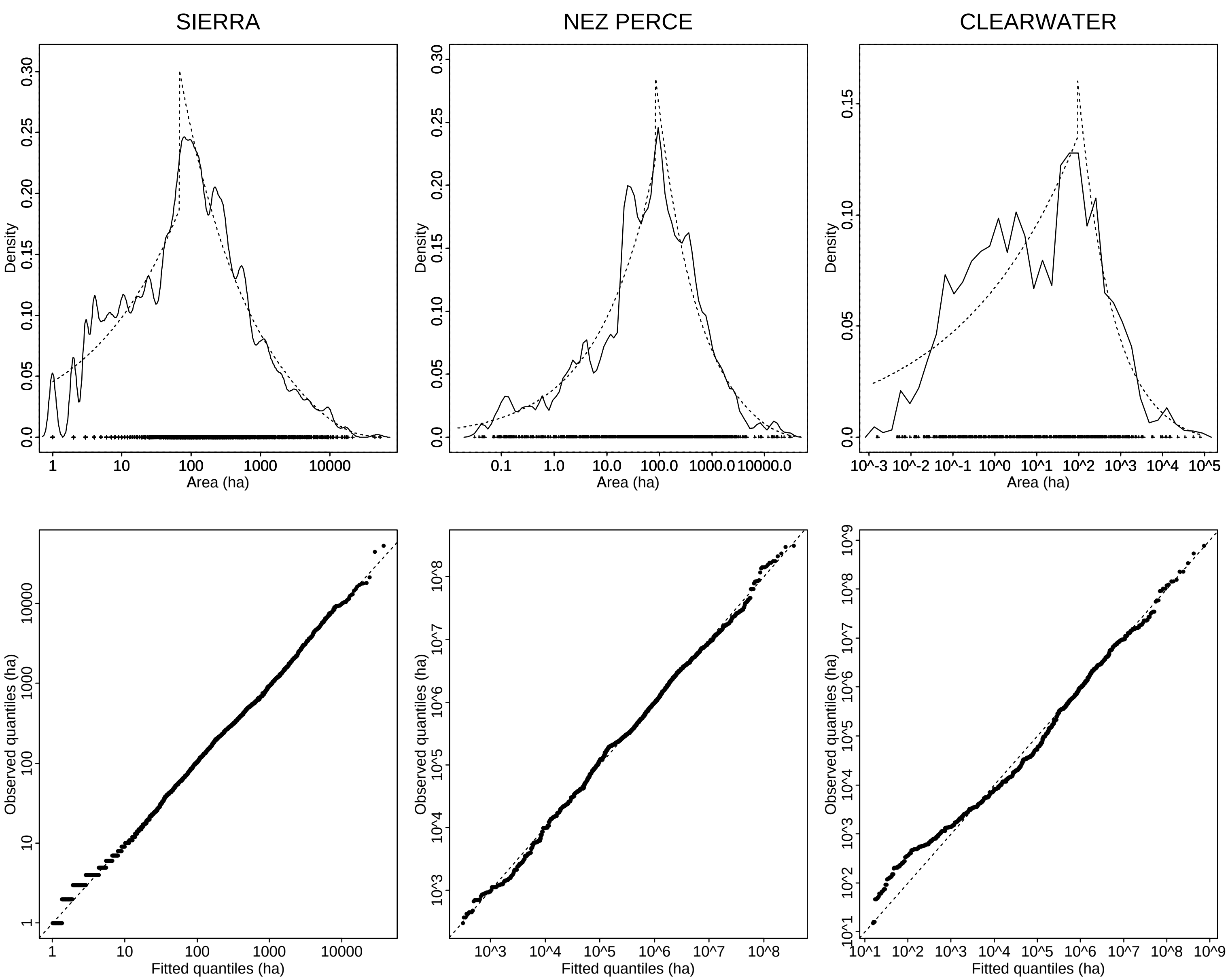

Figure 1: Top panels show non-parametric density estimates (smoothed histograms) for fire areas (log scale) in three regions of Western U.S.A (see text) along with the ML estimates of the density (1) derived from percolation theory (dotted lines). Also shown at the bottom (crosses) are the observed areas. The bottom panels show the quantiles of the empirical area distributions (vertical axis - log scale) against the quantiles of the fitted distribution derived from percolation theory. The closeness to the 45° line (especially for Nez Perce and Sierra) indicates the exceptionally good fit of the theoretical distribution.

