We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



6

Statistical Techniques for Exploring Possibly Increasing Trend of Hurricane Activity

Jesper Rydén
Uppsala University
Sweden

1. Introduction

Discussions on climate change are intense; the following quotation, by G.H. Liljequist (1949), is valid also today:

> The recent climatic fluctuation, which seems to embrace the whole earth, has been the object of many discussions and investigations. This is quite natural: its consequences have been far-reaching and have come to influence the everyday life of man in many ways.

Broadly speaking, one aspect is whether certain events, for instance storms or hurricanes, have been more common recently; another if such events (possibly in addition) have become more severe and violent. A difficult problem related to the latter issue (not discussed in this chapter) is the relation to climate change: are changes to occur statistically speaking, as a part of the inherent variability of nature, or are these an effect from human activity.
In this chapter, we present some statistical techniques for investigation of possible trends in time series, as applied to hurricane data: the yearly number of hurricanes as well as the strengths of hurricanes.

2. Background

Focussing on hurricanes in the North Atlantic, the season 2005 resulted in many broken records: 27 named storms, 15 hurricanes, 7 major hurricanes, 4 hurricanes of Category 5. Moreover, the hurricane Katrina implied very large and severe losses. Naturally, questions arise on possible trends of hurricane activity and this had been discussed even earlier. A comprehensive statistical analysis of hurricanes in the North Atlantic was given by Elsner and Kara (1999), and these authors state that "... caution is advised in interpreting the trends." Some researchers agree that a change in intensity has occurred in recent years, see for instance Webster et al. (2005), but careful definitions have to be made. When investigating a possible trend over time for the yearly number of hurricanes, the statistical analyses are often based on regression models where the number of hurricanes is supposed to be a random variable dependent on covariates like North Atlantic oscillation index (NAO) (see e.g. Elsner, Jagger and Niu 2000). Landsea et al. (1999) found that only weak linear trends can be ascribed to hurricane activity and that multi-decadal variability is more characteristic of the region.

The statistical tools in the literature when analysing possible trends of the total yearly number of hurricanes typically are based on some form of generalized linear model, out of which Poisson regression is a special case (clearly, such models are natural, since one wants to model counts of a quantity). As an example, see Solow (1989). Also, more recent statistical techniques have been employed, e.g. Bayesian methodology in a Markov Chain Monte Carlo approach was used by Elsner and Jagger (2004).

The study of damage caused by hurricanes constitute another aspect of hurricane research, where statistical methodology has been used to a large extent. For instance, Katz (2002) introduced a compound Poisson process with two components: one governing the occurrences of events, the other specifying damages associated with the individual events. Jagger et al. (2008) gave a review and presented a strategy for simulating annual insured losses based on pre-season values of NAO and Atlantic sea-surface temperature. The extreme losses were modelled with a generalized Pareto distribution. Further, hurricanes may cause severe power outages. A risk analysis was performed by Han et al. (2009), making use of a generalized additive model in the statistical modelling.

In this paper, we take a pragmatic approach and focus on the question of increasing trend of the yearly number of hurricanes by modelling only this quantity of interest, hence no covariates involved. The purpose of the paper (see also Rydén, 2010) is to point out what can be deduced by an exploratory graphical method with origin in scale-space theory and compare with more traditional statistical methods, and apply it in a case study of data sets related to hurricane activity. This method is known as the SiZer (Significance of Zero crossings of the derivative) map, which can be used to find significant features in data (Chaudhuri and Marron 1999). In environmental data, such features typically depend on the level of detail for which the time series is considered. As a benchmark, we compare with what a conventional method like Poisson regression implies; in that case, a statistical hypothesis test only tells whether to reject or not the hypothesis of a constant trend. The advantage of the SiZer approach is the possibility to investigate several time scales at the same time, presented in one single image. An application to climatology published recently was given by Weckström et al. (2006) and a review of the SiZer and related techniques was given by Godtliebsen, Olsen and Winther (2003).

The paper is organised as follows: Next, in Section 3, the data set of hurricanes is briefly presented. Further, a data series of an index measuring the strength of hurricanes (ACE index) is introduced. Thereafter, in Section 4, the SiZer methodology is reviewed and applied to these two data sets. In Section 5, a comparison is made with other methods, for instance a standard approach based on Poisson regression. Finally, some conclusions are given.

