On Forecasting Conflict in Sudan: 2009-2012

Bessler, David and Kibriya, Shahriar and Chen, Junyi and Price, Ed

Texas AM University

August 2014

On Forecasting Conflict in the Sudan: 2009 – 2012

David A. Bessler, Shahriar Kibriya, Junyi Chen, and Ed Price[1]

August 2014

Abstract: The paper considers univariate and multivariate models to forecast monthly
conflict events in the Sudan over the out-of-sample period 2009 – 2012. The models used
to generate these forecasts were based on a specification from a machine learning
algorithm fit to 2000 – 2008 monthly data. The idea here is that for policy purposes we
need models that can forecast conflict events before they occur. The model that includes
previous month's wheat price performs better than a similar model which does not include
past wheat prices (the univariate model). Both models did not perform well in forecasting
conflict in a neighborhood of the 2012 "Heglig Crisis". Such a result is generic, as "outlier
or unusual events" are hard for models and policy experts to forecast.

---

[1] The authors are associated with the Conflict and Development Center at Texas A&M University

On Forecasting Conflict in the Sudan: 2009 – 2012

There is considerable evidence from around the world on a fundamental link between food insecurity and conflict. Food prices in particular appear to be particularly important focus in much of the literature – a result consistent with historical assessments offered in non-technical literature on food prices and conflict.[2] Hendrix and Brinkman (2013) review this evidence, much of which is observational, based on correlation or degree of "fit" between observations on conflict events and poverty (or some metric of food scarcity). Our purpose in this note is to consider the forecasting merit of a particular fit relationship: wheat prices and conflict in the Sudan over post 2000 data. As noted by Clive Granger (1980, page 348), "it is generally accepted that to find a model that apparently fits better than another is much easier than to find one that forecasts better." If "fit" relationships found in early periods do not carry-over to improved forecasting ability over later periods, the practical merits of the "fit" results are questionable.

This paper is organized in three sections. In the next section we offer a brief review of literature on forecasting of conflict. We follow with estimated (fit) models, both univariate and multivariate, of conflict, fatalities from conflict, and wheat prices on monthly observations from 2001 through 2008. These models are specified using a recently minted algorithm from the machine learning literature. In terms of "fit" relationships we show that wheat prices in period t (wheat price(t)) are not influenced by either conflict or fatalities numbers, but are best modeled as a first order autoregressive process (wheat price (t-1)). Conflict numbers in period t (conflict(t)) are modeled as first order autoregressions on themselves (conflict (t-1)) and wheat prices (t-1). Finally, fatalities from conflict are modeled as a function of current conflict numbers (conflict(t))

---

[2] See Carlyle (1849) and Lynn (2014) on bread prices and the French revolution and the *Economist* (2012) on food prices and the Arab Spring.

and two lag of fatalities from conflict (fatalities (t-1) and (t-2)). These "fit" results are generally in line with results reported in Chen, *et al*. (2014). Based on the models "fit" with 2001 – 2008 data we explore out-of-sample forecast performance of both univariate and multivariate forecasts of wheat price, conflict and fatalities from conflict over 2009 – 2012 monthly data. Both point (mean) forecasts and probability forecasts are studied.

This paper makes a substantive contribution in two areas: it is the first paper on probabilistic forecasts of conflict in the Sudan. By investigating the overall "goodness" of such forecasts we are able to assess the likelihood that certain periods of high or low conflict numbers were attributable to grain (in particular wheat) prices. Second, our paper demonstrates where improvement in knowledge and understanding of the causes of conflict are needed. While we do offer evidence that wheat prices are a mover of conflict numbers, we demonstrate, as well, time periods where wheat prices are clearly not the cause of conflict events.

What emerges from this paper is the need to develop models capable of offering credible forecasts of extreme events. In our case both models under consideration do a poor job of forecasting conflict numbers associated with the June 2012 "Heglig Crisis". While we are able to offer convincing evidence that the large jump in conflict numbers from May to June 2012 is not due to wheat price innovations or due to business as usual, we are not able to offer a model capable of offering credible "prevision" on its occurrence. While over our entire "out-of-sample" forecasting period our probabilities do pass a statistical test of credibility, they suffer from a form of over-confidence at particular critical periods. Our discussion offers a suggestion for overcoming this weakness (in our models) and perhaps weaknesses of others operating at the interface of conflict, data analysis, and human behavior.

Literature

The literature on forecasting conflict is relatively recent. We do not provide a thorough review here but rely on that offered in Brandt, Freeman and Schrodt (2011). They offer a set of five attributes that "should" characterize an ideal conflict forecasting model (tool):

1. Such a model should produce accurate forecasts;

2. It ought to produce near real-time forecasts;

3. The ideal model ought to model full dynamics (with feedback among belligerents);

4. Expert knowledge should be easily incorporated;

5. The model should be capable of making contingent (hypothetical "what if") forecasts.

Brandt, Freeman and Schrodt (2011) review several recent forays into the area. Three threads characterizing these models are: expected utility-game theoretic models; logistic regression; and Bayesian Vector Autoregressions. The expected utility-game theoretic models (Feder (1995) and Bueno de Mesquita (2010)), postulate preferences between belligerents and can solve for conflict probabilities. And, while they do allow for experts to help specify preferences and other model parameters, they do not produce real time forecasts. A second thread follows the "State Failure Project" and work done at the CIA. Papers by King and Zeng (2001) and Goldstone, et al. (2010) use logistic regression on highly aggregated data. While such models can also produce a probability of conflict, the aggregation level of data used makes real world application somewhat questionable (although it's not exactly clear what the CIA uses in its real time efforts). A third strand of literature reviewed by Brandt, Freeman and Schrodt (2011) is the structural Bayesian Vector Autoregression work described in Brandt and Freeman (2006) and extended in Brandt, Freeman and Schrodt (2011). This work models the dynamic evolution of conflict relevant

variables as a vector autoregression. Probabilistic forecasts are offered and evaluated. Expert opinion can be incorporated through an initial prior. If the data are measured with high frequency, and available in a timely manner, this model can produce near real time forecasts. Further, if experts (human or artificial) can provide contemporaneous structure among innovations, reliable conditional (what if type analysis) probability forecasts are capable of generation.

