Analysis of statistical power reconciles drought-conflict results in Africa

Solomon M. Hsiang, Marshall Burke, Edward Miguel, Kyle C. Meng, and Mark A. Cane

Working Paper Series
Center for Effective Global Action
University of California

This paper is posted at the eScholarship Repository, University of California. http://escholarship.org/uc/cega_wps
Copyright © 2012 by the author(s).

The CEGA Working Paper Series showcases ongoing and completed research by faculty affiliates of the Center. CEGA Working Papers employ rigorous evaluation techniques to measure the impact of large-scale social and economic development programs, and are intended to encourage discussion and feedback from the global development community.

Recommended Citation:

Analysis of statistical power reconciles
drought-conflict results in Africa

Solomon M. Hsiang*1,2, Marshall Burke3, Edward Miguel1,2, Kyle C. Meng4, and
Mark A. Cane5,6

1University of California, Berkeley
2National Bureau of Economic Research
3Stanford University
4University of California, Santa Barbara
5Columbia University
6Lamont-Doherty Earth Observatory of Columbia University

*shsiang@berkeley.edu
Introduction

Whether changes in climate substantially shape human conflict is a question of considerable recent interest to both academics and policymakers. Despite a substantial body of evidence for a strong association between climate and conflict,\(^1\) it remains widely claimed that large-sample empirical evidence linking climatic conditions and modern human conflict in Africa is mixed and thus any reported evidence for a strong association should be discounted. Theisen, Holtermann, and Buhaug\(^2\) (henceforth THB) is one of the studies used to support this claim. Here we show that the results in THB are not inconsistent with earlier studies that report a substantial effect of climate on conflict. We demonstrate using power calculations and Monte Carlo simulations that even if a large association between climate and conflict existed in the data, the approach of THB would not be able to reliably distinguish this association from a null effect, indicating that the approach taken by THB is statistically underpowered in this context. Therefore THB’s analysis provides no basis for discarding earlier analyses and THB’s conclusions drawn from this analysis overstate the extent to which they disagree with the literature.\(^3\)

We also demonstrate that THB’s stated advantage from using exceptionally high resolution data is unlikely to be realized in their analysis, since high resolution rainfall data was not actually collected in the majority conflict zones studied by THB. Although unremarked in the original analysis, the rainfall data in THB are interpolations of sparse and incomplete rainfall measurements.

Consistency of THB with earlier findings

THB break up Africa into 10,671 0.5° × 0.5° pixels which they assign a measure of rainfall or water availability and a measure of conflict for the years 1960-2004. They then regress the pixel-by-year measure of conflict on the measures of climate and evaluate whether the coefficient on their climate variable is statistically significant. Because they find these coefficients are not significant, THB conclude that

---


tween drought and civil war onset,” (p. 105). However, the results presented in THB have such large uncertainty that they are simultaneously consistent with both “no effect” of climate on conflict and effects that are as large as or larger than statistically significant results reported in earlier studies.4

To illustrate why THB’s conclusion is inconsistent with their analysis, we replicate THB’s Figure 3, which examined the change in relative risk of conflict when moving from the 10th to 90th percentile of their many different drought measures (Figure 1). But unlike THB, we plot the x-axis (relative risk) without using a logarithmic scale, allowing us to more clearly observe the large confidence intervals around their point estimates. As shown in Figure 1, parameter estimates for 19 out of 21 of their drought variables are consistent with a 50% increase in conflict, an effect similar in magnitude to results of other studies that report a link between climate and conflict.5 Moreover, 14 out of 21 estimates are consistent with a 100% increase of conflict risk, and only 1 out of 21 of their estimates can reject a 10% increase in conflict. This wide range of estimates does not support THB’s claim that the effect of climate on conflict is exactly zero. Had THB begun their analysis using the null hypotheses that “drought increases conflict risk by 50%” – a logical null given previous findings6 – they would have found that over 90% of their estimates were consistent with this hypothesis.