3. Data sets

The region of the Atlantic basin has been studied in this paper. Data are courtesy of Tropical Prediction Center, NOAA, and are conveniently presented at a homepage of Unisys Weather, http://weather.unisys.com/hurricane/atlantic/. In the sequel hurricanes from 1950–2009 are studied. Seasons officially run from 1 June to 30 November. Hurricane activity has been tabulated back to 1871, but normally the portion dating from about 1944 is considered reliable. Use of aircraft monitoring was then introduced. Before the advent of satellites and aircraft monitoring it was difficult to detect storms that did not affect land or ships. This should be kept in mind when analysing older data: the activity in some seasons before the middle of the 20th century might be under-estimated.

3.1 Hurricanes and categories

Hurricanes can be divided into 5 categories according to the Saffir–Simpson scale which is a 1-5 rating based on the present intensity of the hurricane. The scale is used to give an estimate of the potential property damage and flooding expected along the coast from a hurricane landfall. Wind speed is the determining factor in the scale. Major hurricanes are usually defined as those in Category 3 or higher. For reference, the definitions of categories are given in Table 1.

Category	1	2	3	4	5
Wind speed (km/h)	119-153	154-177	178-209	210-249	> 249

Table 1. Saffir–Simpson scale defining categories of hurricanes.

In this paper, the total yearly numbers of hurricanes in categories 1–5 are investigated. In Fig. (1), top panel, a times series is shown.

3.2 ACE index

The National Oceanic and Atmospheric Administration (NOAA) uses an Accumulated Cyclone Energy (ACE) index as a measure of total seasonal activity. This refers to the collective intensity and duration of named storms and hurricanes in the North Atlantic during a given season. The ACE index is defined as the sum of the squares of the maximum sustained surface wind speed measured every six hours for all named systems while they are at least tropical storm strength. The ACE for the 60 seasons 1950–2009 was investigated.
In Fig. (1), bottom panel, a times series of ACE is displayed. Moreover, in Fig. (2), a scatter plot is shown. Obviously a strong correlation is to be expected between total yearly number of hurricanes and ACE, and is also seen from the figure. However, the plot is interesting as also other factors influence the ACE; for instance, the number of named storms.

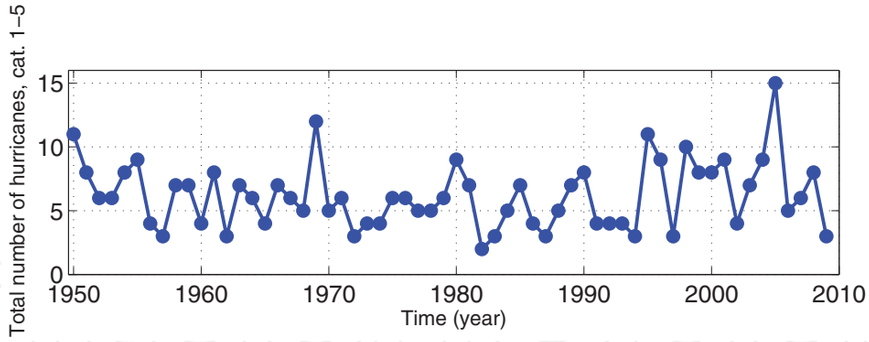
4. Analysis by SiZer

4.1 Statistical background

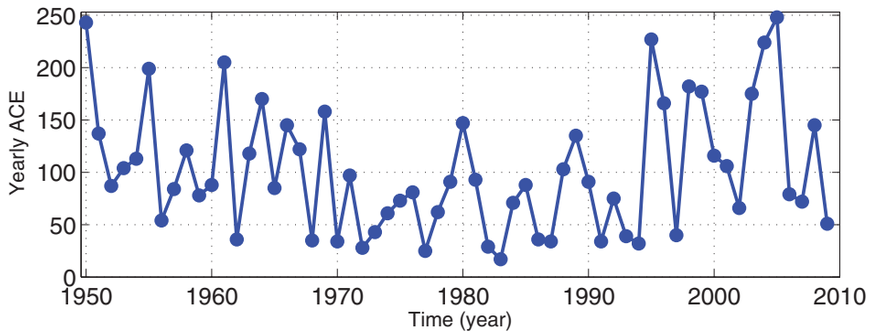
The SiZer technique can be considered a smoothing method. A central question to be posed when facing data is which features are significant, as opposed to sampling artifacts. We here give a brief review; see the original papers by Chaudhuri and Marron (1999), (2000) for full details.
In the simplest SiZer methodology, a random sample $(x_1, y_1), \ldots, (x_n, y_n)$ of independent observations is assumed (this assumption will be reasonable for our datasets under consideration, as will be discussed in the sequel). In a non-parametric regression framework, the data are smoothed to give a curve that can be regarded an estimated conditional mean $f(x) = \mathsf{E}[Y|X = x]$. For the data, the conditional expectation is estimated by

