Special Issue Paper

Environmetrics

Received: 30 March 2014, Accepted: 7 April 2014, Published online in Wiley Online Library:

(wileyonlinelibrary.com) DOI: 10.1002/env.2279

An exploratory data analysis of the temperature fluctuations in a spreading fire

David R. Brillinger[a]* and Mark A. Finney[b]

A series of fire experiments were carried out in a wind tunnel at the United States Forest Service's Fire Science Laboratory in Missoula, Montana. The experiments involved tines cut out of pieces of cardboard. The pieces were laid out in comb-like strips parallel to each other along a testbed. They were ignited at the windward end of the testbed. The progress of the fire was monitored by thermocouples, recording temperature, set out equidistantly up the middle of the testbed. Goals of the experiment included improved understanding of wildfire spread and the development of practical tools for wild land fire managers to employ. There was to be a search for regular pulsing in the series and any other interesting phenomena. This paper presents the results of a variety of exploratory data analyses meant to elicit information concerning the series before commencing probability modeling. Published 2014. This article is a U.S. Government work and is in the public domain in the USA.

Keywords: nonstationary; exploratory data analysis; fire progression; signal-generated noise; temperatures; wind tunnel

1. INTRODUCTION

In 2012, a suite of fire experiments were carried out in a wind tunnel at the United States Forest Service's Fire Science Laboratory in Missoula, Montana. The experiments involved tines cut on of pieces of cardboard. The pieces were comb-like strips. They were laid out parallel to each other along a testbed. They were ignited at the windward end of the testbed. The progress of the fire was monitored by thermocouples spread out equidistantly up the middle of the testbed. The data recorded and available for analysis were time series of temperatures.

This paper uses the methods of exploratory data analysis (EDA) on the temperature series to gain insight and understanding. There is an end goal of building an analytic model for flames. The paper begins with some remarks that John Tukey (hereafter referred to as JWT) made through the years concerning EDA, EDA being his creation. He was a longtime contributor to methods for the analysis of scientific data. He wrote the book "Exploratory Data Analysis" (Tukey, 1977).

Following are from that book and other writings. They illustrate what he took the field to be. Jones (1986) included many of his writings on the topic of EDA. Pungent and pertinent JWT quotes include the following:

> WARNING: It will almost never to be wise to fail to use some E_{kr} in a digital computation. (Page 49 in Blackman and Tukey, 1958).
>
> Exploratory Data Analysis is detective work—numerical detective work—or counting detective work—or graphical detective work. (Page 1 in Tukey, 1977)
>
> Exploratory data analysis isolates patterns and features of the data and reveals these forcefully to the analyst. (Page 1 in Hoaglin *et al.*, 1983)
>
> Exploratory data analysis … does not need probability, significance or confidence. (Tukey, 1973a)
>
> The basic general intent of data analysis is … to: seek through a body of data for interesting relationships and information and to exhibit the results in such a wayas to make the recognizanle to the data analyzer and recordable for posterity (Tukey, 1973b)
>
> Graphical presentation appears to be at the very heart of insightful data analysis. (Page 1 in Jones, 1986)
>
> I hope that I have shown that exploratory data analysis is actively incisive rather than passively descriptive, with real emphasis on the discovery of the unexpected. (Tukey and Wilk, 1966)

One notes JWT's skill with words and the challenging spirit of these remarks. There are many more. Brillinger (2010) and Morgenthaler (2009) provided expositions on EDA.

The data from the Missoula experiments have been analyzed previously, as described in Finney *et al.* (2013) and Brittany *et al.* (2013). Beyond the thermocouple data, these researchers used measurements from videos and employed physical models from fluid mechanics. They could then provide estimates of a broad variety of pertinent physical parameters beyond the reach of the data of this paper.

The work here is in part preparatory to building detailed stochastic models that, in particular, take note of the layout of the fuel and its consumption. Also, the data from more experiments and thermocouples will be studied and pertinent stochastic models developed.

* *Correspondence to: D. R. Brillinger, Department of Statistics, University of California, Berkeley, CA 94720-3860, U.S.A. E-mail: brill@stat.berkeley.edu*

a *Department of Statistics, University of California, Berkeley, CA 94720-3860, U.S.A.*

b *USDA Forest Service Fire Sciences Laboratory, Rocky Mountain Research Station, P.O. Box 8089, Missoula, MT 59807, U.S.A.*

The sections of the present article are as follows: 1. Introduction; 2. The Experiments and the Data; 3. Some Time-side Exploratory Data Analyses; 4. Some Frequency-side Exploratory Data Analyses; and 5. Summary and Discussion. References are provided at the end of the paper.

2. THE EXPERIMENTS AND THE DATA

Laboratory experiments with fire were carried out in a wind tunnel at the USDA Forest Service Fire Sciences Laboratory, Missoula, Montana. The researchers there were interested in studying fire spread. It is clear that when considering fire spread, the geometry and amount of available fuel are important. The geometry in the present case is that of fuel tines in combs laid out in a regular grid laid out in parallel lines along a testbed (Figure 1).

The first comb was ignited at the windward end of the testbed in the presence of a wind speed of $0.67\,\mathrm{cm\,s^{-1}}$. The progress of the fire temperature was monitored by thermocouples spread equidistantly up the middle of the testbed. Figure 2 is a photo taken from the side of the testbed. It shows the flames extending over a number of the tines.

Buoyancy and turbulence of the flames are apparent in the figure. An explanation of the apparent turbulent behavior is that cold air is rushing in irregularly underneath flames to replace the hot air rising. There were 34 experiments in the present study. The data of experiment 13 thermocouple 1 provide the principal example for this paper. The data from thermocouple 5 are also employed to estimate the velocity of the advance of the fire later in the paper.

Figure 3 is a graph of the principal data, that is, the temperatures recorded at thermocouple 1. The sampling interval is 1/500 s. There are 25,001 data points in all in this case.

