On the Attribution of a Single Event to Climate Change

GERRIT HANSEN

Potsdam Institute for Climate Impact Research, Potsdam, Germany

MAXIMILIAN AUFFHAMMER

Department of Agricultural and Resource Economics, University of California, Berkeley, Berkeley, California, and National Bureau of Economic Research, Cambridge, Massachusetts

ANDREW R. SOLOW

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

(Manuscript received 4 June 2014, in final form 16 September 2014)

ABSTRACT

There is growing interest in assessing the role of climate change in observed extreme weather events. Recent work in this area has focused on estimating a measure called attributable risk. A statistical formulation of this problem is described and used to construct a confidence interval for attributable risk. The resulting confidence is shown to be surprisingly wide even in the case where the event of interest is unprecedented in the historical record.

1. Introduction

Climate change is predicted to increase the frequency of extreme weather events like intense hurricanes (Webster et al. 2005) and heat waves (Meehl and Tebaldi 2004). It is natural, therefore, to ask when an event such as the European heat wave in 2003 or Hurricane Sandy in 2012 occurs if it can be attributed to climate change. This attribution question has gained some prominence with efforts to assess liability for weather-related damages due to climate change (Allen 2003). Recent work on single-event attribution has focused either implicitly or explicitly on a quantity known as attributable risk (Bindoff et al. 2014; Rahmstorf and Coumou 2011; Stott et al. 2004). The purpose of this note is to present a statistical formulation for attributable risk and to discuss its estimation with a particular emphasis on the construction of a confidence interval.

2. A statistical formulation

A natural statistical formulation of single-event attribution is in terms of a stochastic point process (Cox and Isham 1980). A stochastic point process is the classical model of a series of events occurring in some way randomly through time. Such models have been used to describe a variety of extreme weather events including heat waves (e.g., Furrer et al. 2010) and hurricanes (e.g., Jagger and Elsner 2006). We note that the definition of the events of interest can include features such as intensity, location, and seasonality: for example, wintertime exceedances of a temperature threshold or category 5 hurricanes above a certain latitude.

A point process is partially characterized by a rate function that gives the instantaneous frequency of events. When this rate function is constant, the point process is said to be stationary. For a stationary point process with constant rate μ, the expected number of events in a period of length T is μT. For simplicity, we will focus here on the case where climate change causes a shift from one stationary point process to another stationary point process. As discussed below, however, the results presented in this paper also apply to the nonstationary case.

Single-event attribution asks the following: Given that an event has occurred after the climate has changed, was it or was it not caused by climate change? This question implies that, once climate has changed, the point process of events represents the superposition of a point process of events that would have occurred in the absence of

Corresponding author address: Gerrit Hansen, Potsdam Institute for Climate Impact Research, Research Domain 1, Telegrafenberg A26, 14473 Potsdam, Germany.
E-mail: hansen@pik-potsdam.de

climate change and a point process of events that would not have occurred in the absence of climate change and are, therefore, attributable to climate change. Moreover, these point processes must be independent; otherwise, the former would inherit a climate change effect through the latter.

Suppose that the rate before climate change is μ. Following climate change, this rate increases to $\beta\mu$ with $\beta \geq 1$. It is straightforward to show that, conditional on an event occurring after the climate has changed, the probability that it was caused by climate change is

$$p = 1 - 1/\beta. \tag{1}$$

It is this probability that most recent papers on single-event attribution seek to assess. Borrowing from epidemiology, the probability in (1) is referred to as the risk attributable to climate change or simply the attributable risk (Walter 1976). The definition of attributable risk only makes sense if $\beta \geq 1$: that is, if climate change increases the rate of events. In cases where climate change decreases this rate, the quantity $1 - \beta$ is the risk attributable to the absence of climate change for an event that occurred prior to climate change.

For convenience, we refer to a comparison of event rates before and after climate change. In practice, it is common to compare the rate in an earlier period to the rate in a later period without assuming that the former is completely free from the effect of climate change. In that case, the issue is one of attribution to a change in climate that has occurred between the two periods.

3. Estimation of attributable risk

In practical applications, attributable risk is not known and has to be estimated. In this section, we discuss this estimation with a particular focus on the construction of a confidence interval.

Let the random variable X be the number of events in a pre–climate change period of length T_1 and the random variable Y be the number of events in a post–climate change period of length T_2. The counts X and Y can be based either on historical records or on simulations from a climate model. In the former case, it is important that the event of interest not be selected because of its rarity in the pre–climate change record. We will assume that both pre– and post–climate change events follow stationary Poisson processes. For a stationary Poisson process, the numbers of events in non-overlapping periods are independent Poisson random variables with means proportional to the lengths of the periods (Cox and Isham 1980). As noted below, the Poisson model can be extended to allow for a non-stationary rate function. Although not all point

processes are Poisson processes, there is theoretical support for their use in modeling rare events (Barbour 1988).