$$\hat{f}_h(x) = \underset{a}{\operatorname{argmin}} \sum_{i=1}^{n} [y_i - (a + b(x_i - x))]^2 K_h(x - x_i)$$

where h is a smoothing parameter called bandwidth and K_h results from a kernel function K, $K_h(\cdot) = (1/h)K(\cdot/h)$. A common choice of K is a Gaussian kernel. The estimate $\hat{f}_h(x)$ is wiggly when h is small and very smoothed when h is large.
The visual assessment of significance of peaks and valleys in a family of smooths is based on confidence limits for the derivative $f'_h(x)$ in scale space. In the SiZer colour map, blue portions



(a) Yearly hurricanes, 1950-2009.



(b) ACE for hurricanes, 1950-2009.

Fig. 1. Time series: yearly hurricanes and ACE.

indicate locations of x and h where $\hat{f}'_h(x)$ is significantly positive, red portions locations where $\hat{f}'_h(x)$ is significantly negative. Furthermore, confidence limits for $f'_h(x)$ are computed (for details, see Chaudhuri and Marron 1999).

4.2 Check of assumptions

SiZer methodology handles independent observations but in time series, dependence is often present. We here study empirically if the assumption seems reasonable. Autocorrelation functions for the two data sets are shown in Figure 3.
The 95% confidence limits, marked as dotted lines, follow from approximative normality and are given by $-1/(n-1) \pm 1.96/\sqrt{n}$. None of the serial correlations fall outside the limits so as a reasonable approximation, we assume independence and proceed using the SiZer. Methodology for handling of dependent observations have been developed, see Rondonotti et al. (2007) or for a method denoted SiNos, Godtliebsen et al. (2003).

4.3 Analysis of data

In Figure 4 the data set of yearly number of hurricanes is analysed. In the top panel a family of smoothed curves are shown; the original observations are shown as dots and the solid lines show smooths resulting from various choices of bandwidth. In the bottom panel, the derivatives are visualized with respect to both location and scale (bandwidth). The purple colour in the upper part of that plot indicates that the derivative is not found to be significantly different from zero, while the darker grey colour at the lower part is used at locations with too