Meeting Brandt, Freeman, and Schrodt's (2011) first attribute listed above begs the question: "accurate compared to what?" In another forecasting setting Casillas-Olvera and Bessler (2006) accuracy was measured in terms of performance of the Bank of England's forecasters relative to a set of "other forecasters", who regularly issued subjective (expert opinion) forecasts on inflation rates and GDP growth. Here we are not aware of a set of alternative forecasters to serve as a benchmark from which to judge the forecasts forthcoming from our data-based models. Accordingly, below we compare our point (mean) and probability forecasts from a multivariate model of conflict (conditioning on past wheat prices) against a model that does not condition on previous wheat prices.

Chen, et al. (2014) provide "fit" evidence that grain prices in the Sudan Granger-cause conflict numbers in the Sudan based on monthly data. Using a model similar in its dynamic structure as that of Brandt, Freeman and Schrodt (2011), Chen et al. (2014) show that wheat and grain sorghum prices explain a considerable proportion of the (within sample) forecast error variance of conflict numbers in the Sudan using monthly data observed over 2001 – 2012.

In undivided Sudan wheat was primarily consumed in the Northern part (Abdelrahman, Ali Hasab 1998) especially by the urban dwellers. Following the trend of other East African countries wheat has become the staple food of non-subsistence households in Sudan during the

last decade (Mkumbwa 2011, Mostafa et al., 2013). Rapid urbanization in Sudan shifted consumer preference from sorghum to processed bread made from wheat. Although the production of wheat has increased marginally Mostafa et al., (2013) report that the average consumption of wheat has increased to 1770 thousand tons in the 2000's from 743 thousand tons in the 1980s. The lack production of wheat in Sudan implies larger imports, higher price volatility and government intervention.

Data and Models

The data used in the current paper are monthly wheat prices, conflict numbers and fatalities from conflict numbers from the Sudan for the period 2001 – 2012. The nominal price data per ton of wheat reflect the monthly wholesale prices in Khartoum port. The source of the dataset is Global Information and Early Warning System (GIEWS) Food Price Data and Analysis Tool, a real time data steaming source of Food and Agricultural Organization (FAO) of the United Nations. While Chen et al. (2014) use data on both North and South Sudan after the 2010 separation of the two countries; we use only data on North Sudan, as Chen et al. (2014) find a break point in the combined data from both countries in 2011, in a neighborhood of the division of the countries. Plots are given in figure 1.

The data for the number of conflict events are obtained from Armed Conflict Location & Event Dataset (ACLED) (Raleigh, Linke, Hegre, and Carlsen 2010) over the same period. The ACLED dataset provides time specific information on recorded and reported battles, killings, riots, and recruitment activities of rebels, governments, militias, armed groups, protesters and civilians. And finally, Chen et al. (2014) use the procedure presented in Hsiao (1979) to specify a subset

vector autoregression; we follow the causal search literature, using PC algorithm (Spirtes, Glymour and Scheines (2000)) for model specification.

The graphical pattern among current and lagged wheat prices, conflict numbers and fatalities from conflict numbers found with PC algorithm[3] is given in Figure 2. Notice first that there are several contemporaneous relations identified in the figure. Wheat price (T-2) causes Conflict numbers (T-2). Fatalities (T-2), (T-1), and (T) and Conflict (T-2), (T-1), and (T) have contemporaneous relationships (the bi-directed arrow between conflict and fatalities indicate the presence of a latent variable that may move both conflict and fatalities in contemporaneous time). As these contemporaneous relationships cannot be exploited for forecasting purposes, we do not explore them further here. In terms of lagged (time delayed) relationships we see wheat prices (W) are generated as a first order (univariate) autoregression. Conflict (C) numbers are generated by one lag of Wheat prices (W (T-1)) and one lag of Conflict (C (T-1)) numbers. Fatalities from conflicts (F(T)) are generated by two lag of Fatalities (F(T-1), F(T-2)). The models fit[4] over 2001 – 2008 monthly data are given as equations (1), (2) and (3) below.

(1) $W(T) = 6.31 + .92\ W(T-1);\quad R^2 = .83;\ Q_{df=23} = 30.4,\ \text{P-value} = .14$

$\qquad\qquad (3.58)\quad (.06)$

---

(2) $C(T) = -5.65 + .17\,W(T-1) + .30\,C(T-1)$; $R^2 = .47$; $Q_{df=23} = 18.5$, P-value = .73

$\quad\quad\quad\quad (2.40) \quad (.04) \quad\quad\quad\quad (.11)$

(3) $F(T) = 94.13 + .32\,F(T-1) + .17\,F(T-2)$; $R^2 = .17$; $Q_{df=23} = 16.8$, P-value = .81

$\quad\quad\quad\quad (38.13) \quad (.30) \quad\quad\quad\quad (.08)$

Equation (1) indicates that wheat prices (W) are fit with a modestly high coefficient of determination (.83), with no serious residual autocorrelation (as indicated by our failure to reject the null hypothesis of white noise residuals at usual (.10 or .05) p-values). The degree of fit on the conflict equation (2) is much weaker ($R^2 = .47$) than that for wheat price. Interestingly wheat prices lagged one period show a larger simple t-statistic (t= .17/.04) than that associated with the estimated coefficient on conflict lagged one period in the second equation (t=.30/.11). No strong evidence of autocorrelation amongst residuals is seen in equation number 2, as well. Finally, equation (3) has even poorer fit (than equation (2)) as R^2 is just .17. It is the second lag of fatalities which shows the largest t-ratio (t=.17/.08) in this third equation. Again, no evidence of autocorrelation in residuals is noted (p-value on Q is .81).