Let x and y be the observed values of X and Y, respectively. Przyborowski and Wilenski (1940) gave an expression for the joint distribution of independent Poisson random variables. For the model outlined here, the probability of observing x and y can be decomposed as

$$\text{prob}(X = x, Y = y) = \frac{[\mu(T_1 + \beta T_2)]^{x+y} \exp[-\mu(T_1 + \beta T_2)]}{(x+y)!}$$
$$\times \frac{(x+y)!}{x!y!}[\tau/(\tau+\beta)]^x[\beta/(\tau+\beta)]^y, \tag{2}$$

where $\tau = T_1/T_2$. The first term is the Poisson probability of observing a total of $x + y$ events and the second term is the conditional probability that x of these events occurred in the pre–climate change period and y occurred in the post–climate change period. This latter probability is given by the binomial distribution with $x + y$ trials and success probability $\tau/(\tau + \beta)$. The maximum likelihood (ML) estimates of μ and β are the natural ones,

$$\hat{\mu} = x/T_1 \quad \text{and} \tag{3}$$

$$\hat{\beta} = \tau y/x, \tag{4}$$

where x and y are the observed values of X and Y and the ML estimate of attributable risk p is

$$\hat{p} = 1 - x/\tau y. \tag{5}$$

Because there is positive probability that $Y = 0$, this estimate has neither finite mean nor variance. This can be avoided by conditioning on the event that $Y > 0$ so that Y has a so-called zero-truncated Poisson distribution. This conditioning seems reasonable as at least one post–climate change event must have occurred to trigger the attribution exercise. Rather than pursue this here, we will instead focus on the construction of a confidence interval for p.

We will proceed as follows: If the lower and upper bounds of a $1 - \alpha$ confidence interval for β are β_L and β_U, respectively, then the lower and upper bounds of a $1 - \alpha$ confidence interval for p are $1 - 1/\beta_L$ and $1 - 1/\beta_U$, respectively, so that a confidence interval for p can be constructed from a confidence interval for β. Under the model outlined above, β is the ratio of Poisson means. The literature on constructing a confidence interval for the ratio of Poisson means dates back at least to Chapman (1952) and several approaches are described in Price and Bonett (2000). Here, we will adopt the common approach

of eliminating the nuisance parameter μ by conditioning on the observed value $n = x + y$ of $X + Y$. As noted, the conditional distribution of X and Y given that their sum is $x + y$ is binomial with $x + y$ trials and success probability $\tau/(\tau + \beta)$. If L and U are the lower and upper bounds, respectively, of a $1 - \alpha$ confidence interval for this probability, then the corresponding lower and upper bounds of a $1 - \alpha$ confidence interval for β are

$$\beta_L = \tau(1 - U)/U \quad \text{and} \tag{6}$$

$$\beta_U = \tau(1 - L)/L, \tag{7}$$

respectively. In this step, we will use the approximate confidence interval originally proposed by Wilson (1927) and recommended by Brown et al. (2001) for a binomial probability with

$$L = \frac{x + z^2/2}{n + z^2} - \frac{z\sqrt{n}}{n + z^2}\sqrt{\hat{q}(1 - \hat{q}) + z^2/(4n)} \quad \text{and} \tag{8}$$

$$U = \frac{x + z^2/2}{n + z^2} + \frac{z\sqrt{n}}{n + z^2}\sqrt{\hat{q}(1 - \hat{q}) + z^2/(4n)}, \tag{9}$$

where z is the upper $\alpha/2$ quantile of the standard normal distribution and $\hat{q} = x/n$. Finally, the corresponding lower and upper bounds of an approximate $1 - \alpha$ confidence interval for p are

$$p_L = 1 - 1/\beta_L \quad \text{and} \tag{10}$$

$$p_U = 1 - 1/\beta_U, \tag{11}$$

respectively.

The actual coverage of the Wilson confidence interval is close to its nominal level unless x is close to 0. For x small but positive, Brown et al. (2001) described a modification that improves coverage. Although we will not consider this here, we will consider the important case where $x = 0$ (i.e., the events of interest are without precedent prior to climate change). In this case, the upper bound of an exact $1 - \alpha$ confidence interval for $\tau/(\tau + \beta)$ is $1 - \alpha^{1/n}$ (Jovanovic and Levy 1997). It follows that the lower bound of an exact $1 - \alpha$ confidence interval for p when $x = 0$ is

$$p_L^0 = 1 - (1 - \alpha^{1/n})/(\tau\alpha^{1/n}). \tag{12}$$

As noted, although we have focused on the case where both the pre– and post–climate change Poisson processes are stationary, the results of this section extend to the case where either or both is nonstationary. Briefly, for a nonstationary Poisson process with time-varying rate function $\mu(t)$, the number of events in the interval (u, v) has a Poisson distribution with mean $\overline{\mu}(v - u)$, where

TABLE 1. Maximum likelihood estimate $\hat{p}$ of attributable risk and lower p_L and upper p_U bounds of an approximate 0.95 confidence interval for p for selected values of τ, x, and y.