Summary values for these data include the following:

Temperatures recorded (°C)

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
13.98	37.53	73.48	148.40	152.80	1369.00

Successive differences

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
−62.06	−1.034	−0.0755	0.00002	0.37940	102.80

Figure 1. The combs of tines are laid out regularly along a straight testbed

Figure 2. This photo was taken from the side of the testbed. The wind is blowing from the left, and the flames are moving to the right. Some tines are still standing

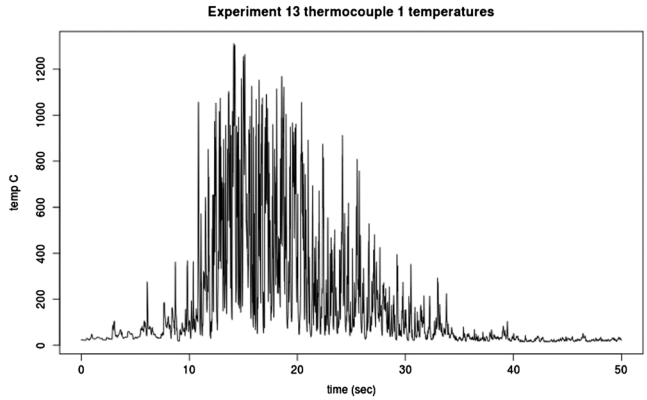

Figure 3. The temperatures (°C) recorded at thermocouple 1. There are 25,001 measurements in all

It is noteworthy to see a value of a change of 102.80 °C in 1/500 of a second. Figure 2 shows dark regions bordering on light. A jump between the two can lead to a large change.

In Figure 3, one sees the temperature starting low and growing. It is then sustained for a period and finally dies off slowly as embers are providing the heat. The dieing off results because the amount of fuel is limited. At the beginning, not many tines are alight but steadily more become alight. Eventually, all the tines have burnt through, and one is left with embers. Finney *et al.* (2013) recognized three stages: pre-ignition, burning, and glowing.

The figure shows turbulent-like variability about an imagined smoothed line. The level of variability starts low and then increases, is sustained, and then dies off slowly.

The experimental testbed in the experiment studied here was horizontal. Other experiments use inclined testbeds and periodic pulsing is sometimes observed. Results for inclined testbeds are considered in Malalsekera *et al.* (1996); Atkinson *et al.* (1995); Dupuy *et al.* (2011), and Malalsekera *et al.* (1996).

The main EDA method employed in this section was visualization.

3. SOME TIME-SIDE EDA ANALYSES

3.1. Time-side EDA analyses

It is of interest to estimate a smooth central trend line for the temperature plot of Figure 3. One way to learn about such a line and to notice special features is to prepare parallel boxplots. To do this, the data set of 25,000 data points was split into 20 contiguous time segments of 1250 points. Boxplots were computed for each segment and are set up in parallel in Figure 4.

In Figure 4, one sees the overall shape of Figure 3. A method to obtain a trend line would be to link together the middle lines in the boxes. These are the medians of the data of the segments. Overall, there is a quick rise followed by a slow falloff. The vertical dashed lines are called whiskers. The points falling outside are called outliers. There are many of these extra-large values. The distances between the tops and bottoms of the boxes provide the inter-quartile ranges (IQRs)—a measure of spread. One sees that the size of the spread mimics the medians. There is a form of signal-generated noise. It might be associated with turbulence.

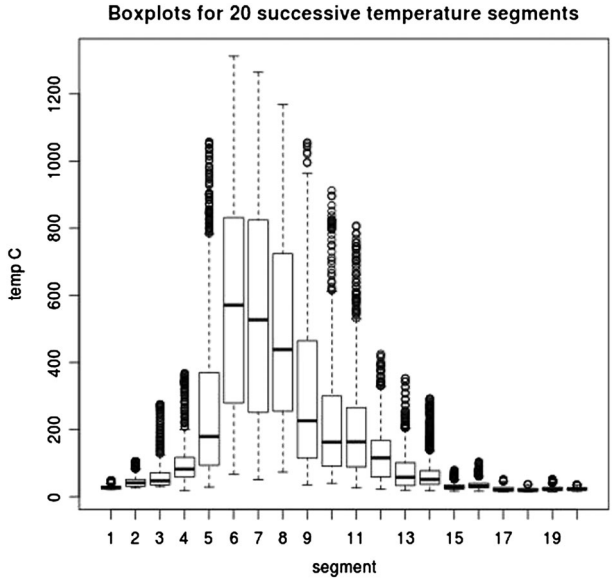

Figure 4. The plot shows parallel boxplots for 20 successive segments of the data of Figure 3. There are many outliers indicated by the open circles in the figure

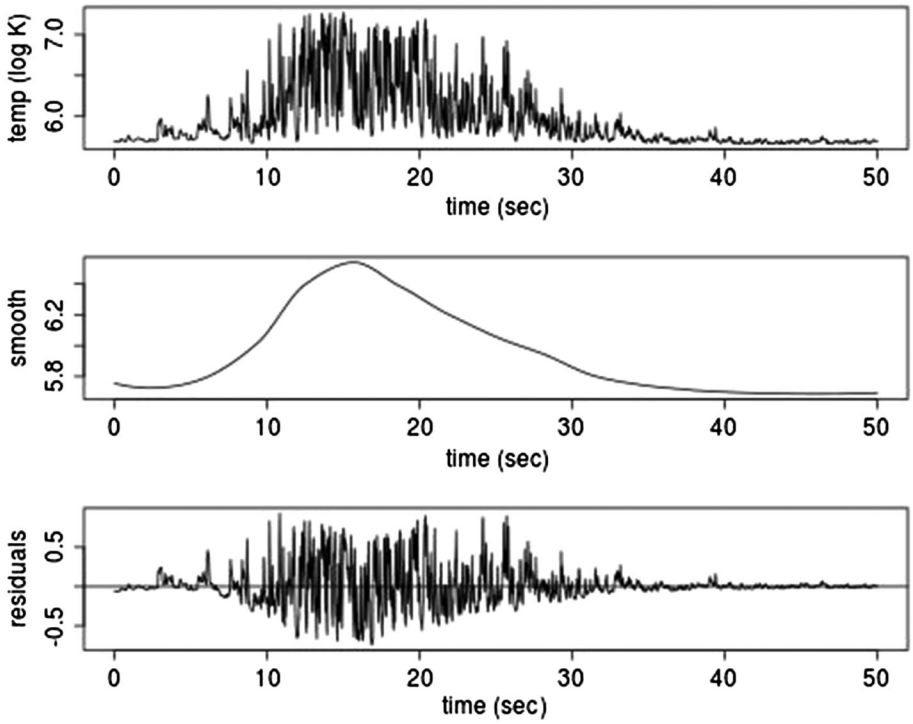

Figure 5. The series shown are, respectively, the data, the smoothed values obtained by loess, and the difference between these two, the residuals. The data are Kelvin values