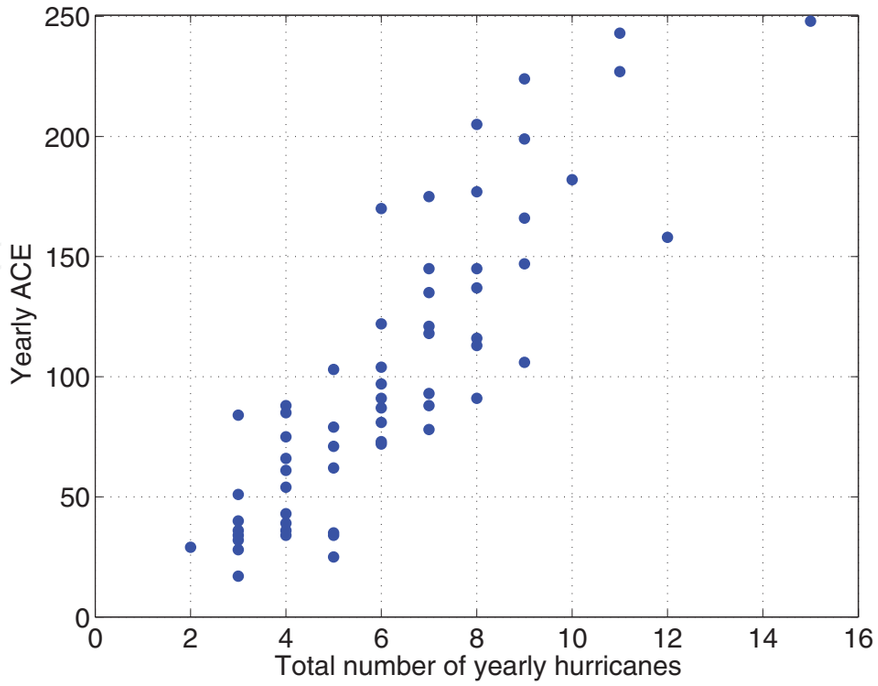


Fig. 2. Scatter plot: Total yearly number of hurricanes, yearly ACE.

few data available to do inference. Note that there are neither blue portions (corresponding to significant increases) nor red portions (corresponding to significant decreases). The solid horizontal bar in the SiZer plot shows a good data-driven bandwidth as suggested by Ruppert et al. (1995), here corresponding to approximately $h = 10^{0.66} = 4.6$ years.
The corresponding visualization for the ACE data is shown in Figure 5. The top panel shows original data and a family of smooths. Now, in the bottom panel, there are portions found with significant features: a decrease in the period 1950-1970 and an increase from about 1990 and onwards. These are valid for intermediate scales, values of $\log_{10}(h)$ in the interval $[0.9, 1.3]$. In summary, we are not able to conclude a significant change over longer time scales, which is of interest when discussing climate changes.

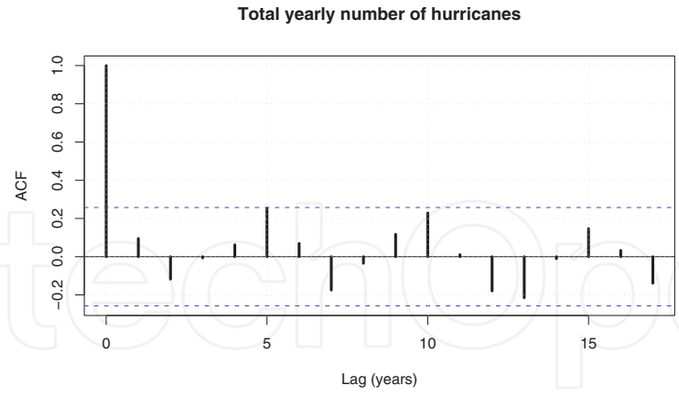
5. Other approaches

In this section we analyse the data sets by some other approaches. The purpose is to compare with the outcome of the SiZer technique; in statistical modelling of geological and environmental phenomena, it is often of interest to compare results from several methods and assumptions.

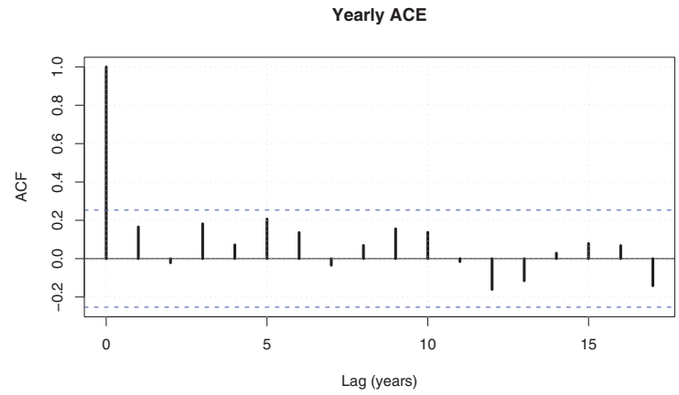
5.1 Poisson regression

Assume first that we study the total yearly number of hurricanes. Clearly, we model a discrete response variable, Y say, assumed to have a Poisson distribution. Suppose that the mean of Y is μ. In its most common form, the Poisson regression model is a special case of a generalized linear model which relates the mean to a linear function of covariates. In our case, we formulate the model

$$\mu_i = t_i \exp(\beta_0 + \beta_1 x_i), \quad i = 1, \ldots, n \tag{1}$$



(a) Autocorrelation function for total yearly number of hurricanes.