τ	x	y	$\hat{p}$	p_L	p_U
0.5	1	4	0.5	-2.33	0.93
	10	40		0.01	0.75
	25	100		0.23	0.68
	1	10	0.8	-0.21	0.97
	5	50		0.51	0.92
	25	250		0.70	0.87
1.0	2	4	0.5	-1.33	0.89
	20	40		0.15	0.71
	50	100		0.30	0.64
	2	10	0.8	0.19	0.95
	10	50		0.61	0.90
	50	250		0.73	0.85
2.0	4	4	0.5	-0.82	0.86
	40	40		0.23	0.68
	100	100		0.34	0.62
	4	10	0.8	0.40	0.93
	20	50		0.67	0.88
	100	250		0.75	0.84

$$\overline{\mu} = \int_u^v \mu(t)\, dt/(v - u) \tag{13}$$

is the mean rate during this interval. It follows that β_L and β_U in (6) and (7) are the bounds of a $1 - \alpha$ confidence interval for the ratio of the mean rate in the post–climate change period to the mean rate in the pre–climate change period and consequently that p_L and p_U in (10) and (11) are the bounds of a $1 - \alpha$ confidence interval for attributable risk based on these mean rates. Of course, if the rate function increases continuously during the post–climate change period, the attributable risk for events late in this period is greater than that for events earlier in the period. It is possible to develop a continuous measure of attributable risk, by modeling the rate function, but the construction of a confidence interval would be more challenging.

4. Results

To illustrate the calculations outlined in the previous section, Table 1 presents the ML estimate $\hat{p}$ and the bounds of the approximate 0.95 confidence interval for p for selected positive values of x and y and selected values of τ. A negative lower confidence bound in Table 1 indicates that the confidence interval for β contains values less than 1 (i.e., a decrease in the rate of events cannot be ruled out). Table 2 presents the lower bound of a 0.95 confidence interval for p for selected values of y with $x = 0$ and selected values of τ. In all cases in Table 2, the point estimate of p is equal to 1 as is the upper bound of the confidence interval. Again, a negative lower

TABLE 2. Lower bound p_L^0 of a 0.95 confidence interval for attributable risk for selected values of τ and y with $x = 0$.

τ	y	p_L^0
0.5	1	−37.0
	2	−5.94
	5	−0.64
	10	0.30
1.0	1	−18.0
	2	−2.47
	5	0.18
	10	0.65
2.0	1	−8.5
	2	−0.74
	5	0.59
	10	0.83

confidence bound indicates that a decrease in the rate of events cannot be ruled out. It is clear that even establishing that the rate of events has increased with climate change (i.e., $p > 0$) may not be possible when the events are rare. This is true even if the events are without precedent in the pre–climate change record. Even if this basic fact can be established, a surprisingly large number of events may be needed before attributable risk can be estimated with high confidence.

As a further illustration, we applied the methods of the previous section to data documenting intense (categories 4 and 5) hurricanes in the North Atlantic over the period 1950–2012. These data were extracted from the Atlantic hurricane best-track dataset maintained at the U.S. National Hurricane Center (NOAA 2014). The effect of climate change on the frequency of such hurricanes and the quality of the historical data remain unsettled (Knutson et al. 2010), and we stress that this is intended as an illustration. Over the 30-yr period 1950–79, there were a total of 39 intense North Atlantic hurricanes while over the following 33-yr period, 1980–2012, there were 53 such hurricanes. If we assume that the effect of climate change over the entire 63-yr period was to increase the rate of these hurricanes, then the ML estimate of the estimated probability that a hurricane in the later period is attributable to climate change is 0.19 and an approximate 0.95 confidence interval for this probability is $(-0.17, 0.44)$. The negative lower bound of this confidence interval indicates that a decline in the rate of intense hurricanes between these periods cannot be ruled out. At the same time, the upper bound of 0.44 indicates that neither can a near doubling of this rate be ruled out. It is worth noting that the rate of intense hurricanes varies over both the seasonal time scale and the interannual time scale (e.g., due to ENSO variability), so this is an example of an application to a nonstationary process.