Point Forecasts

Recursive forecasts of wheat price, conflict numbers and fatalities from conflict are generated from these equations (as recursively re-estimated) for the monthly data 2009 – 2012. Forecasting "goodness" results for point forecasts (mean of the under-lying probability distribution at each date) for one step, two-steps and twelve-step ahead horizons are given in table 1. These numbers are provided for both the Vector Autoregression (VAR) model summarized as equation (1), (2) and (3), as well as for the univariate forecasts (UNIV). As equation (1) and (3) are univariate, this comparison model is only different for equation (2), where lagged values of wheat price in the

conflict equation are dropped. Accordingly, the results reported in Table 1 under the row labels

Vector Autoregression and *Univariate* are the same for *wheat price* and *fatalities*. The entries

differ on *conflict* entries.

In terms of Mean Squared Error (MSE) measures we see the usual relationship between

forecast performance and horizon (steps ahead). As we look further out for all three variables our

forecast performance diminishes (as measured by higher MSE values). Note as well that MSE

measures for conflict forecasts associated with the Vector Autoregression are lower than

corresponding measures associated with the univariate model at all horizons (steps ahead). This

latter result is consistent with Granger's notion of causality (Granger 1980): Wheat Price causes

Conflict numbers. As a footnote to table 1, we report mean absolute percentage error at the one-

step-ahead horizon. These metrics indicate that wheat prices are much easier to forecast (MAPE

= .07), relative to conflict numbers (MAPE=.92) and fatalities (MAPE=13.57) – a result

consistent with the relative magnitudes of the R-squareds reported on each equation (1), (2) and

(3) above.

While the root mean squared error metrics presented in table 1 indicate that forecasts of

conflict numbers based on models that include past wheat prices in the model are "better" than

forecasts of conflict numbers that do not include past wheat prices (MSE's are lower for the

former at all horizons), there may not be a significant difference between the measures compared.

In table 2 we present forecast encompassing results of the hypothesis that forecasts from the

vector autoregression encompass forecasts from the univariate model and vice versa.

Encompassing tests address the question: can we dispense with the encompassed model and

focus decision-making on the encompassing model? We offer results at all three horizons: do

forecasts from the VAR encompass forecasts from univariate model? Do forecasts from the

univariate model encompass forecasts from the VAR model?   Each row of table 2 offers a test of

the encompassing hypothesis (see the footnote to the table for the specification of the test). We

see that it is only at one-step ahead that we do not reject the encompassing hypothesis: forecasts

from the VAR encompass forecasts from the univariate model at one-step ahead. At all other

horizons the forecasts of each model do not encompass those of the other. This provides evidence

that the "wheat price causes conflict numbers" is a short-run phenomena (consistent with results

reported in Chen et al. (2014) and in the graphical search results in figure 2).

Continuing with the study of the relative merits of the two models (VAR and Univariate)

for forecasting conflict numbers we look at the d-separation (conditional independence) relation

among forecasts from the univariate model, the VAR model and the actual realization. We do

this for just the one-step horizon (given the outcomes of the encompassing test just reported).

The d-separation[5] results look for "blockage of information flow" amongst variables. Figure 3

shows that forecasts of conflict numbers emanating from the univariate model are "blocked" in

their path to the actual realization of conflict numbers by forecasts from the VAR model –

suggesting that one need not condition his/her decision-making with respect to conflict numbers

on the univariate model. The VAR forecasts capture all the relevant information (that's available

from these two forecasting models) about next month's actual conflict numbers.

---

[5] D-separation (directional separation) is due to Judea Pearl (2000 page 16) and is key to sorting-out what needs to be conditioned on and what need not be conditioned on in decision-making environments. Bessler and Wang (2011) explore it use in forecast evaluation for agricultural prices and US GDP and inflation rates.

Probability Forecasts

Of course, no point forecast can capture the extent of uncertainty of a future realization, whether such a forecast comes from a large-scale (many equations) model, a simple autoregression or from a human expert. Below we offer comment on probability forecasts from the univariate autoregression and the vector autoregression, whose point forecasts were studied above.

The literature on probability forecasts from models begins with Dawid (1984), although an earlier and certainly helpful literature on probability forecasting in psychology pre-dates Dawid's efforts by several years.[6] Here we follow Kling and Bessler (1989) and bootstrap forecasts emanating from equations (1), (2), and (3) and the univariate form on equation (2). Uncertainty in errors on each equation and parameter uncertainty of each estimated parameter is accommodated by drawing on observed (historical) error terms in 1000 simulations at each out-of-sample data point. This gives us 1000 forecasts of each variable, wheat price, conflict numbers and fatalities, at each date, January 2009 through December 2012. Again, as in our point forecasting results, the model parameters are up-dated sequentially, by re-estimating at every month, as it moves through the out-of-sample data, January 2009 through December 2012.

Lichtenstein, et al. (1982) suggests calibration as an indicator of "goodness" of probability forecasts. So for example, if our model (or our mind) issues a probability of p=.25 100 times we should expect, after the fact, that 25 of these 100 outcomes should be realized (be "true"). Dawid (1984) builds on this calibration idea and the probability integral transform to offer a test of calibration through the associated cumulative density function (CDF). That is to say, if our issued probability are placed into their CDF form ($\int_{-\infty}^{x} P(x)dx$), where *P(x)* is the probability

---

[6] See, for example, Lichtenstein, Fischhoff and Phillips (1982, page 306-334).

distribution[7] on conflict being *x*, generated from the 1000 draws simulation described above, then

the realized fractile (fractile on event that actually occurs, after the fact), will be uniformly

distributed if the issued distribution is well-calibrated. A chi-squared test of well-calibration is

offered in Dawid; its application to our conflict study is presented below.