(b) Autocorrelation function for ACE.

Fig. 3. Autocorrelation functions for time series of total number of yearly hurricanes (top) and yearly ACE index (bottom)

where n is the number of years considered, t_i is the exposure time for each observation (one year, hence $t_i = 1$, all i), and x_i is an explanatory variable which is year. This explanatory variable is introduced since the purpose is to test for a possible trend over time: is the number of hurricanes increasing over the time period studied?
The statistical test for testing if $\beta_1 = 0$ is made by using the *deviance*. Parameter estimates are returned as ML estimates and l being the log-likelihood function, the deviance is given by

$$D = 2(l(\hat{\beta}_0, \hat{\beta}_1) - l(\beta_0, \beta_1))$$

This is used as follows: If $D > \chi^2_1(\alpha)$, reject the hypothesis. Comparison is made with table values $\chi^2_1(0.05) = 3.84$, $\chi^2_1(0.01) = 6.63$. The original data set results in $D = 0.35$, hence the hypothesis of zero slope cannot be rejected. However, when studying the regression diagnostics of the resulting model based on the original data, it is found that at least one outlier is present. Cook's distance is frequently used as a measure of leverage and values are presented in Figure 6. When again fitting a model, omitting years 1950 and 2005, the deviance $D = 0.103$ is found. Hence, again, the hypothesis is not rejected.

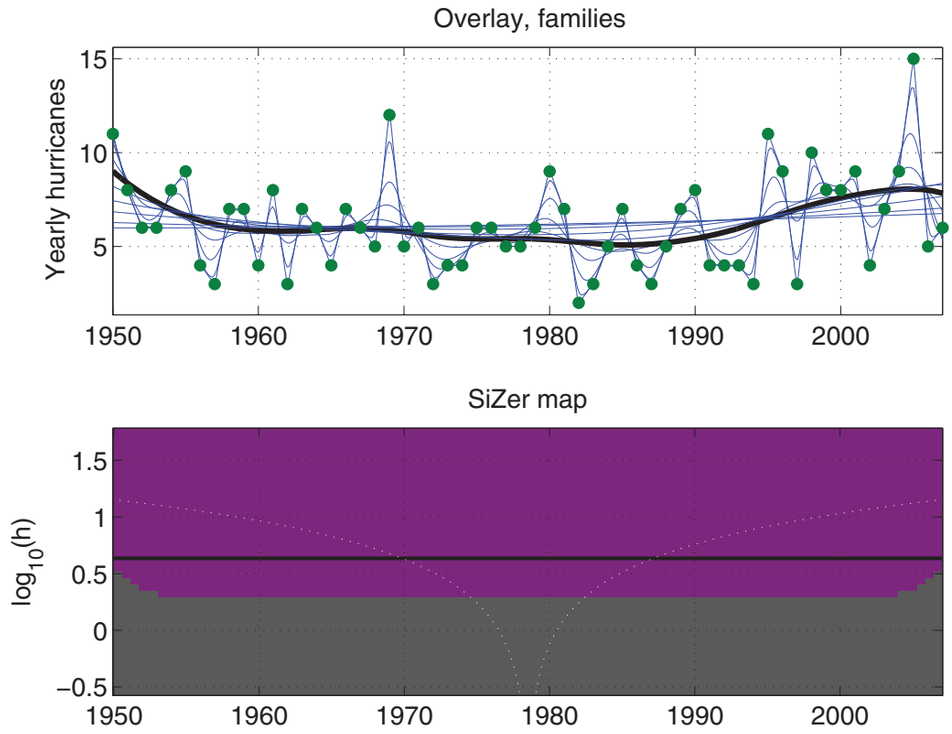


Fig. 4. SiZer analysis of total number of yearly hurricanes. Top: Overlay of family of smooths. Bottom: SiZer map.

To justify the use of Poisson regression, the possible presence of over-dispersion should be checked. A simple confidence interval for the ratio $\mathsf{V}[Y]/\mathsf{E}[Y]$ can be constructed following Rychlik and Rydén (2006), Section 7.3 (see also Brown and Zhao 2002). We find the 0.95 confidence interval $[0.61, 1.28]$ and hence the hypothesis of Poisson distribution is not rejected since 1 is within the interval.