Sample size and statistical power of THB

The complete sample in THB contains 363,811 pixel-by-year observations, of which 59 are coded as experiencing conflict (a binary variable) using a definition proposed by THB. Thus these conflicts are extraordinarily rare events: the unconditional probability that any location exhibits conflict in a randomly selected year is

\[
\frac{59}{363,811} = 0.00016 = 0.016\%.
\]

Because the likelihood of conflict is so rare in this sample (a 1 in 6,250 event), a large proportional change in the risk of conflict still constitutes a very small change in the overall likelihood of conflict. If some variable caused the risk of conflict to

---


5Hsiang, Burke, and Miguel, “Quantifying the influence of climate on human conflict”.

6Hsiang, Burke, and Miguel, “Quantifying the influence of climate on human conflict”.

---
Figure 1: Replication of THB Figure 3. Estimates and confidence intervals indicate THB's estimated change in relative risk of conflict when moving from the 10th to 90th percentile of various drought measures. The three vertical dashed lines indicate that a measure of drought increases the relative risk of conflict to 1.1, 1.5 and 2, indicating that conflict risk rises by 10%, 50%, and 100%, respectively.
Figure 2: The sample size required to distinguish a minimal effect of drought from a null of “no effect”. Dashed horizontal line indicates the sample size used in THB.

double, then the change in the probability of conflict would only rise by about a fiftieth of a percentage point. Using statistical techniques to reliably distinguish such small changes from “no effect” requires either that sample sizes are very large or that the level of noise in data is very small. In this section we compute the sample size that THB would need in order to reliably reject the findings of earlier studies in this context.

Most of THB’s drought measures—as well as their preferred measure—are binary, and binary treatment variables lend themselves to sample size calculations that are easily implemented using standard statistical packages. These calculations determine, for a given effect size $\theta$ (i.e. $\%$ change from baseline) and significance level (i.e. willingness to suffer a Type I error, typically denoted $\alpha$), the sample size required to distinguish a null of no effect from $\theta$ with $\kappa$ confidence – where $\kappa$ (or “power”)
is the probability that the null is rejected when it is indeed false. To perform these calculations in the simplest possible scenario, the analyst needs (1) an estimate of the baseline mean and variance of the outcome of interest in the population of interest, (2) to set $\alpha$ and $\kappa$, and (3) to choose a minimum effect size $\theta$ that must be distinguished from the null of “no effect.”

We choose the standard significance level $\alpha = 0.05$ and $\kappa = 0.8$, a relatively generous tolerance for Type II error. The standard $\kappa = 0.9$, would only further increase the necessary sample sizes. We assume that there is no intraclass correlation in outcomes within a given grid-cell over time and assume that 16.8% of grid-cell-years experience a drought – consistent with THB’s $\text{sp}_{6}\text{dum}$ variable. Baseline conflict risk is kept fixed at the 0.016% used by THB.

Figure 2 shows the sample sizes that are needed to distinguish the null hypothesis that drought has “no effect” ($\theta = 0$) from hypotheses that the effect $\theta$ ranges in size from 10-600% above average levels of conflict risk. THB’s sample size in their regressions (17,393 observations) is shown as a horizontal line. We calculate that THB would need 51.4 million observations to detect or reject a 10% increase in conflict, and they would need about 2 million observations to detect or reject a 50% increase in conflict—an effect consistent with earlier studies. We estimate that given the level of noise in their data and their 17,393 observations, the smallest effect of drought that THB could reliably reject is an increase of roughly 550%, an effect well beyond the range of estimates proposed in previous studies.

**Evaluating the method in THB with Monte Carlo simulations**

As a final demonstration that the approach of THB does not provide insight into the presence or absence of a relationship between climate and conflict, we generate synthetic data with the general structure of the data used by THB, ensuring that in this data set conflict and climate are related by construction. We then test whether the approach of THB can recover this fact. We repeat this exercise many times and compute the likelihood that THB would have correctly rejected the null hypothesis that climate and conflict are unrelated in favor of the hypothesis that these variables are related—which we know is true in our synthetic data. Our procedure is as follows.

