The story behind the paper

Diego Rybski
*Potsdam Institute for Climate Impact Research — PIK,
Member of Leibniz Association, P.O. Box 60 12 03, Potsdam 14412, Germany
University of California Berkeley, Department of Environmental Science,
Policy and Management, 130 Mulford Hall #3114, Berkeley, CA 94720, USA and
Complexity Science Hub Vienna, Josefstädterstrasse 39, A-1090 Vienna, Austria*
(Dated: August 2, 2021)

Most scientific publications tell a different story about their findings than the one that really led to them. Usually, there is a kind of path which the researchers follow and on which they step by step approach the result. However, when presenting the results and findings the story is usually told from a different perspective, namely from what is known at the end, and the process of understanding is often omitted. In this essay, I would like to give the background of one of my recent publications [1], about the Forest Fire Model (FFM) [2, e.g.]. I taught the FFM over many years in a lecture to students from the Earth and Environmental Sciences based on secondary literature [3]. At the time I had a rather superficial understanding and mostly explained qualitatively.

The FFM [4] is an iterative simulation on a grid. Each cell can have three states, namely “empty”, “tree”, and “fire”. New trees grow on empty cells with a given probability and lightnings strike tree cells with another probability. A fire cell is always converted to an empty cell and a tree cell catches fire if at least one neighbor is a fire cell. In this version of the model, let’s call it DS-FFM after the authors [4], there can be regrowth while the fire is still burning. Another version, the H-FFM after [5], is more realistic since an entire connected tree-cluster turns into empty cells if at least one tree is hit by a lightning.

During my PhD I investigated long-term persistence in climate records. It is a special form of memory in time series, where the auto-correlation function decreases as a power-law and accordingly correlations can be found over very long time-scales [6, e.g.]. While teaching the FFM I always wondered if it also generates long-term persistence. My intuition said yes, but the memory should vanish because of the finite system size – there is simply not enough “space” to store the information infinitely.

When I came to Berkeley [7], in 2020 we experienced the smoke of the largest wildfires that were ever registered in California (Fig. 1). I took this as an occasion to finally implement the FFM. The version with spreading fires (DS-FFM) is really simple to implement and for the version where the entire clusters burns in one iteration (H-FFM) I could use code from our City Clustering Algorithm [8, 9]. Then I extracted time series of the shares of tress or fires from the H-FFM.

For my PhD I had used Detrended Fluctation Analysis (DFA), a powerful method to quantify long-term correlations. When analyzing the FFM data with DFA I always found oscillations. DFA is sensitive to oscillatory trends

FIG. 1. The sky over Berkeley, CA on 9.9.2020.

[10] and in this case they made it impossible to quantify the persistence. This was a bit disappointing because whatever I did, there was no simple way to get rid of the oscillations.

Then I started thinking and asked myself what the reason for those oscillations could be. It reminded me of the logistic map [11, e.g.] where for certain parameter ranges also periodic patterns occur. The question was, how a corresponding equation could look like for the FFM. I came up with

$$y_{t+1} = y_t + p(N^2 - y_t) - qaN^2(p_c - y_t/N^2)^{-\gamma}\,, \quad (1)$$

where y_t is the share of trees at time step t, p and q are the parameters of tree growth and lightning, respectively, N is the system size, a and γ are parameters, and $p_c \approx 0.593$ is the critical percolation concentration [12]. The second term on the right hand side represents the regrowth and the third term represents the disappearing clusters, based on the average size of a finite cluster as known from percolation theory [12]. For $N = 1$ and $q' = qa$ it can be rewritten as

$$y_{t+1} = (1 - p)y_t + p - q'(p_c - y_t)^{-\gamma} \quad (2)$$

which has some qualitative similarity with the logistic map $x_{t+1} = rx_t - rx_t^2$.