5.2 Further models for count data

Several extensions of the model for Poisson regression have been presented in the literature, and although over-dispersion seems not to be a problem in our case, we use some of these. An expose is given by Zeileis, Kleiber and Jackman (2008), where hurdle and zero-inflated regression models are discussed along with the implementation in the R system for statistical computing. A model based on quasi-likelihood and a negative-binomial model were fitted to our data, but differences were small compared to the standard Poisson regression model.
In the literature, one also encounters zero-inflated Poisson regression (see e.g. Lambert, 1992). Focussing at only hurricanes of category 5, we face without doubt a data set with excess zeros. Using the package pscl in R, a zero-inflated model was fitted to the data of hurricanes of category 5, and the model was compared to the corresponding model obtained from conventional Poisson regression by a Vuong test. However, the zero-inflated model was not a significant improvement over the standard model.

5.3 Mann Kendall test

In environmental applications, a frequently used non-parametric test for trend detection is the Mann Kendall test. For our data sets, this results in the two-sided p-values 0.99 for the data

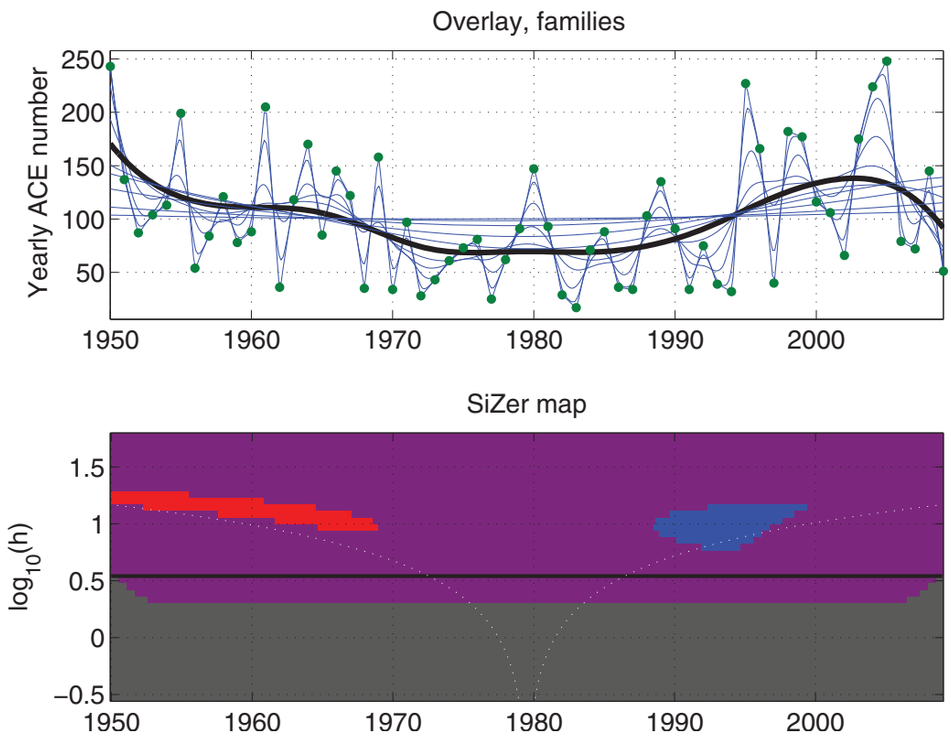


Fig. 5. SiZer analysis of ACE. Top: Overlay of family of smooths. Bottom: SiZer map.