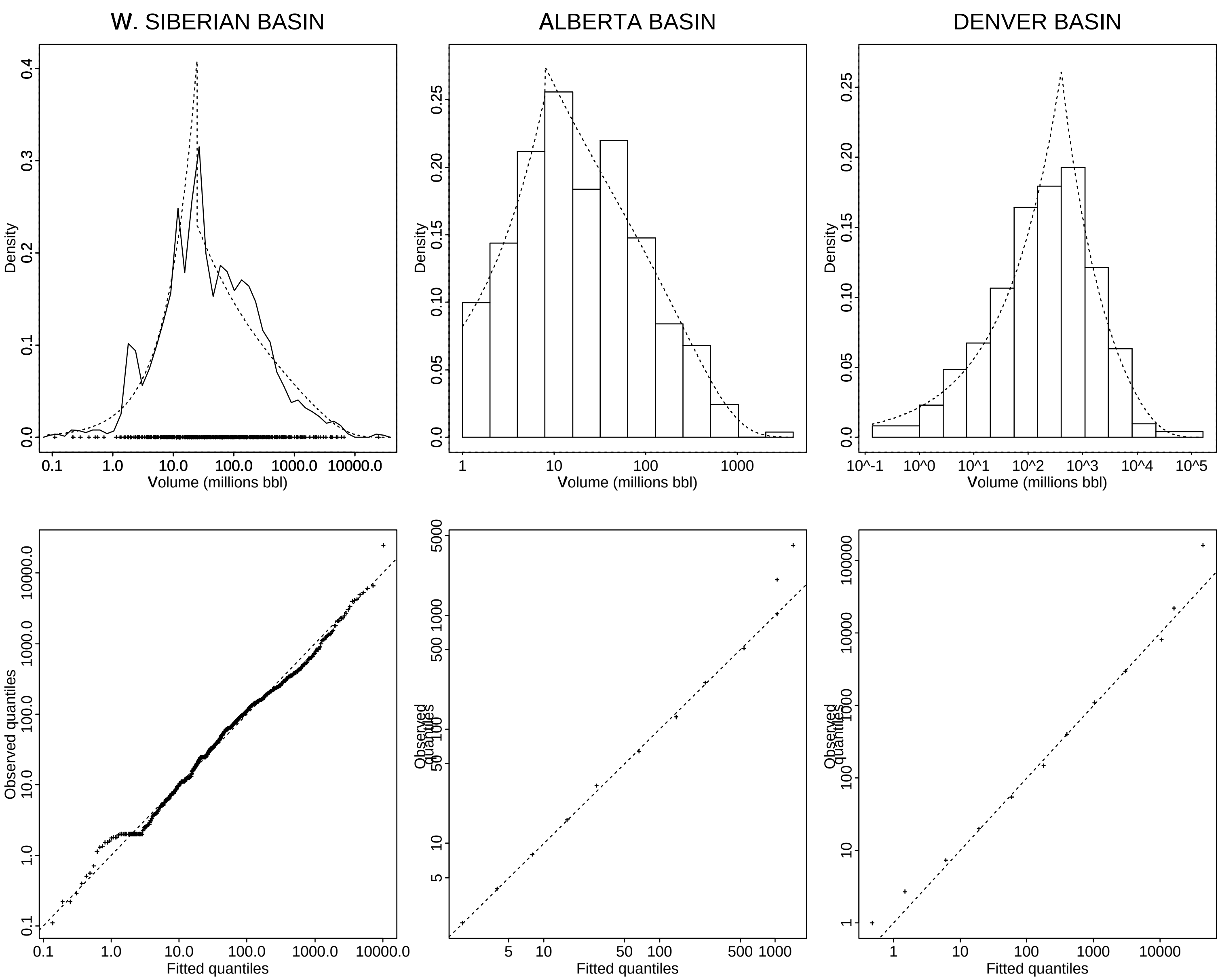

Figure 2: Top left-hand panel show a non-parametric density estimate (smoothed histogram) for oil field size (log scale) in West Siberian Basin along with the ML estimates of the density (1) derived from percolation theory (dotted line). Also shown at the bottom (crosses) are the observed volumes. The other two upper panels show density histograms (from grouped data) and ML estimates of the density (1) derived from percolation theory (dotted lines) for Alberta and Denver Basins. The bottom left-hand panel shows the quantiles of the empirical volume distributions (vertical axis - log scale) against the quantiles of the fitted distribution derived from percolation theory. The other two lower panels show the upper bounds of the cells into which data are grouped (providing upper bounds for the quantiles of the empirical distribution) against the corresponding quantiles of the fitted distribution derived from percolation theory. The closeness of the points to the 45° line indicates the good fit of the theoretical distribution.