Consequently, I calculated the Feigenbaum-diagram for Eq. (2), an example is shown in Fig. 2. I was excited although it was clear that Eq. (2) is not correct – the third term only quantifies the finite clusters but above the critical threshold ($y > p_c$) the infinite cluster plays an important role (and it was clear that the system

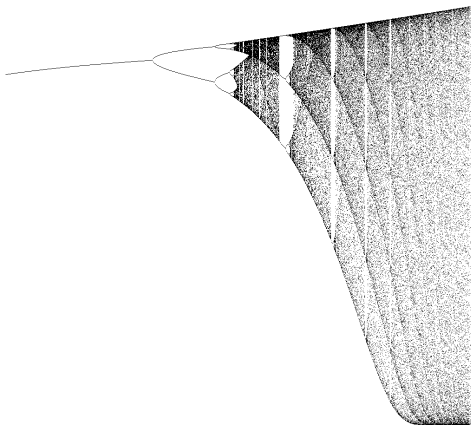

FIG. 2. Examplary Feigenbaum-diagram of Eq. (2), where p is plotted horizontally and $\exp(y)$ vertically, q' is constant.

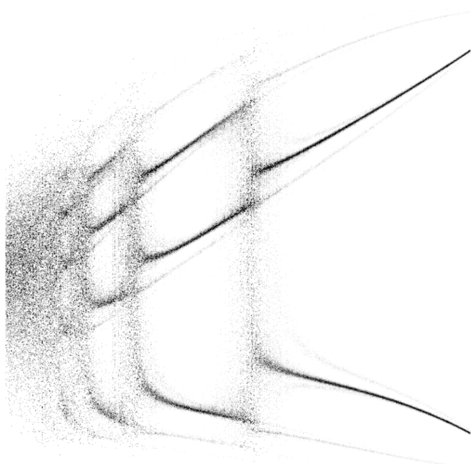

FIG. 3. Exemplary Feigenbaum-diagram for the H-FFM [5] for a system of size 200×200, p between 0.1 (left) and 0.9 (right), and $q = 0.0001$. The tree density is plotted between 0 (bottom) and 1 (top). It is one of the first Figures that I generated in the context.

could enter this regime). Nevertheless, I was speculating that the FFM might exhibit similar rich complexity as the logistic map.

Next I drafted a one-pager summarizing my thoughts and sent it to a few colleagues including some who published in the field. They did not seem to share me excitement, which is understandable since the model is 30 years old and numerous papers have been published. My PhD adviser said I should continue investigating only if I can relate it to real-world data. Someone else correctly noted that Eq. (2) is a mean-field approximation. However, no one was aware of any previous study going in this direction. The only colleagues who engaged more were my co-workers with whom we then wrote the paper.

I continued working and an important step was to calculate a Feigenbaum-diagram for the actual H-FFM and not for the mean-field approximation. An early result is shown in Fig. 3. At that point it was clear, something interesting is going on. By exhibiting a chaotic regime and periodic regimes, this result indeed shares some similarities with the logistic map, but also dissimilarities. Of course there was also luck involved. Had I not used the H-FFM, where entire clusters disappear when struck by a lightning, I had not found this result, since the simpler DS-FFM, where fires spread from cell to cell, does not lead to pictures like Fig. 3 (at least not according to preliminary tests).

Subsequently, I decided to intensify my efforts for this work. The decision was not per-se clear since usually I work in a different field, namely with cities as complex systems, and this FFM work would divert my research focus and further widen my already rather broad research profile. However, in favor of freedom of research I felt some obligation to write a paper on the topic.

We then systematically investigated the system and created nicer Feigenbaum-diagrams. Using Fast Fourier Transform we analyzed the power spectra of the tree density records and indeed found quasi-oscillatory fluctuations in the chaotic regime (small p). Interestingly, such oscillatory behavior emerges without being externally imposed. We further found that the frequency is approximately equal to the tree growth probability p for which we also provide an analytical expression. At lower frequency (while still $p \gg q$) the power spectra exhibit scaling – but contrary to $1/f$ noise (as mentioned in various papers [2, e.g.]) it is $1/f^2$ noise, i.e. a form of Brownian Motion. Accordingly, my initial hypothesis about long-term correlations in the system was confirmed. However, strictly speaking it is a special form, since Brownian Motion can be obtained by summing-up uncorrelated noise. For very long records, the power spectra become flat, indicating the absence of persistence, as also expected. Last we studied how the average and the standard deviation of the tree density depend on the tree growth probability p and found that the latter shows some regularity.