We generate 363,811 normally distributed $\text{rainfall}$ observations with mean=0 and $\sigma=1$. The $\text{precdev}$ rainfall variable used by THB is similar to normal but has greater kurtosis, which makes estimates noisier. This implies that our assumption of normal $\text{rainfall}$ makes our approach relatively more forgiving, in terms of signal detection, than that of THB$^7$.

---

$^7$The variable that THB prefer is a dummy variable $\text{sp}_{6}\text{dum}$ that is 1 under drought conditions, however it has even
We then prescribe rainfall to exert an influence on the risk of conflict that is consistent with a recent meta-analysis of this relationship: a 1σ increase in rainfall raises conflict risk 10%. We do this by constructing synthetic observations of conflict\_risk, a latent variable that is not observed

\[\text{conflict\_risk} = 0.00016217 \times (1 + 0.1 \times \text{rainfall}) + \epsilon\]  

where the mean is set to 0.00016217 since it is the unconditional risk of conflict in THB. All non-climatic factors that influence conflict risk are described by \(\epsilon\), a normally distributed disturbance with standard deviation of 0.01. We select a standard deviation of 0.01 because the mean conflict risk in THB is 0.00016 and if conflict risk is distributed Poisson (a reasonable assumption because it is an infrequent event that can occur at any moment) then we would expect the variance in risk to equal its mean—suggesting a standard deviation of 0.00016\(^{0.5} = 0.013 \sim 0.01\).

Because only actual conflict is observed, not the latent variable conflict\_risk, we use conflict\_risk to determine which observations are most likely to experience actual conflict. To ensure that our data has the same properties as that of THB, we code the 59 observations with highest conflict risk as \(\text{conflict} = 1\), with the remaining 363,752 observations coded as \(\text{conflict} = 0\). The subsequent data set thus contains exactly the same number of observations and conflicts as the data used by THB, but is constructed with an underlying pattern of conflict risk that rises 10% per 1σ in rainfall. The question we now ask is whether a regression approach can recover this rainfall signal in the conflict data.

To do this, we repeat the above procedure 10,000 times and each time mimic THB’s approach by estimating the logistic regression

\[\Pr(\text{conflict}|\text{rainfall}) = \frac{\exp(\beta \times \text{rainfall})}{1 + \exp(\beta \times \text{rainfall})}\]  

Following THB, we use a t-test to evaluate whether \(\beta\) is statistically significant at the 5% level and find that this approach achieves a significant result only 5.4% of the time (see Figure 3A), even though we know with certainty that rainfall influences conflict in this synthetic data. 94.6% of the time we incorrectly fail to reject the null hypothesis that rainfall and conflict are unrelated (false negatives).

Because the reliability of THB’s approach depends on the signal-to-noise ratio of the data, and lower variance than precdev, making signal detection even more difficult.

Hsiang, Burke, and Miguel, “Quantifying the influence of climate on human conflict”.  

9We use this latent variable approach to construct synthetic data because it matches the data generating process that THB assume, implicitly, by using logistic regression.
Figure 3: (A) Monte Carlo results evaluating whether the approach of THB can reliably detect a true increase of conflict by 10% for a 1σ increase in rainfall (similar to meta-analysis results\textsuperscript{10}) when the variance in conflict risk equals its mean (it is approximately Poisson). Histogram displays the p-values from 10,000 simulations. P-values above 0.05 indicate cases where the approach of THB (incorrectly) fails to reject the null hypothesis of “no effect.” (B) Same as Panel A, except the true effect of rainfall on conflict is 100% per 1σ.
this ratio is determined by the data generating process that produces conflict risk, we adjusted the assumptions used in Equation 1 to see whether THB’s approach would exhibit substantially greater statistical power under more favorable data conditions. To do this, we repeated our procedure but altered the coefficient of rainfall in Equation 1 so that it was 10 times larger. This implies a relationship where conflict risk rises 100% for a 1σ increase in rainfall. Amplifying the climate signal in the conflict data improves the likelihood of signal detection, however the regression approach of THB still fails to detect this effect 89% of the time (see Figure 3B).