Calibration plots on conflict forecasts are presented in figure 4. A calibration plot on well-

calibrated forecasts would show the realized fractiles fall on a 45 degree line, in a plot of realized

fractile against the relative frequency (Bunn 1982). Forecasts from both the univariate and VAR

models of conflict have realized fractiles not on the 45 degree line, but generally above it.

Visually, the univariate plots are somewhat closer to the line, relative to the VAR plots. A formal

test of well-calibration on both sets of forecasts is given in table 3. Here we do not reject the null

hypothesis of well-calibration for both the univariate and the VAR forecasts of conflict at a 5%

level of significance. Similarly we do not reject the null of well-calibration on the forecasts on

wheat price and fatalities from conflict (table 3).

Given the decision reached in table 2, that our forecasts pass a Chi-squared test of well-

calibration, we investigate the issued probabilities on conflict for particularly interesting data

realizations. First, we note that June 2009 was a particularly good month in that no conflicts

occurred in the data set. The previous months of April and May saw 6 and 7 conflict numbers

reported. The issued probabilities for June 2009 are given in histogram form in figure 5. Both the

VAR and the Univariate model issue probabilities that contain the actual realization in their

modal class ($0 \le C_{\text{June 2009}} < 10$); where $C_{\text{June 2009}}$ is the conflict number in June of 2009.[8] The VAR

---

[7] Actually we use summation operations (not integrals) as our work, below, involves discrete probabilities, not continuous.
[8] Defining classes or bins for aggregating probabilities is somewhat arbitrary. There is an extensive literature that attempts
to decide on the number of probability bins from historical data (see, for example, Scott (1979) and Wand (1997)). Here
we argue that the proper number and size of each bin (the size does not have to be the same across all bins) is "best"

gave slightly preferred density forecasts, in the sense that its issued probability for this bin-class

was .595; the univariate model issued its probability on this class as .588.

A case where both models show less or no forecasting quality is found for the date June

2012. This date corresponds to the largest number of conflict numbers in our data set (153

conflicts). Figure 6 shows the probability histograms for both models are centered on bin 3 =

$P(20 \leq C_T < 30)$, with neither model issuing non-zero probabilities for the realized event (conflict for June

2012 =153). This date corresponds to the "Heglig Crisis"- an oil related set of conflict events

between Sudan and South Sudan. There is no hint of this extreme number in our data. Conflict

numbers for the preceding two months (May and April) are reported as 38 and 19, respectively.

Further wheat price was within its recent historical range at 170 for May 2012 and 165 for April

2012, both down from the high of $181/ton for December 2011. Clearly both models did not offer

credible forecasts of conflict for June 2012.

Another date which may be of interest is December 2011. This date corresponds to the

highest wheat price observation in our post-fit, forecasting, period of study. Here wheat price (as

indicated above) was $181/ton. Figure 7 shows the associated histograms on conflict numbers for

this month. Recall the VAR model has wheat price in its information set; while the univariate

model does not. The actual number of conflicts for December 2011 was 25. Conflict numbers for

the preceding two months were 12 and 11, respectively. Both models issue their modal class as

---

defined by subject matter experts (most likely in the field) who must deal with consequences from conflict events. If we follow the statistics literature (and there is not universal agreement on selection of bin width) we find bins in the neighborhood of 25 conflict numbers (using the result derive in Wand (1997), h = 3.49 $(\sigma)/N^{1/3}$, where h is bin width, σ is the standard deviation of the data (prior to our forecasting interval) and N refers to the number of observations prior to our forecasting interval. Our subjective judgment is that bin width should be a bit narrower at low levels of conflict and perhaps higher at higher levels of conflict. Needless to say this is an area for further research.   We use the outcome bins in Figures 5, 6 and 7 defined as follows: bin 1=$P(0 \leq C_T < 10)$; bin 2= $P(10 \leq C_T < 20)$; bin 3 = $P(20 \leq C_T < 30)$; bin 4 = $P(30 \leq C_T < 40)$; bin 5 = $P(40 \leq C_T < 80)$; and bin 6 = $P(80 \leq C_T < \infty)$; where C_T is conflict at time T.

bin 2= $P(10 \leq C_{\text{December 2011}} < 20)$; however the VAR model places more density in this bin class

relative to the univariate model (.141>.096). The latter places more density in bin 1 and less in bin

3 (the realized bin) relative to the VAR model.

Discussion

Here we have explored the ability of two models in forecasting conflict numbers in the Sudan.

An earlier paper found a "fit" relation between wheat prices and conflict numbers in the Sudan.

We use similar data and explore the out-of-sample forecasting ability of similar models over 2009

– 2012 monthly data on wheat prices, conflict numbers and fatalities in North Sudan. Both point

(mean) and probability forecasts are studied from two candidate models: a univariate model,

which expresses conflict numbers at time t as a function of conflict numbers at time t-1 and a

vector autoregression model which expresses conflict numbers a time t as dependent on conflict

numbers in t-1 and wheat prices in time t-1. Point forecasting results indicate the VAR model

preforms better in terms of root mean squared forecast errors and forecast encompassing. Both

models offer well-calibrated probability forecasts over our post fit period. Clearly, knowledge of

wheat prices in month t is helpful in forecasting conflict numbers for month t+1; a result that

coheres well with 'fit' results found on earlier data via a machine learning algorithm.