Re-expression or transformation of a measurement is a common tool of EDA. The suggestion of signal-generated noise leads to consideration of working with the logarithm of Kelvin temperatures. This is done henceforth. The top panel of Figure 5 shows the variability changing as the smooth does. The bottom panel is the residuals (or rough) given by the additive EDA decomposition (Tukey and Wilk, 1966):

$$\text{observation} = \text{fit} + \text{residual}$$

In a more formal approach to obtaining a smooth central curve, one can employ Cleveland's "loess" smoothing procedure (Venables and Ripley, 2003). When it is applied to the temperature series, one obtains the smooth curve of the central panel of Figure 5. It shows a curve with a rapid rise and a slow falloff. The loess tool involves a sliding weighted regression analysis relating temperature to time. Figure 5 shows the three components: data, fit, and residual. Each may be considered in some detail.

The use of log Kelvin values has led to a residual series more symmetric about 0. The residuals also appear to make local behavior clearer. It now seems pertinent to compute a smoothed version of the spreads. This is the middle curve of Figure 6. The computations involved applying loess to the absolute values of the residuals. (One might have used the IQRs.) This leads to the middle curve of Figure 6. Dividing the individual values of this curve by the corresponding values of the middle panel leads to the bottom panel of the figure. This procedure may be viewed as a form of prewhitening. Considering the final panel of Figure 6 suggests that it has been reasonably effective as hoped. The step is based on the following paradigm (Tukey and Wilk, 1966):

$$\text{observation} = \text{fit} \times \text{residual factor}$$

These quantities, the residual factor in the aforementioned equation, will be called scaled residuals. They are seen to show little trend. They are considerably more time-homogeneous than the residuals of Figure 5 but do remain auto-correlated. Homogeneous residuals can obviate the need for using weighted least squares. They can be used to investigate symmetry, skewness, kurtosis, and other related qualities. They can be fed into other EDA analyses such as normal probability plotting and density estimation to examine departures from standard distributions. It is impressive how similar the middle curves of Figures 5 and 6 are.

In the left-hand panel of Figure 7, one sees evidence for substantial departure of distribution of the scaled residuals from a normal curve. The specific character is clearer from the right-hand plot. It indicates that the residuals are strongly skewed to the right.

Turning to the case of several thermocouples, it is to be noted that experiment 13 actually involved measurements at 64 thermocouples. Next data for thermocouples 1 and 5 will be used as an illustrative example. They are plotted in Figure 8.

The two series are very similar. This is not surprising because thermocouple 5 is downwind from number 1 and close to it. These two series provide an opportunity to estimate the velocity with which the heat energy is moving down the testbed. This will be done next.

It is important to estimate the velocity with which the heat front is advancing, for then one can interchange time with location, albeit approximately. In a paper that considers the case of location x on the surface of a sphere, Brillinger (1997) provided a review of various methods to estimate velocity of motion given data $\{Y(x,t), t=0,\ldots,T-1, x=0,\ldots,X-1\}$. Projection pursuit regression (PPR) is one method. It is discussed and investigated in Friedman and Tukey (1974), Huber (1985), Xia and An (1999), Hastie *et al.* (2001), and Venables and Ripley (2003), and it is used next.

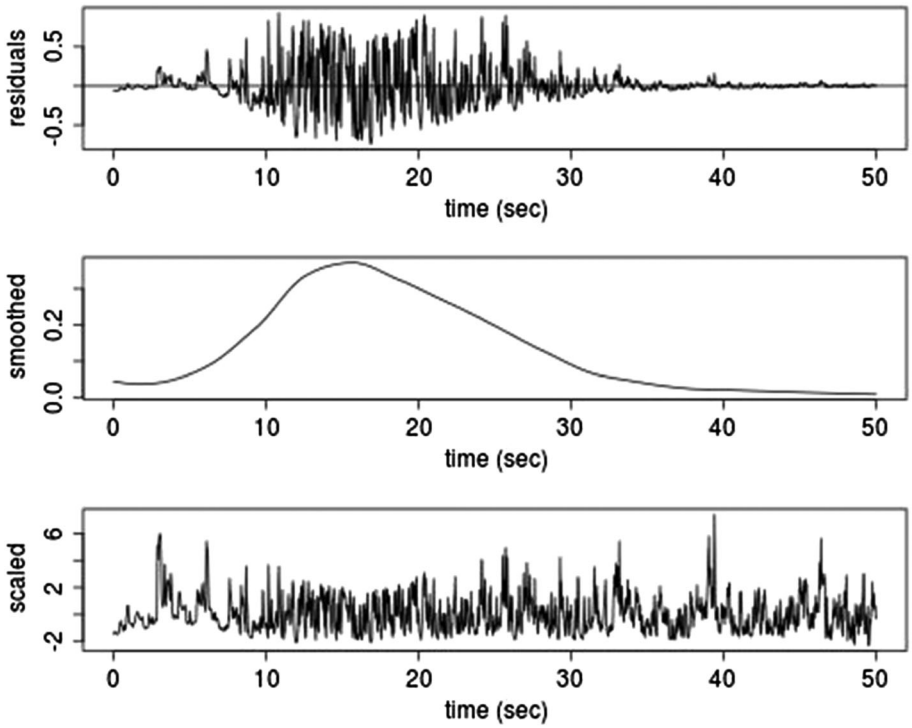

Figure 6. The panels are, respectively, the residuals of the last panel of Figure 5, the smooth of the absolute residuals, and the result of dividing these latter into the corresponding values of the top plot