5. Discussion

This note has outlined a statistical formulation of the attribution of a single event to climate change and has used this formulation to provide a confidence interval for attributable risk. Formulating single-event attribution in this way raises two fundamental issues. First, as noted, underlying the concept of attributable risk is a dichotomy between events that would have occurred in the absence of climate change and events that would not. This dichotomy makes sense in epidemiology (and in other contexts). For example, some cases of lung cancer are caused by smoking, others are not, and it is natural to ask about the risk of lung cancer attributable to smoking. It is not so clear, however, that attributable risk makes sense in the context of climate change. While the effects of smoking are confined to the smoker (and perhaps those around him), the effect of climate change is pervasive and the notion that, once the climate has changed, some weather events would have occurred exactly as they did in its absence may not be tenable. To be clear, this is not at all to say that a change in the rate of events cannot be attributed to climate change, only that the superposition argument on which attributable risk is based may not be tenable.

Second, even if the notion of attributable risk makes sense in the context of climate change, the quantity p is simply a function of the rates of events before and after climate change and not particularized to an individual event. To put it another way, the attributable risk is the same for all post–climate change events. In this sense, it is not really single-event attribution.

Turning to the results of the previous section, it is clear that uncertainty about attributable risk can remain high unless both the number of observed events and the effect of climate change are large.

REFERENCES

Allen, M., 2003: Liability for climate change. *Nature,* **421,** 891–892, doi:10.1038/421891a.

Barbour, A. D., 1988: Stein's method and Poisson process convergence. *J. Appl. Probab.,* **25,** 175–184, doi:10.2307/3214155.

Bindoff, N. L., and Coauthors, 2014: Detection and attribution of climate change: From global to regional. *Climate Change 2013: The Physical Science Basis,* T. F. Stocker et al., Eds., Cambridge University Press, 867–952.

Brown, L. D., T. T. Cai, and A. DasGupta, 2001: Interval estimation for a binomial proportion. *Stat. Sci.,* **16,** 101–117, doi:10.1214/ss/1009213286.

Chapman, D. G., 1952: On tests and estimates for the ratio of Poisson means. *Ann. Inst. Stat. Math.,* **4,** 45–49, doi:10.1007/BF02949788.

Cox, D. R., and V. Isham, 1980: *Point Processes.* Chapman and Hall, 188 pp.

Furrer, E. M., R. W. Katz, M. D. Walter, and R. Furrer, 2010: Statistical modeling of hot spells and heat waves. *Climate Res.,* **43,** 191–205, doi:10.3354/cr00924.

Jagger, T. H., and J. B. Elsner, 2006: Climatology models for extreme hurricane winds near the United States. *J. Climate,* **19,** 3220–3236, doi:10.1175/JCLI3913.1.

Jovanovic, B. D., and P. S. Levy, 1997: A look at the rule of three. *Amer. Stat.,* **51,** 137–139.

Knutson, T. R., and Coauthors, 2010: Tropical cyclones and climate change. *Nat. Geosci.,* **3,** 157–163, doi:10.1038/ngeo779.

Meehl, G. A., and C. Tebaldi, 2004: More intense, more frequent, and longer lasting heat waves in the 21st century. *Science,* **305,** 994–997, doi:10.1126/science.1098704.

NOAA, cited 2014: Atlantic hurricane database (HURDAT2) 1851–2013. [Available online at http://www.nhc.noaa.gov/data#hurdat.]

Price, R. M., and D. Bonett, 2000: Estimating the ratio of two Poisson means. *Comput. Stat. Data Anal.,* **34,** 345–356, doi:10.1016/S0167-9473(99)00100-0.

Przyborowski, J., and H. Wilenski, 1940: Homogeneity of results in testing samples from Poisson series: With an application to testing clover seed for dodder. *Biometrika,* **31,** 313–323.

Rahmstorf, S., and D. Coumou, 2011: Increase of extreme events in a warming world. *Proc. Natl. Acad. Sci. USA,* **108,** 17 905–17 909, doi:10.1073/pnas.1101766108.

Stott, P. A., D. A. Stone, and M. R. Allen, 2004: Human contribution to the European heatwave of 2003. *Nature,* **432,** 610–614, doi:10.1038/nature03089.

Walter, S. D., 1976: The estimation and interpretation of attributable risk in health research. *Biometrics,* **32,** 829–849, doi:10.2307/2529268.

Webster, P. J., G. J. Holland, J. A. Curry, and H.-R. Chang, 2005: Changes in tropical cyclone number, duration, and intensity in a warming environment. *Science,* **309,** 1844–1846, doi:10.1126/science.1116448.

Wilson, E., 1927: Probable inference, the law of succession, and statistical inference. *J. Amer. Stat. Assoc.,* **22,** 209–212, doi:10.1080/01621459.1927.10502953.