While we do not reject well-calibration for both models' probability forecasts, both do not

perform well in their forecasts conflict for June 2012. This month corresponds to the "Heglig

Crisis"- an oil related set of conflict events between North and South Sudan. There is no

information in either the history of conflict numbers or wheat prices to hint of the large number of

conflicts that occurred at this date. Both models issue probability forecasts of zero on conflict

events in a neighborhood of the number actually observed. This exposes clearly the need to rely

14

on more than formal models on historically relevant data in forecasting conflict (or in terms of forecasting human subject outcomes, in general). In our literature review we cite the need to bring experts into the forecasting exercise, as suggested by Brandt, Freeman, and Schrodt (2011).

Some information variables are measured and reported regularly – market prices are one such variable. And, as food is of particular importance for human existence, its price can offer us prevision of up-coming conflict. Other information is not reported regularly, as it occurs in a seemingly unsystematic pattern. The "Heglig Crisis" which erupted in extreme conflict in June of 2012 was clearly not a wheat price caused event. Most discussions cite oil as the root cause (see the UN report cited below). Here subjective judgment may well have improved our model's forecasts. In May 2012 the "Security Council Report" from the United Nations discusses this crisis. This discussion indicates clearly the existence of evidence that a rise in conflict was eminent. The following quote of the UN report supports that such evidence was communicated to both parties of the potential (and finally realized) conflict: "A key issue is whether and how the Council can exert sufficient leverage on the parties to deter them from expanding their conflict, induce them to cease fighting, and convince them to return in good faith to the negotiating table. Since February, the Council has produced two press statements and two presidential statements regarding the situation in Sudan and South Sudan with what appears to be minimal impact on the calculations of the parties."

Even if we can forecast increases in conflict, can we find some way to prevent or mitigate the possibility of such events coming to pass? Such efforts at mitigation would, in turn, require more elaborate forecasting models, as we would now have to offer our forecasts on the

probability of conflict, given a particular mitigation strategy has been undertaken. Much of such

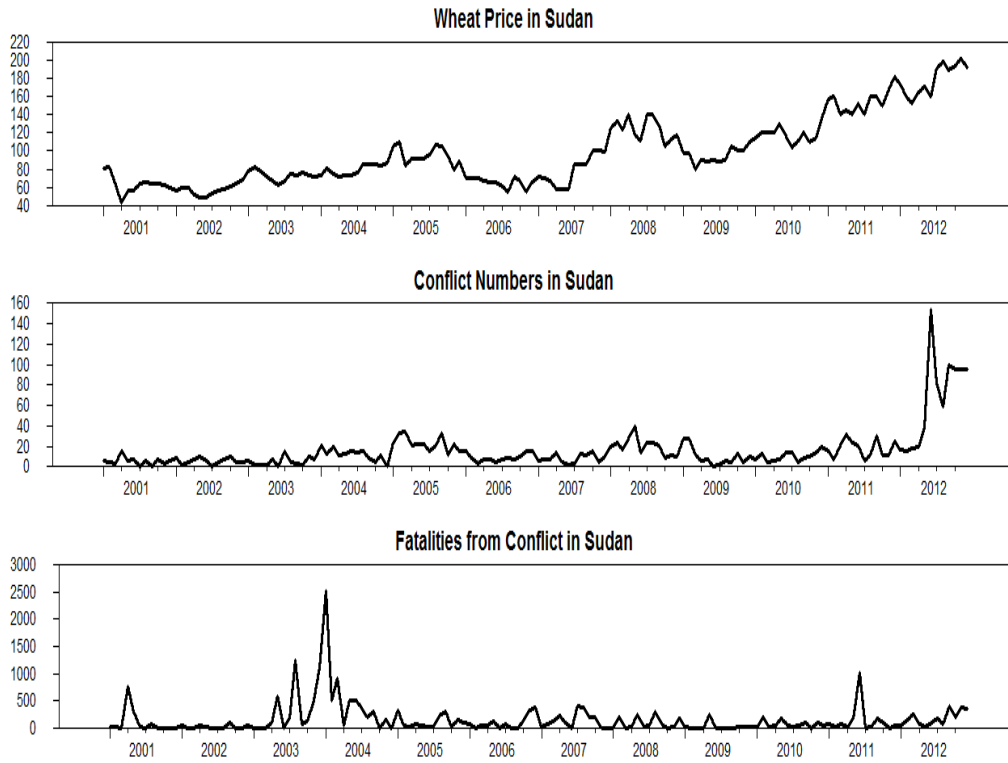conditioning information will likely be subjective, varying from episode to episode.

Figure 1. Time Series Plots of Wheat Price, Conflict Numbers, and Fatalities from Conflict in Sudan, 2001 – 2012, monthly data.
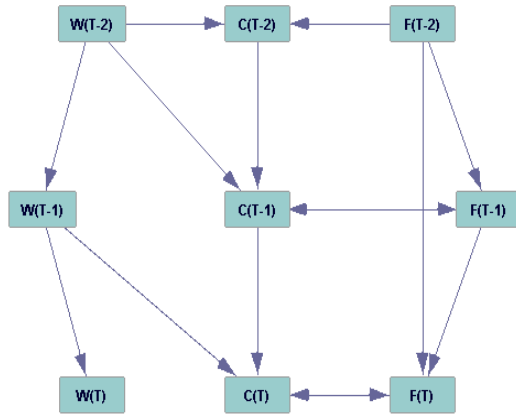
Figure 2. Graphical Pattern Found with PC Algorithm.[9]

---

[9] Variable symbols are as follows: Wheat Price in period T is represented as W(T), Conflict in period T as C(T) and Fatalities from Conflict in period T as F(T). Lags of each of the above are indicated using T-1 and T-2 for one and two lags, respectively. Data used to fit the graph given here are monthly on Wheat Prices, Conflict Numbers and Fatalities from Conflict in the Sudan 2001 – 2008. The p-value used in the PC Algorithm search was 10%. Details on the Algorithm are given in Spirtes, Glymour and Scheines (2000).