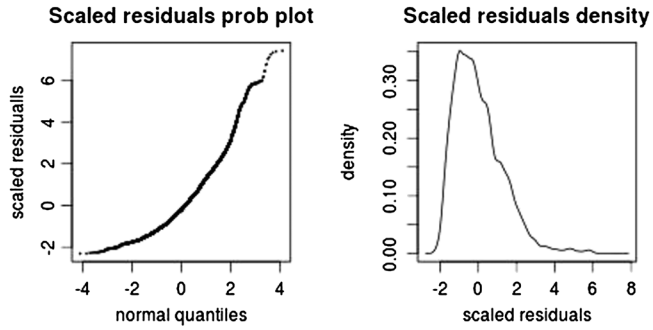

Figure 7. The left-hand plot here provides a normal quantile plot, and the right hand, a density estimate of the scaled residuals

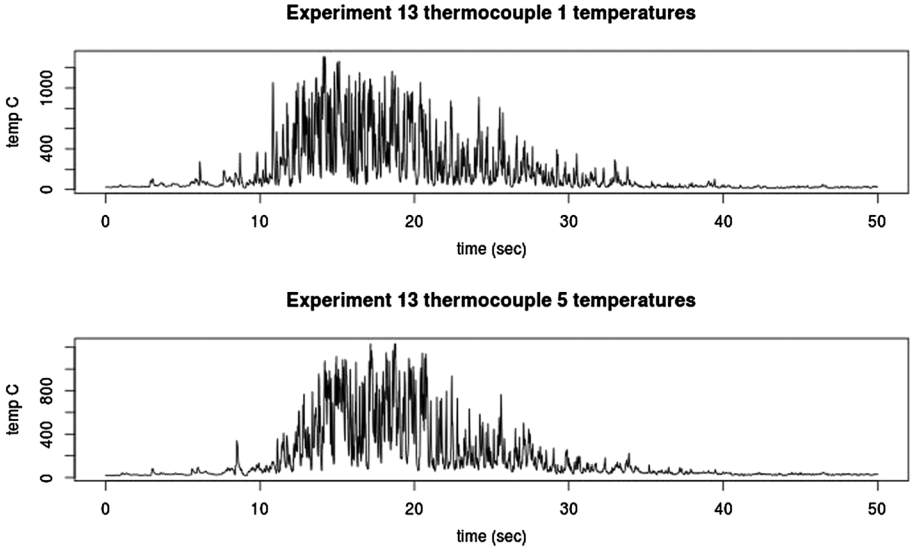

Figure 8. The temperatures recorded by thermocouples 1 and 5, respectively. The wind comes from the left. The lower panel looks like the top panel shifted somewhat to the right

Movement can be described by wave equations. The wave equation on the line has solutions $g(x - vt)$ with t time, x location, v velocity, and $g(.)$ the wave shape. In the present case, let $y_i(t)$ denote temperature series recorded at the ith thermocouple, which is located at x_i. Assuming that the presence of the smooth function $g(.)$, one can suppose that

$$y_i(t) \sim g(x_i - vt) \quad i = 1, \ldots, 64$$

With appropriate definitions, this last is the simplest form of a PPR model. The model was fit for $i = 1, 5$ using the procedure ppr() of the library MASS of the statistical package R (Venables and Ripley, 2003). This tool fits $g(x - vt)$ to (x,t) by least squares with the data to be $y(x_i,t_{ij})$ and (x_i,t_{ij}), $i = 1, 5$ and $j = 1,\ldots,25{,}001$. The results follow.

Using ppr(.), the estimate of v was found to be $2.188\,\mathrm{cm\,s^{-1}}$. The estimate of the function $g(.)$ is provided in Figure 9. One sees that it follows the general shape of the central function of series 1 and 5 seen in Figure 8. It is to be noted that the shape, g, will undoubtedly be changing somewhat as the heat energy moves along the testbed because of local variability.

Now that an estimate of the velocity, $2.188\,\mathrm{cm\,s^{-1}}$, is available, the time values may be replaced by distance values. Thereby, one has a snapshot from the side of the testbed of the values $y(x)$ with $x = vt$. One is able to convert time series into a spatial series and then a spatial–temporal series, $y(x,t)$. The situation now looks more like the common one studied in turbulence. The spatial snapshot is graphed in Figure 10. It is to be remembered that the velocity used is an estimate.

The result is that the graph is no longer moving but varies in space. One could now estimate spatial correlations.

To end this presentation of time-side analyses, it can be noted that it is intended to make later use of the locations of tines on a rectangular grid, as in Figure 1. The flames start at the beginning of the fuel and then move towards and through the thermocouples. Taking individual tines as heat sources, the heat arriving at a thermocouple will start low and then increase and finally die as the fuel is exhausted. The distances of the tines to the thermocouple are computable. The heat at a thermocouple will be a sum of the heats arriving from various tines. It is anticipated that smooth curves like that of the middle one of Figure 5 will be produced.

3.2. Point process methods

The possibility of regular pulses in the data has been mentioned. To begin to study this possibility, attention turns to detecting and recording locations of pulses in the original data. The following is an algorithm for finding pulses. Consider 7-tuples $(y_1, y_2, y_3, y_4, y_5, y_6, y_7)$ of adjacent values. Pulses might be located by looking for 7-tuples $(y_1, y_2, y_3, y_4, y_5, y_6, y_7)$ with $y_1 < y_2 < y_3 < y_4$ and $y_4 > y_5 > y_6 > y_7$. When one is found, the location of a pulse is taken to be y_4. For the series of the bottom panel of Figure 5, 522 pulses were detected amongst 25,001 series values. (Of course, longer tuples could be considered.)

The locations of these pulses may be viewed as part of a realization of a point process. Point process analysis often includes graphing the step function of the cumulative counts of the number of points as a function of time. Such a plot is provided in Figure 11.