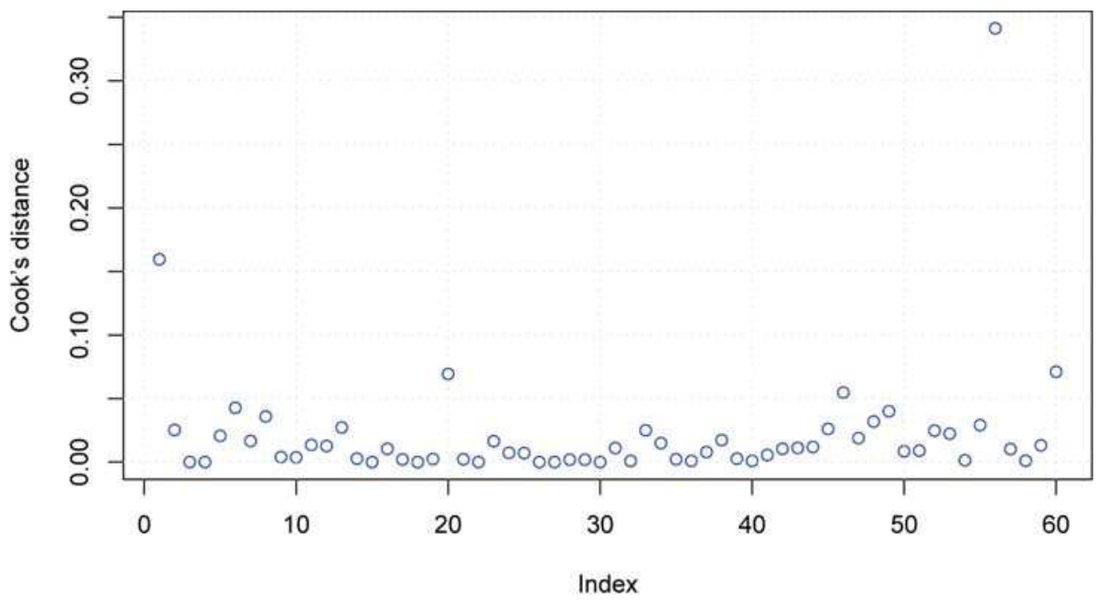


Fig. 6. Cook's distances after Poisson regression

of total yearly hurricanes, and 0.00014 for the ACE data. Again, for the total number of data, one cannot reject the hypothesis of no trend. For the ACE data, the test certainly indicates significant trend; however, keeping the result regarding time scales from the SiZer analysis in

mind and recalling that the data set is quite small, there is not enough evidence to point at changes at longer time horizons.

6. Conclusions

Detection of trends are important and common problems in statistical analysis of environmental data. An exploratory analysis by the SiZer method provides a useful complement to other statistical techniques and deserves to be more well-known to statisticans and researchers in the field. The debate on increasing trend of hurricane activity has now and then been intense in the research community. The result provided by this tool, originating from kernel-estimation methodology and scale-space theory, renders a comprehensible visualization of data and in addition a statistical test for significance of slope. For applications of the type presented in this paper, by first applying the SiZer analysis when exploring data, a hint of interesting time periods can be obtained and further statistical tests in those regions can then be performed by other methods, for example techniques based on Poisson regression.

7. Acknowledgements

I am grateful for the publicly available software for SiZer analysis in the form of Matlab routines written by Prof. James S. Marron, Department of Statistics, University of North Carolina, Chapel Hill. These were used to a large extent for the analysis carried out in this paper.

8. References

Brown, L.D., Zhao, L.H. (2002). A test of the Poisson distribution. *Sankhya*, Vol. 64, 611-625.

Chaudhuri, P., Marron, J.S. (1999). SiZer for exploration of structures in curves. *Journal of the American Statistical Association*, Vol. 94, 807-823.

Chaudhuri, P., Marron, J.S. (2000). Scale space view of curve estimation. *Annals of Statistics*, Vol. 28, 408-428.

Elsner, J.B. & Kara, A.B. (1999). *Hurricanes of the North Atlantic: Climate and Society.* Oxford University Press.

Elsner, J.B., Jagger, T., Niu, X.-F. (2000). Changes in the rates of North Atlantic major hurricane activity during the 20th century. *Geophysical Research Letters* Vol. 27, 1743-1746.

Elsner, J.B. & Jagger, T.H. (2004). A hierarchical Bayesian approach to seasonal hurricane modeling. *Jouurnal of Climate*, Vol. 17, 2813-2827.

Godtliebsen, F., Olsen, L.R., Winther, J.-G. (2003). Recent developments in time series analysis: Examples of use in climate research. *Geophysical Research Letters*, Vol. 30, 1654-1657.

Han, S.-R., Guikema, S.D. & Quiring, S.M. (2009). Improving the predictive accuracy of hurricane power outage forecasts using generalized additive models. *Risk Analysis*, Vol. 29, 1443-1453.