We then tried improving the signal-to-noise ratio by lowering the amount of statistical noise in the synthetic data. We did this by using the original form of Equation 1, but reduced the standard deviation of ϵ by a factor of ten to 0.001. This implies that the residual variance in conflict risk (0.001² = 0.000001) is 100 times smaller than its mean, implying a distribution of conflict risk that is much narrower than a Poisson distribution (extreme under-dispersion). Even under this optimistic assumption, the regression approach failed to detect the influence of rainfall on conflict 88% of the time. Mathematically, this case is essentially identical to the case above and results mirror those in Figure 3B.

We conclude that under both reasonable and extremely generous assumptions, the approach used by THB cannot determine whether climate influences conflict or not. Under either scenario, the results of THB will almost certainly look as if there is no statistically significant association.

What went wrong

Many studies are able to detect a clear statistical signal for climatological forcing of human conflict in Africa, using both aggregated data11 and highly disaggregated data similar to that in THB12—so why do THB run into a problem with low statistical power when these other studies do not?

As mentioned earlier, the probability of conflict in THB’s sample is 0.00016, a very small number. Large proportional changes in this number (such as a 100% increase in risk) lead to very small level changes in the probability of conflict, and very small level changes require extremely large samples to detect using statistical techniques. Earlier studies did not use data sets that exhibit such a low

---

11Burke et al., “Warming increases the risk of civil war in Africa”; Hsiang, Meng, and Cane, “Civil conflicts are associated with the global climate”; Hendrix and Salehyan, “Climate change, rainfall, and social conflict in Africa”.
probability of conflict, so large proportional changes in the risk of conflict correspond with sizable changes in the probability of conflict—changes that are large enough to be detected with statistics.

The likelihood of conflict in THB’s sample is much lower than other studies because of two of their modeling choices: (1) they study the onset of large-scale civil conflicts, which are relatively rare occurrences and (2) they dramatically increase the resolution of their analysis relative to prior studies but do not increase the number of conflict events in proportion, instead keeping the total number of conflict observations the same as prior studies and coding each conflict as occurring in a single $0.5^\circ \times 0.5^\circ$ (roughly $55 \times 55$ km $\approx 34 \times 34$ mile) pixel within a country with many pixels. This approach creates a large number of “non-conflict” observations that occur within countries that experience civil conflict, differing from previous studies that code an entire country (or many pixels in a country) as experiencing conflict when there is a conflict anywhere within that country. To retain the statistical power exhibited in earlier studies, THB would have needed to describe the spatial extent of conflict in their high-resolution framework such that the fraction of pixels exhibiting conflict is similar to the fraction of observations exhibiting conflict in prior studies, somewhat analogous to the approach of Levy et al. Other studies that successfully detected climatic effects on conflict at the subnational level in Africa do so by examining types of smaller scale conflicts, such as riots or inter-group conflicts, that occur much more frequently than large scale civil conflict. Thus, similar to THB, these subnational studies have many more observations in total than the national-level studies, but unlike THB, they also have many more conflict events and thus a much higher probability of conflicts than the 0.00016 in THB’s data.