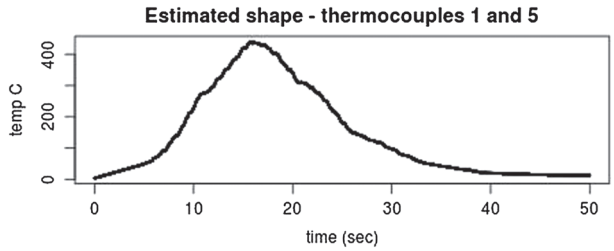

Figure 9. The projection pursuit estimate of the smooth function $g(.)$

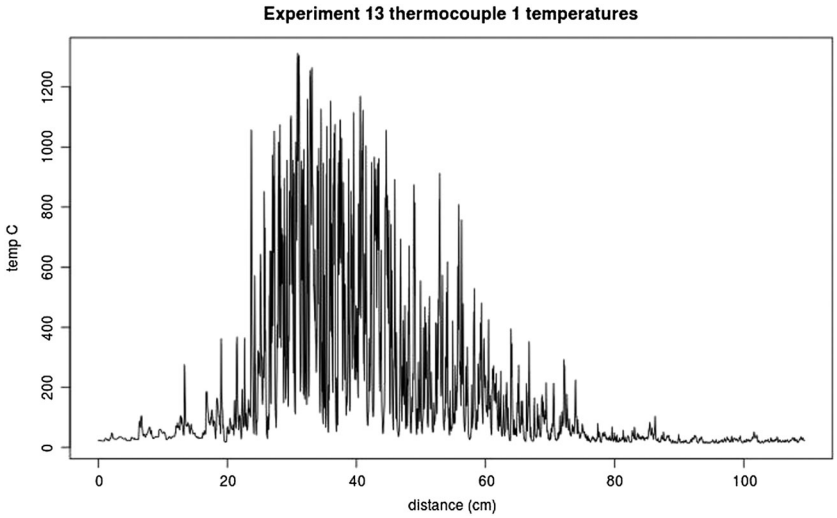

Figure 10. The figure provides a plot of the temperatures versus distance along the testbed

The curve in Figure 11 is a step function with 522 jumps of size 1. The slope of the curve at a given time point gives the local rate of occurrence of points/pulses. Here, the rate of pulses is seen to be slower for the first 10 s and then increasing until slowing down around time 40 s. From the graph, one can see that there are about 50 pulses in the first 10 s.

If there were sufficient pulses the same distance apart, then they might show up if one looked at the successive differences between the times of the estimated pulses. A histogram type plot is provided by a stem and leaf display shown in Figure 12. It is typically more useful than a histogram in that it displays more detail concerning the data up to a round off error. The smallest value is 9, and the largest is 499 in units of a 500^{th} of a second. No pulsing stands out.

There is a latent marked point process here. Denote the time series values by $\{y(t)\}$. Denote the pulse locations by $\{\sigma_j\}$. Take marks to be $\{y(\sigma_j)\}$. The marked point process is the collection of values $\{\sigma_j, y(\sigma_j)\}$.

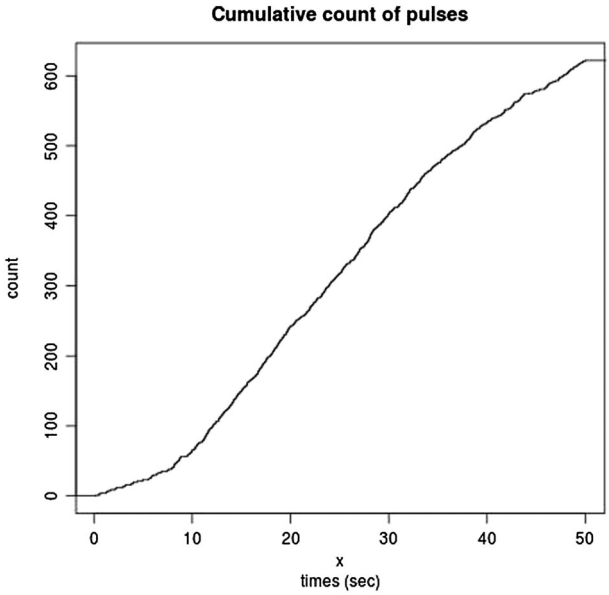

Figure 11. The cumulative count of the 522 pulses as a function of time is shown

```
The decimal point is at the |

 0 | 969
 2 | 356990146
 4 | 3390014456789
 6 | 00114663345566789
 8 | 012456781122235668899
10 | 00023455678112345567789
12 | 00111233444678899001223445567888
14 | 00112444556788001123345568 9
16 | 01112334556679990111234456679999
18 | 00112233455567891222335666 79
20 | 00144556678902244567 8899
22 | 00123346678800112234445667 7899
24 | 00112235567899123344456 78
26 | 01134556678890000123455667 7999
28 | 01123344456678899033445578 99
30 | 00233456789000112345677899
32 | 02234566678890112336678 9
34 | 022334456789001335789
36 | 011245699123388899
38 | 023455667778800233456699
40 | 1122345589911235667
42 | 0123355557013346789
44 | 11235890034557899
46 | 011123456880134567799
48 | 2344567788990135667899
```

Figure 12. Stem and leaf display of the time differences between successive pulse locations

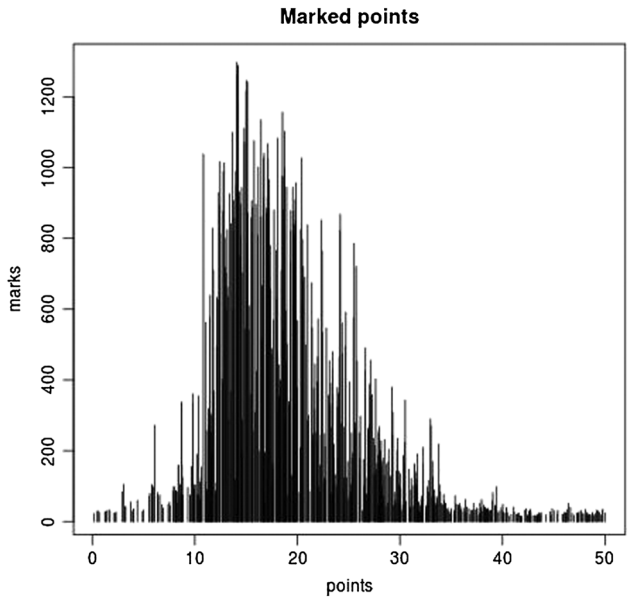

Figure 13. The points $\{\sigma_j, y(\sigma_j)\}$ $j = 1,\ldots,522$ are shown with vertical lines drawn from $\{\sigma_j, 0\}$ to $\{\sigma_j, y(\sigma_j)\}$

The graph of Figure 13 shows vertical lines connecting the $\{\sigma_j, 0\}$ to the $\{\sigma_j, y(\sigma_j)\}$. The figure has much the same character as Figure 3. In a sense, it is the skeleton of Figure 3. The figure provides much the same information as Figure 3 but requires only 522 points instead of 25,001. One could reduce the number further by making the definition of a pulse more restrictive. Here, it involved a 7-tuple.