Katz, R.W. (2002). Stochastic modeling of hurricane damage. *Journal of Applied Meteorology*, Vol. 41, 754-762.

Jagger, T.H., Elsner, J.B. & Saunders, M.A. (2008). Forecasting U.S. insured hurricane losses. In: *Climate Extremes and Society*, Diaz, H.F. & Murnane, P.J., 189-208, Cambridge University Press.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics*, Vol. 34, 1-14.

Landsea, C.W., Pielke Jr, R.A., Mestas-Nunez, A.M., Knaff, J.A. (1999). Atlantic basin hurricanes: indices of climatic changes. *Climatic Change*, Vol. 42, 89-129.

Liljequist, G.H. (1949). On fluctuations of the summer mean temperature in Sweden. *Geografiska Annaler*, Vol. 31, 159-178.

Rondonotti, V., Marron, J.S., Park, C. (2007). SiZer for time series: A new approach to the analysis of trends. *Electronic Journal of Statistics*, Vol. 1, 268-289.

Ruppert, D., Sheather, M.J., Wand, M.P. (1995). An effective bandwidth selector for least squares regression. *Journal of the American Statistical Association*, Vol. 90, 1257-1270.

Rychlik, I. & Rydén, J. (2006). *Probability and Risk Analysis. An Introduction for Engineers.* Springer–Verlag.

Rydén, J. (2010). Exploring possibly increasing trend of hurricane activity by a SiZer approach. *Environmental and Ecological Statistics*, Vol. 17, 125-132.

Solow, A.R. (1989). Statistical modeling of storm counts. *Journal of Climate*, Vol. 2, 131-136.

Webster, P.J., Holland, G.J., Curry, J.A., Chang, H.-R. (2005). Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science*, Vol. 309, 1844-1846.

Weckström, J., Korhola, A., Erästö, P., Holmström, L. (2006). Temperature patterns over the past eight centuries in Northern Ferroscandia inferred from sedimentary diatoms. *Quaternary Research*, Vol. 66, 78-86.

Zeileis, A., Kleiber, C. & Jackman, S. (2008). Regression models for count data in R. *Journal of statistical software*, Vol. 27, No. 8, 1-25.

IntechOpen



Recent Hurricane Research - Climate, Dynamics, and Societal Impacts

Edited by Prof. Anthony Lupo

ISBN 978-953-307-238-8
Hard cover, 616 pages
Publisher InTech
Published online 19, April, 2011
Published in print edition April, 2011

This book represents recent research on tropical cyclones and their impact, and a wide range of topics are covered. An updated global climatology is presented, including the global occurrence of tropical cyclones and the terrestrial factors that may contribute to the variability and long-term trends in their occurrence. Research also examines long term trends in tropical cyclone occurrences and intensity as related to solar activity, while other research discusses the impact climate change may have on these storms. The dynamics and structure of tropical cyclones are studied, with traditional diagnostics employed to examine these as well as more modern approaches in examining their thermodynamics. The book aptly demonstrates how new research into short-range forecasting of tropical cyclone tracks and intensities using satellite information has led to significant improvements. In looking at societal and ecological risks, and damage assessment, authors investigate the use of technology for anticipating, and later evaluating, the amount of damage that is done to human society, watersheds, and forests by land-falling storms. The economic and ecological vulnerability of coastal regions are also studied and are supported by case studies which examine the potential hazards related to the evacuation of populated areas, including medical facilities. These studies provide decision makers with a potential basis for developing improved evacuation techniques.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jesper Rydén (2011). Statistical Techniques for Exploring Possibly Increasing Trend of Hurricane Activity, Recent Hurricane Research - Climate, Dynamics, and Societal Impacts, Prof. Anthony Lupo (Ed.), ISBN: 978-953-307-238-8, InTech, Available from: http://www.intechopen.com/books/recent-hurricane-research-climate-dynamics-and-societal-impacts/statistical-techniques-for-exploring-possibly-increasing-trend-of-hurricane-activity

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

www.intechopen.com

IntechOpen

IntechOpen

© 2011 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike-3.0 License, which permits use, distribution and reproduction for non-commercial purposes, provided the original is properly cited and derivative works building on this content are distributed under the same license.

IntechOpen

IntechOpen