To be clear why increasing resolution without a proportional increase in conflict counts causes a signal-to-noise issue for THB, consider the elements of Equation 1 above:

$$conflict\_risk = \frac{N_C}{N_U}(1 + \theta \times rainfall) + \epsilon$$

(3)

where $N_C$ is the total number of conflicts in the sample, $N_U$ is the total number of observational

---

13 Burke et al., “Warming increases the risk of civil war in Africa”.
14 Levy et al., “Freshwater availability anomalies and outbreak of internal war: Results from a global spatial time series analysis”; Burke et al., “Warming increases the risk of civil war in Africa”; Hsiang, Meng, and Cane, “Civil conflicts are associated with the global climate”; Hendrix and Salehyan, “Climate change, rainfall, and social conflict in Africa”.
15 Levy et al., “Freshwater availability anomalies and outbreak of internal war: Results from a global spatial time series analysis”.
17 With the exception of Levy et al.
units in the sample, and $\theta$ is the proportional change in conflict risk associated with a unit change in the centered variable rainfall. A meta-analysis of the literature suggests $\theta \approx 0.1$ per $1\sigma$ change in climate measures,\(^{18}\) such as rainfall. As THB increase the resolution of their data, they cut large observational units, such as the countries-by-year observations used in Burke et al. (2009), into a large number of small grid-squares. The effect is to increase $N_U$ to 363,811—contrast this value with the 889 observational units in Burke et al (2009). However, because THB hold the number of discreet conflict events fixed at the 59 events recorded in the UCDP/PRIO database, $N_C$ does not grow as THB increase their resolution. Thus, for a fixed $\theta$, increasing resolution as THB do drives down the signal $\frac{N_C}{N_U}\theta$ but does not alter the unexplained variation in risk $\epsilon$, causing the signal to noise ratio to fall. Had THB coded conflicts as spanning the larger regions over which they are actually observed, regions that include multiple grid-cells and are not constrained to the small $55 \times 55$ km ($34 \times 34$ mi) of an arbitrarily defined grid-cell, then $N_C$ would grow roughly in proportion to $N_U$ as the resolution of the analysis increased. This was the approach employed by Levy et al. (2005), which recovered results relatively similar to the rest of the literature.\(^{19}\)

There could be benefits of using a small-pixel approach in the study of civil conflict. For example, high resolution data containing much more localized information than national-level averages could provide new information about the climatological conditions in locations where conflicts begin. However, the data in THB are unlikely to accomplish this, as most of the “high-resolution” rainfall data in THB is not actually high-resolution information. Although the quality of rainfall data is not discussed in THB, the collection of rainfall data over Africa during 1960-2004 is inconsistent and scattered in space, as many governments lacked the resources or capacity to regularly collect and record weather data throughout their territory. Research groups which aggregate ground-level observations to create gridded climate data for Africa, such as the the Global Precipitation Climatology Centre (GPCC) that generated the data used in THB, therefore must interpolate the few real weather observations in Africa to estimate local historical weather conditions throughout much of the continent.\(^{21}\) The top panel of Figure 4 displays the extent of the actual rainfall observations available to the GPCC. Most pixels in Africa have no actual observations in the record, with only scattered coverage in most places other than densely populated regions of South Africa, Botswana, Namibia and West Africa—and in all of

\(^{18}\)Hsiang, Burke, and Miguel, “Quantifying the influence of climate on human conflict”.

\(^{19}\)Hsiang, Burke, and Miguel, “Quantifying the influence of climate on human conflict”.