Figure 3. Graphical Pattern on Forecasts of Conflict and Actual Conflict Numbers in the Sudan from Monthly data 2009 – 2012.[10]
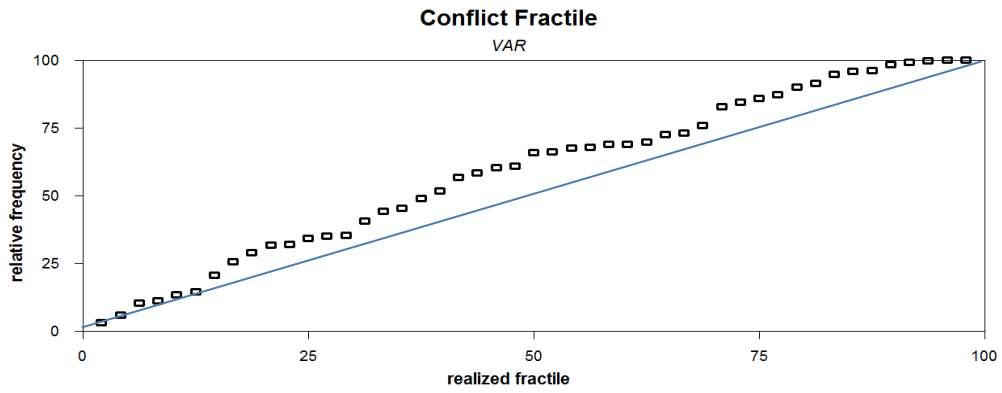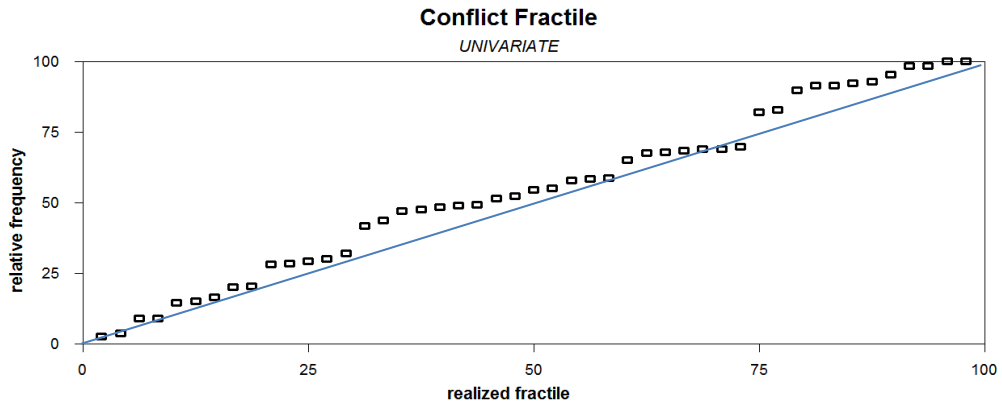
Figure 4. Realized Fractiles and Relative Frequencies on Conflict Events from the Univariate Model and VAR Model Probability Forecasts: Sudan, January 2009 – December 2012.
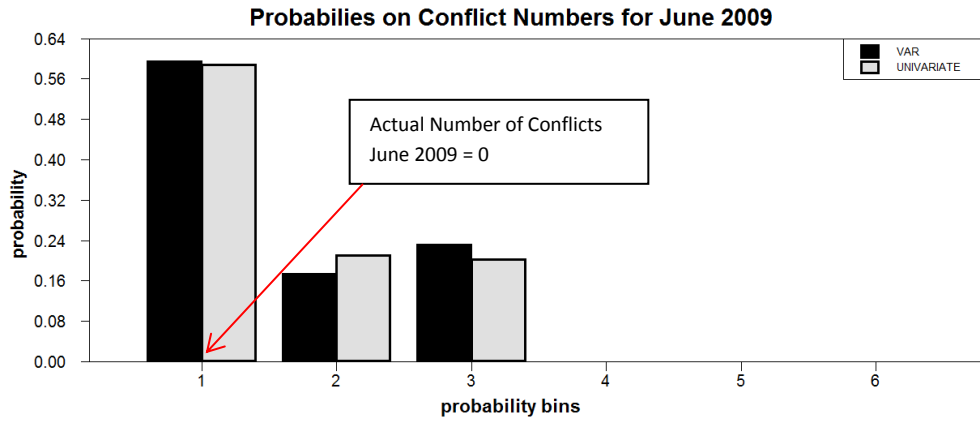
Figure 5. Probabilities Issued on Conflict Numbers from Univariate and VAR Model for June 2009.[11]

---

[11] The probability bins in Figure 5 are defined as follows: bin $1 = P(0 \leq C_{June\ 2009} < 10)$; bin $2 = P(10 \leq C_{June\ 2009} < 20)$; bin 3 = $P(20 \leq C_{June\ 2009} < 30)$; bin 4 = $P(30 \leq C_{June\ 2009} < 40)$; bin 5 = $P(40 \leq C_{June\ 2009} < 80)$; and bin 6 = $P(80 \leq C_{June\ 2009} < \infty)$.