Overall, we found (i) periodic attractors for large p, which motivated us to introduce *self-organized multistability*, (ii) Brownian motion at large time scales, (iii) quasi-periodic oscillations, where the frequency is approximately proportional to p, and (iv) square-root dependence of the standard deviation on p and the system size. We ended up not using the wrong mean-field approximation Eq. (2). I am curious how the community will accept our work and which follow-up papers it will trigger.

To give an idea of the time-scales I want to list some key-dates. I started sending out the one-pager on 30.9.2020 and less than a week later (5.10.2020) I generated Fig. 2. Jan W. Kantelhardt presented our results at a physics colloquium in Halle (14.1.2021) and at the spring meeting of the German Physical Society (23.3.2021, DY 19.4). The first submission of the paper we sent to the best physics journal on 23.4.2021 and it was desk-rejected (i.e. without review) on 3.5.2021 as the "paper does not meet the [. . .] criteria of impact, innovation, and interest". On 4.5.2021 we submitted the paper to PRE Letters and received the reports on 29.5.2021. We resubmitted on 17.6.2021, the paper was accepted on the same day, and published on 29.7.2021.

One of the referees wrote "In particular, the authors should make it clear why the behavior observed has never been observed before." We can only speculate why these properties have never been reported before. (i) It seems like the colleagues who studied the model in the 1990s and early 2000s focused on the fire size distributions and by doing so completely neglected temporal properties. (ii) Previous work studied a different parameter range. The quasi-periodic attractors only become clearly visible for quite large tree-growth probabilities unless the power spectrum is analyzed in detail.

I am content with the work and a bit proud of having made the discoveries. In summary, I think it took curiosity, intuition, luck, and a bit of courage. Overall, I am thankful to have the freedom to be able to investigate something that is beyond any project proposal. I am particularly thankful to have co-workers who supported me and the work at the various stages.

[1] D. Rybski, V. Butsic, and J. W. Kantelhardt, Self-organized multistability in the forest fire model, Phys. Rev. E **104**, L012201 (2021).

[2] S. Clar, B. Drossel, and F. Schwabl, Forest fires and other examples of self-organized criticality, J. Phys.: Condens. Matter **8**, 6803 (1996).

[3] M. E. J. Newman, Power laws, Pareto distributions and Zipf's law, Contemp. Phys. **46**, 323 (2005).

[4] B. Drossel and F. Schwabl, Self-organized critical forest-fire model, Phys. Rev. Lett. **69**, 1629 (1992).

[5] C. L. Henley, Statics of a "self-organized" percolation model, Phys. Rev. Lett. **71**, 2741 (1993).

[6] C.-K. Peng, S. V. Buldyrev, A. L. Goldberger, S. Havlin, F. Sciortino, M. Simons, and H. E. Stanley, Long-range correlations in nucleotide sequences, Nature **356**, 168 (1992).

[7] D. Rybski thanks the Alexander von Humboldt Foundation for financial support under the Feodor Lynen Fellowship.

[8] H. D. Rozenfeld, D. Rybski, J. S. Andrade Jr., M. Batty, H. E. Stanley, and H. A. Makse, Laws of population growth, Proc. Natl. Acad. Sci. U. S. A. **105**, 18702 (2008).

[9] H. D. Rozenfeld, D. Rybski, X. Gabaix, and H. A. Makse, The area and population of cities: New insights from a different perspective on cities, Am. Econ. Rev. **101**, 2205 (2011).

[10] J. W. Kantelhardt, E. Koscielny-Bunde, H. H. A. Rego, S. Havlin, and A. Bunde, Detecting long-range correlations with Detrended Fluctuation Analysis, Phys. A **295**, 441 (2001).

[11] R. M. May, Simple mathematical models with very complicated dynamics, Nature **261**, 459 (1976).

[12] A. Bunde and S. Havlin, Fractals and disordered systems (Springer-Verlag, New York, 1991) Chap. 2, pp. 51–95.