Figure 4: The quality of rainfall data used by THB is limited in locations where THB code conflicts. Top: The fraction of months in THB’s sample when a pixel reports any rainfall observations. Missing data is interpolated.20 Each conflict is coded as occurring in a $0.5^\circ \times 0.5^\circ$ pixel at the center of a red circle. Bottom: The fraction of months in the THB sample when any rainfall data is recorded at a conflict location.
these locations, except South Africa, most of the weather records are incomplete, with long periods in which no observations are recorded. Despite this incompleteness, we might speculate that some high-resolution information could be useful if the specific locations in which conflicts are coded have good observational records. Yet, this is not the case. The locations of conflicts, as they are coded in THB, are the pixel in the center of each red circle in the top panel of Figure 4. Most conflicts tend to occur in remote locations that are not densely populated and are far from regions where governments have a strong capacity to enforce peace and stability. These are often the same locations where governments typically lack the capacity needed to set up and manage weather observatories. We demonstrate this point in the lower panel of Figure 4, where we plot the distribution of conflicts in the THB data according to the fraction of the sample for which there are complete rainfall records. We see that 44 out of the 59 conflicts (75%) occur in pixels that contain zero actual rainfall observations at any time in THB’s sample. Furthermore, no conflicts occur in locations that have a complete rainfall record. Thus, it seems that using a small-pixel approach is unlikely to provide new insights into the highly localized influence of climate on large-scale conflict in Africa, since virtually no new localized climate information is introduced for the locations of interest.

Conclusion

THB conclude that climate and conflict are unrelated in Africa, in disagreement with studies that find large associations at the local, national and continental scales. THB’s conclusion that no association exists is unwarranted and is a misinterpretation of their findings. Their analysis is statistically underpowered, so it cannot detect a reasonably—or even an unreasonably—strong association as statistically significant. In addition, examination of the rainfall observations underlying THB’s data reveals that a small-pixel approach in this sample is unlikely to yield new insights to the cause of civil conflicts in Africa because African governments tend not to record rainfall data in locations where civil conflicts erupt. By attempting a small-pixel analysis, THB sacrifice the statistical power that is

---

24Levy et al., “Freshwater availability anomalies and outbreak of internal war: Results from a global spatial time series analysis”; Burke et al., “Warming increases the risk of civil war in Africa”; Hendrix and Salehyan, “Climate change, rainfall, and social conflict in Africa”.
25Hsiang, Meng, and Cane, “Civil conflicts are associated with the global climate”.

---
critical to their findings being reliable, with little gain in location-specific climate information.

The fallacy in THB is that it does not follow logically that a failure to detect a signal implies that none exists. It only means one did not find it. Still, an accumulation of such failures may persuade some researchers that no such association exists.\textsuperscript{26} This consequence imposes a responsibility to ensure that every effort is made to do statistical analyses properly. This responsibility is heightened in a case like climate and conflict, where such a finding casts doubt upon the many other studies that have found a strong association.\textsuperscript{27} As a rule, a null result should always be accompanied by a power calculation to demonstrate that the null result is not spurious under reasonable assumptions—for example, the United States National Institutes of Health enforces this level of scientific discipline by requiring researchers to present such power calculations before a research project can be considered for funding.

The power calculation presented here demonstrates that THB’s approach is overwhelmingly likely to obtain statistically insignificant results even when a real association between climate and conflict exists. Thus, it is incorrect to interpret the findings of THB as a counter to the accumulating literature\textsuperscript{28} that does find a link between climate and conflict.

References


\textsuperscript{26}Hsiang and Meng, “Reconciling disagreement over climate-conflict results in Africa.”

\textsuperscript{27}Hsiang, Burke, and Miguel, “Reconciling Temperature-conflict Results in Kenya”; Levy et al., “Freshwater availability anomalies and outbreak of internal war: Results from a global spatial time series analysis”; Miguel, “Poverty and witch killing”; Burke et al., “Warming increases the risk of civil war in Africa”; Hsiang, Meng, and Cane, “Civil conflicts are associated with the global climate”; Harari and La Ferrara, “Conflict, Climate and Cells: A disaggregated analysis”; Hendrix and Salehyan, “Climate change, rainfall, and social conflict in Africa”; Fjelde and Uexkull, “Climate triggers: Rainfall anomalies, vulnerability and communal conflict in sub-Saharan Africa”; O’Loughlin et al., “Climate variability and conflict risk in East Africa, 1990–2009”.

\textsuperscript{28}Hsiang, Burke, and Miguel, “Quantifying the influence of climate on human conflict”.