Figure 6. Probabilities Issued on Conflict Numbers from Univariate and VAR Model for June 2012.[12]

---

[12] The probability bins in Figure 6 are defined as follows: bin 1=P($0 \leq C_{\text{June 2012}} < 10$); bin 2= P($10 \leq C_{\text{June 2012}} < 20$); bin 3 = P($20 \leq C_{\text{June 2012}} < 30$); bin 4 = P($30 \leq C_{\text{June 2012}} < 40$); bin 5 = P($40 \leq C_{\text{June 2012}} < 80$); and bin 6 = P($80 \leq C_{\text{June 2012}} < \infty$).
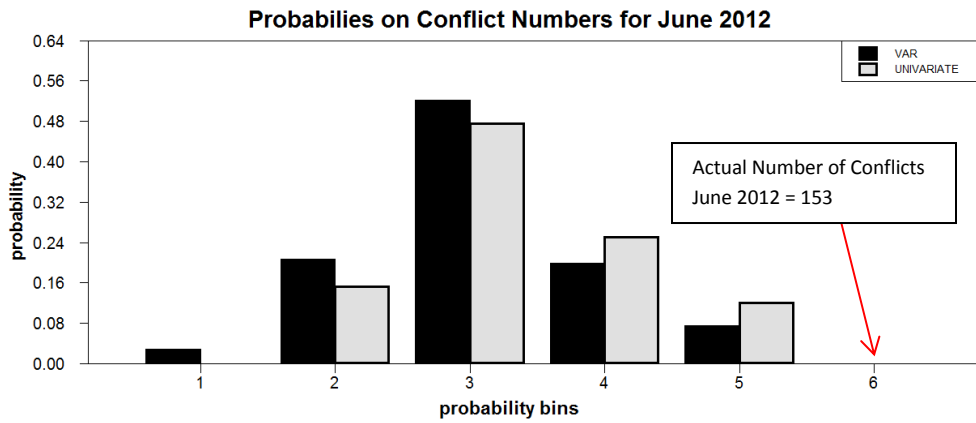
Probabilies on Conflict Numbers for December 2011
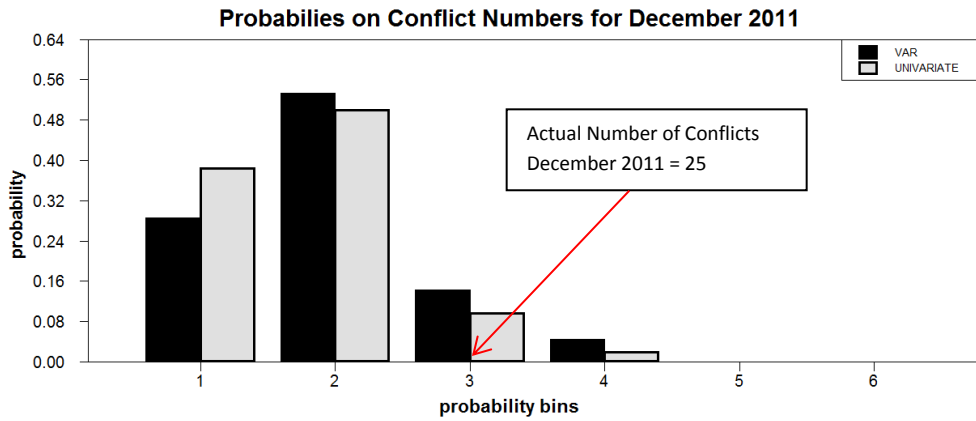
Figure 7. Probabilities Issued on Conflict Numbers from Univariate and VAR Model for December 2011.[1]

---

[1] The probability bins in Figure 7 are defined as follows: bin 1=P($0 \leq C_{\text{December 2011}} < 10$); bin 2= P($10 \leq C_{\text{December2011}} < 20$); bin 3 = P($20 \leq C_{\text{December 2011}} < 30$); bin 4 = P($30 \leq C_{\text{December 2011}} < 40$); bin 5 = P($40 \leq C_{\text{December 2011}} < 80$); and bin 6 = P($80 \leq C_{\text{December 2011}} < \infty$).

Table 1. Forecast Error Metrics on one-step-ahead forecasts of Wheat Prices, Conflict, and Fatalities from Conflict in the Sudan 2009 – 2012, Monthly Data.

Series/*Model*/Steps Ahead	RMSE[1]	Theil U[2]	Number of Forecasts
Wheat			
Univariate			
1-step	11.89	1.03	48
2-steps	16.73	1.07	47
12-steps	45.80	1.49	37
VAR			
1-step	11.89	1.03	48
2-steps	16.73	1.07	47
12-steps	45.80	1.49	37
Conflict			
Univariate			
1-step	23.39	1.05	48
2-steps	29.51	1.08	47
12-steps	39.69	1.05	37
VAR			
1-step	21.55	.97	48
2-steps	25.81	.95	47
12-steps	35.92	.95	37
Fatalities			
Univariate			
1-step	176.82	.80	48
2-steps	179.30	.76	47
12-steps	185.30	.74	37
VAR			
1-step	176.82	.80	48
2-steps	179.30	.76	47
12-steps	185.30	.74	37

---

[1] Root Mean squared error (RMSE) is calculated as: RMSE $=(\sum_1^K (A(t) - F(t))^2/K)^{.5}$, where $A(t)$ is the actual value, $F(t)$ is the forecasted value from either the univariate or VAR model, K is the number of out of sample forecasts studied.

[2] Theil U compares the mean squared error from the model (univariate or VAR) with that emanating from a random walk forecast (last period actual is the forecast of next period's actual). Theil U < 1.0 are preferred to Theil U's >1.0. Mean Absolute Per Cent Error (MAPE) are given at one step ahead as: .07 (wheat price), .92 (conflict numbers and 13.57 (fatalities). Numbers at higher levels are increasing and reflect the ordering:
$MAPE_{wheat} < MAPE_{conflict} < MAPE_{fatalities}$.