Many approaches and powerful methods are available for stochastic marked point processes, for example, the complete intensity function. These will be considered in Section 4 where formal stochastic models for series like the one of Figure 3 are considered.

In this section, analyses have been carried out decomposing the temperature series of Figure 3 into a smooth part and residuals. Next, a method was developed for prewhitening the series, and then locations of pulses in the series were estimated. In the end, a latent marked point process was constructed, the velocity with which the energy is moving is estimated, together with an estimated moving shape. Lastly, a snapshot of a stationary segment of the series was prepared.

4. FREQUENCY-SIDE EDA ANALYSES

Frequency-side analysis that is making extensive use of sine and cosine functions in the work has been basic to time series analysis and signal processing for many years. It has led to important scientific discoveries and new engineering techniques. It has been used to look for periodicities, fractals, and scaling (self-similar) laws. A common empirical quantity that has been employed in these studies is the empirical Fourier transform,

$$d^n(\lambda) = \sum_{t=1}^{n} u(t) \exp\{-i\lambda t\}$$

given the values $\{u(1),\ldots,u(n)\}$with $i = \sqrt{(-1)}$ and λ real-valued. Often, the λ's are of the form $2\pi j/n$, for $j = 1,\ldots,n$.

Taking $\{u(t)\}$ to be the scaled noise values graphed in the top panel of Figure 14, the second panel displays the absolute value $|d^n(\lambda)|$ on a log scale.

The graph is on a log scale which often stabilizes the variability of such spectra. One notices a steady decay of $|d^n(\lambda)|$ as the frequency increases; that is, there remains autocorrelation. One also notes two "birdies" standing up near 200 Hz. This is thought to be electronic noise from the thermocouples.

Given the periodic pulsing noticed in fires moving up a plane, see Atkinson *et al.* (1995), Brittany *et al.* (2013), and Finney *et al.* (2013). One looks for prominent peaks at lower frequencies, for example, in the third panel of Figure 14. None stands out. There is incomplete removal of the trend apparent in Figure 3 and only partially removed by rescaling.

The quantity

$$|d^n(\lambda)|^2$$

is called the periodogram. The quantity.

$$d^n(\lambda)\, d^n(\mu)\, \mathrm{conj}(d^n(\lambda + \mu))$$

is called the biperiodogram. The latter can prove useful when nonlinearities or skewed values are involved (Lii *et al.*, 1976). The next figure shows the log (absolute value) of a smoothed biperiodogram. It was smoothed by splitting the original series up into 20 contiguous disjoint segments, and evaluating the biperiodogram for each and averaging those together.

One sees a rapid decay as one moves away from (0,0). The decay is clearer in the following perspective plot of Figure 16.

Figures 15 and 16 suggest the possibility of employing a linear process model. This will be considered in later work.

To end the examples, attention returns to the nonconstant central level in Figure 3. It has been a steady concern in much of this paper. The spectrogram provides an alternate way of dealing with this. The spectrogram is computed by evaluating periodograms, $|d^n(\lambda)|^2$, for segments

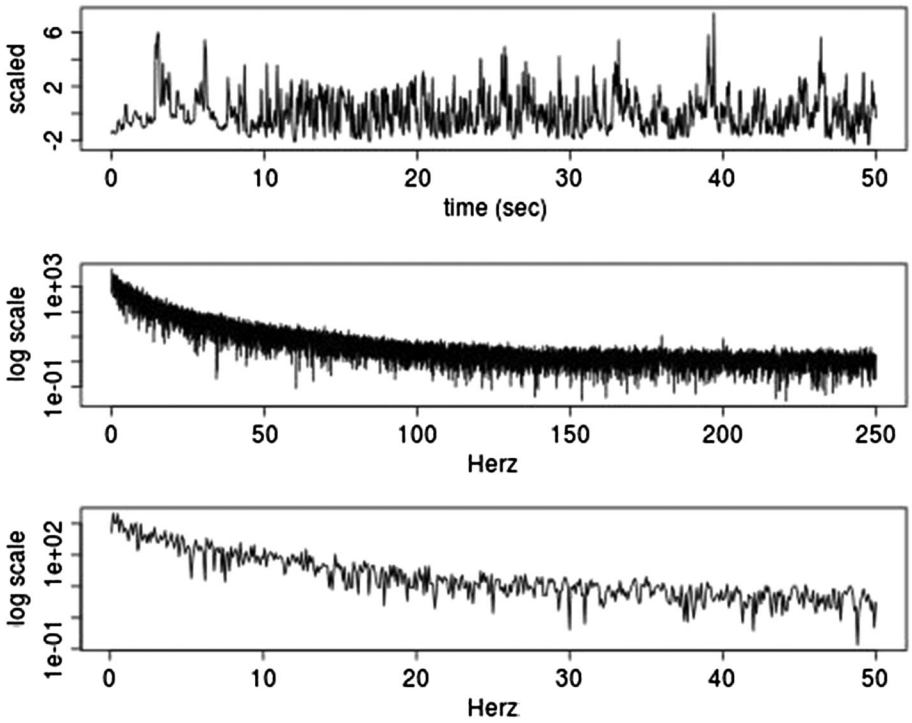

Figure 14. The top plot panel reproduces the rescaled residuals of Figure 6. The middle shows the absolute value of the Fourier transform of these data. The bottom shows the absolute value of the Fourier transform for the first 10 s of the data

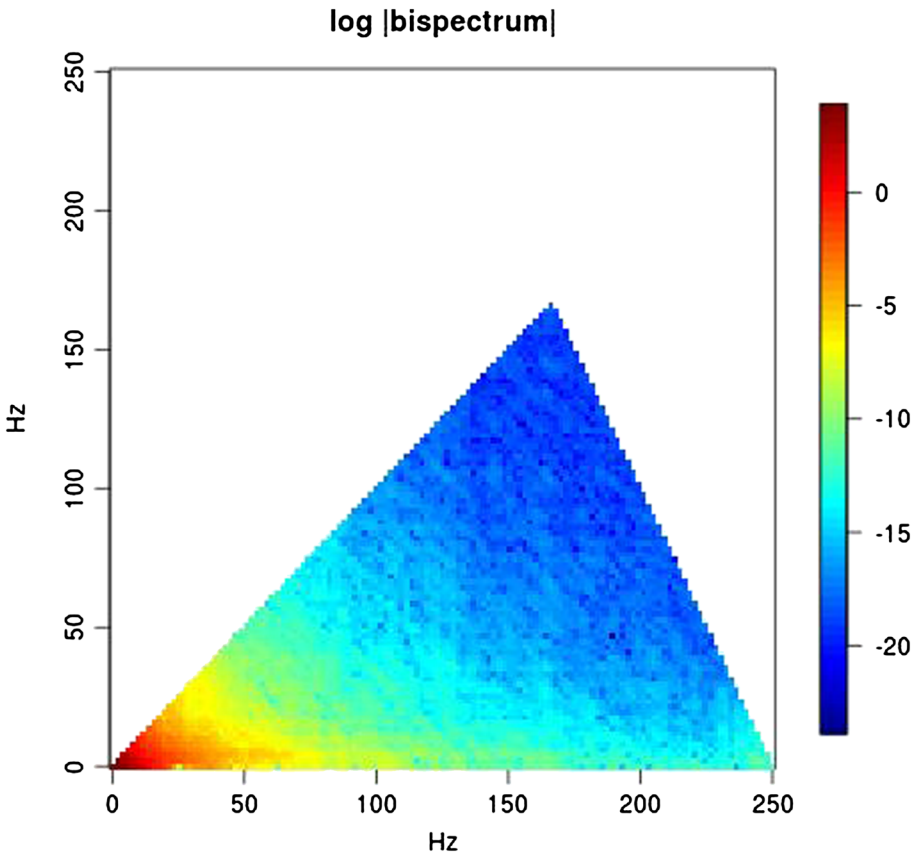

Figure 15. An image plot of the smoothed biperiodogram

sliding along the time axis. It can be effective in highlighting latent frequencies in a nonstationary case. In that case, horizontal ridges appear; see the examples in Brillinger (2012).

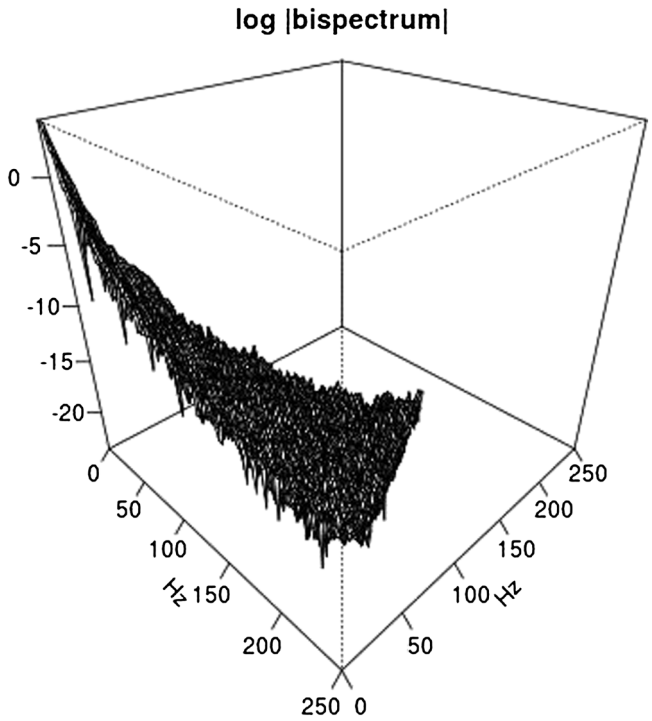

Figure 16. Perspective plot of Figure 15

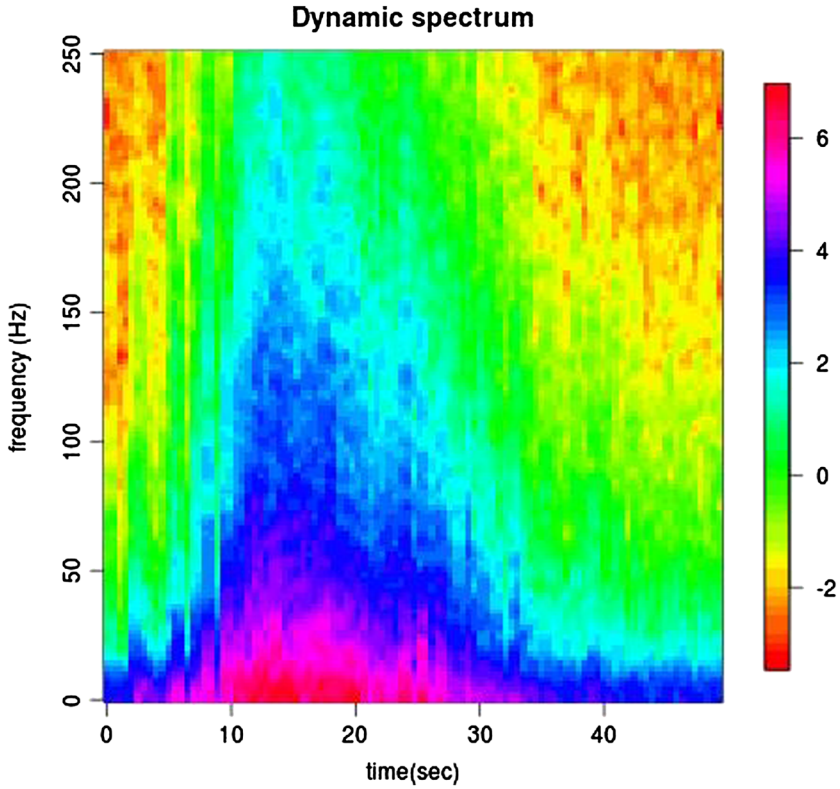

Figure 17. Spectrogram of the time series of Figure 3

Figure 17 provides an image plot as a function of time and frequency. One sees the nonstationarity, the rising and falling of energy, over many frequency bands.

In this section, several frequency-side analyses have been presented. They were meant to search out regular pulsing. The final graph of Figure 14 suggests that if it exists, the amplitude is small as also suggested by Figure 9.

5. DISCUSSION AND SUMMARY

5.1. Summary

There was an investigation of temperature series collected in a wind tunnel experiment on fire progression. The methods of EDA were employed because they are practical tools able to provide a structure to an investigation, to summarize, and to elicit unexpected phenomena. There were procedures for each of these in the R statistical package (Ihaka and Gentleman, 1996).

The EDA methods included the following: graphs/displays, stem-and-leafs, (parallel) boxplots, probability plots, stepfun, re-expression, (5-number) summary, Fast Fourier Transform, spectrogram, smoothing, residuals, and projection pursuit. There were procedures for each of these in the R statistical package (Ihaka and Gentleman, 1996).

5.2. Discussion

The typical statistician has learned from bitter experience that negative results are just as important as positive ones, sometimes more so. (Tukey, 1967)

The negative result in the present work was that EDA methods did not find evidence for regular pulsing in this case of a horizontal test bed. What that work did though was to lead to a focusing on pulses generally. A marked point process was developed. This will perhaps be useful in future work. Another discovery was the presence of signal-generated noise. The noise here appears turbulent. The periodogram and biperiodogram were computed and found to die off steadily as happens with turbulence in other situations. The birdies of Figure 14 were a minor find. There is a need to analyze more of the data sets to confirm results "found." There is a need to run the series for other thermocouples and other experiments. One wonders if the results here may help in developing procedures to simulate similar temperature series.

A frequency domain procedure for fitting signal-generated noise is developed in Ihaka (1993) for the case of a parametric signal.

It can be said that the focus on EDA methods has provided a logical plan of work and suggested items for future work.

Acknowledgements

DRB's research was supported by the NSF Grant DMS-100755. It was further supported in part by the Banff International Research Station (BIRS) with an invitation to speak at the workshop "Managing Fire on Populated Forest Landscapes," 20−25 Oct 2013 in Banff, Alberta.

There are thanks to H. K. Preisler for arranging and editing this collaborative work. We further thank Ryan Lovett, Chris Paciorek, and Ray Spence of the Statistical Computing Facility whose computer systems were made use of and mention the help from the US Forest Service, National Fire Decision Support Center.

REFERENCES

Atkinson GT, Drysdale DD, Yu Y. 1995 Fire driven flow in an inclined testbed. Fire Safety Journal **25**: 141–158.

Blackman RB, Tukey JW. 1958. The Measurement of Power Spectra. Dover: New York.

Brillinger DR. 1997. An application of statistics to meteorology: estimation of motion. In Festschrift for Lucien Le Cam. Springer: New York; 93–105.

Brillinger DR. 2010. Exploratory data analysis. In International Encyclopedia of Political Science. Sage: New York; 530–537.

Brillinger DR. 2012. The Nicholson blowfly experiments: some history and EDA. Journal of Time Series Analysis **33**: 718–723.

Brittany A, Adam N, Akafuah K, Finney M, Forthofer J, Saito K. 2013. A study of flame spread in engineered cardboard fuelbeds, Part II: scaling law approach. *Seventh International Symposium on Scale Modeling (ISSM-7)* Hirosaki, Japan.

Dupuy JI, Marech J, Portier D, Valette J-C. 2011. The effects of slope and fuel bed width on laboratory fire behavior. International Wildlife Journal **20**: 272–288.

Finney MA, Furthofer J, Grenfell IC, Adam BA, Akafuah NK, Saito K. 2013. A flame spread in engineered cardboard fuelbeds, Part I: correlations and observations of flame spread. *Seventh International Symposium on Scale Modeling (ISSM-7)* Hirosaki, Japan.

Friedman JH, Tukey JW. 1974. A projection pursuit algorithm for exploratory data analysis. IEEE Transactions on Computers **C-23**(9): 881–890.

Hastie T, Tibshirani R, Friedman JH. 2001 The Elements of Statistical Learning. Springer: New York.

Hoaglin DC, Mosteller F, Tukey JW. 1983 Understanding Robust and Exploratory Data Analysis. Wiley: New York.

Huber P. 1985. Projection pursuit. Annals of Statistics **13**: 435–476.

Ihaka R. 1993. Statistical aspects of earthquake source parameter estimation in the presence of signal generated noise. Communications in Statistics - Theory and Methods **22**: 1425–1440.

Ihaka R, Gentleman R. 1996. R: a language for data analysis and graphics. Journal of Computational and Graphical Statistics **5**: 299–314.

Jones LD. editor. 1986. Philosophy and Principals of Data Analysis 1965–1986. The Collected Works of John W. Tukey, Volume IV. Wadsworth & Brooks/Cole: Monterey.

Lii KS, Rosenblatt M, Van Atta C. 1976. Bispectral measurements in turbulence. Journal of Fluid Mechanics **77**: 46–62.

Malalsekera WMG, Versteeg HK, Gilchriste K. 1996. A review of research and an experimental study on the pulsation of buoyant diffusion flames and pool fires. Fire and Materials **20**: 261–271.

Morgenthaler S. 2009. Exploratory data analysis. Wiley Online Library. WIRES Computational Statistics DOI: 10.1002/wics.2

Tukey JW. 1967. A statistician's comment. Electronic Handling of Information: Testing and Evaluation, A Kent, OE Taulbee, J Belzer, GD Goldstein (eds), Thompson Book Co: Washington DC; 41–47.

Tukey JW. 1973a. Lags in statistical technology. In *Proceedings First Canadian Conf. on Applied Statistics '71 Canada,* 96–104.

Tukey JW. 1973b. Exploratory data analysis as part of a larger whole. In Proc. 18th Conf. on Design of Experiment in Army Research Development and Training. U.S. Army Research Office: Durham; 1–10.

Tukey JW. 1977. Exploratory Data Analysis. Addison-Wesley: Reading.

Tukey JW, Wilk MB. 1966. Data analysis and statistics: techniques and approaches. Proceedings of the Symposium on Information Processing in Sight Sensory Systems, PW Nye (ed.), California Institute of Technology: Pasadena; 7–27.

Venables WN, Ripley BD. 2003. Modern Applied Statistics with S. Springer: New York.

Xia X, An HZ. 1999 Projection pursuit autoregression in time series. Journal of Time Series Analysis **20**: 693–714.