Table 2. Encompassing Tests on One-, Two-, and Three-Steps Ahead Forecasts of Conflict[1]

Step Ahead	Null Hypothesis	$\hat{\lambda}$	p-value	Decision
1	VAR Encompasses UNIV	-.64	.11	Fail to Reject
1	UNIV Encompasses VAR	1.64	.00	Reject
2	VAR Encompasses UNIV	-.80	.02	Reject
2	UNIV Encompasses VAR	1.80	.00	Reject
12	VAR Encompasses UNIV	-4.49	.00	Reject
12	UNIV Encompasses VAR	5.49	.00	Reject

---

[1] The entries in the table refer to estimated value of lambda ($\hat{\lambda}$) from an ordinary least squares regression:

$$e_{1t} = \lambda(e_{1t} - e_{2t}) + \varepsilon_t ,$$

where e_{1t} refers to forecast errors from the encompassing model, e_{2t} refers to forecast errors from the encompassed model, both for out-of-sample data for period t and ε_t is a white noise disturbance from the encompassing regression. Following Harvey and Newbold (2000) all regression models are estimated with White's "robusterrors" procedure to account for possible heteroskedasticity in errors.   We reject the null hypothesis (stated in the second column from the left for low p-values; i.e. values lower than .05. All estimation was carried-out using RATS (Doan 2009).

Table 3. Chi-Squared Statistics on Calibration of Probability Forecasts of Wheat Prices, Conflict and Fatalities from Conflict in the Sudan 2009 – 2012, Monthly Data.

Series/Model	Calculated[1] χ^2	Critical $\chi^2_{(.05,df=9)}$	Decision with Respect to the Null Hypothesis
Wheat			
Univariate[2]	11.58	16.9	Fail to Reject
VAR[3]	11.58	16.9	Fail to Reject
Conflict			
Univariate[2]	12.83	16.9	Fail to Reject
Vector Autoregression[3]	14.91	16.9	Fail to Reject
Fatalities			
Univariate[2]	8.3	16.9	Fail to Reject
Vector Autoregression[3]	8.3	16.9	Fail to Reject

1. The null hypothesis in every case in the table is that the issued probabilities are well-calibrated. The χ^2 statistics are calculated as follows: $\chi^2 = \sum_{k=1}^{10}(m_i - 4.8)^2/4.8$; where m_i are the number of realized fractiles in category i, where categories are defined as $c_1 = (0,.099)$, $c_2 = (.100, .199)$, …, $c_{10} = (.900, .999)$.

2. The univariate forecast are generated from the fit model $Y_t = \beta_o + \beta_1 Y_{t-1} + … + \beta_k Y_{t-k} + \varepsilon_t$; where Y_t is a scalar holding either observations at time t on conflict or fatalities, β_o, β_1, … β_k are scalars, recursively up-dated at each data point; 1000 draws on the sampling distributions on $\beta_{o,i}$, $\beta_{1,i}$, …, $\beta_{k,i}$ and ε_{t+1} are used to form 1000 forecasts, $\hat{Y}_{t+1,i} = \beta_{o,i} + \beta_{1,i} Y_t + … + \beta_{k,i} Y_{t-k+1} + \varepsilon_{t+1,i}$; i = 1, …, 1000.

3. The vector autoregressive forecasts are generated as: $Y_t = \beta_o + \beta_1 Y_{t-1} + … + \beta_k Y_{t-k} + \varepsilon_{t+1}$; where : Y_t is a vector holding wheat price, sorghum price, conflict and fatalities at time t, β_o is a vector of constants, one for each series, and β_1, …, β_k are matrices holding estimated coefficients, relating Y_t to lags Y_{t-1}, … , Y_{t-k}. The estimated betas are recursively up-dated at each data point; 1000 draws on the sampling distributions on $\beta_{o,i}$, $\beta_{1,i}$, …, $\beta_{k,i}$ and ε_{t+1} are used to form 1000 forecasts, $\hat{Y}_{t+1,i} = \beta_{o,i} + \beta_{1,i} Y_t + … + \beta_{k,i} Y_{t-k+1} + \varepsilon_{t+1,i}$; i = 1, …, 1000.

References

Abdelrahman, Ali Hasab-Elrasoul. (1998) *Trends in Sudanese Cereal Production, Consumption, and Trade*. Center for Agricultural and Rural Development, Iowa State University.

Bessler, D.A. and Z. Wang, (2012) "D-Separation, Forecasting, and Economic Science: A Conjecture," *Theory and Decision*, 295-314.

Brandt, P.T., J.R. Freeman, and P.A. Schrodt (2011) "Real Time, Time Series Forecasting of Inter-and Intra-State Political Conflict," *Conflict Management and Peace Science* 28:41-64.

Bueno, de Mesquita (2002) Predicting Policies, Columbus: The Ohio State University Press.

Bunn, D. *Applied Decision Analysis*, McGraw-Hill 1984.

Carlyle, Thomas, 1837. *The French Revolution: A History*, New York Modern Library (2002).

Casillas-Olvera, G. and D.A. Bessler. 2006. "Probability Forecasting and Central Bank Accountability," *Journal of Policy Modeling* 28:223-234.

Raleigh, Clionadh, Andrew Linke, Havard Hegre and Joackim Karlen (2010). Introducing ACLED: An Armed Conflict Location and Event Dataset. *Journal of Peace Research*, 47, 5:1-10.

Chen, J., S. Kybriya. E. Price and D.A. Bessler (2014) "A Causal Exploration of Conflict and Commodity Prices of Undivided Sudan". Unpublished working paper, Conflict and Development Center at Texas A&M University.

Dawid, A.P., "Statistical theory: A prequential approach," *Journal of Royal Statistical Society*, 147(1984): 278-297.

Doan, T. *RATS*, version 7, Evanston: Estima, 2009.

Feder, S. (1995) "Faction and Policon: New Ways of Analyze Politics, Editor: H. Bradford Westerfeld, *Inside the CIA's Private World: Declassified Articles from Agency's International Journal 1955-1992*, New Haven: Yale University Press.

Granger, C.W.J. (1980) "Testing for causality: A personal viewpoint," *Journal of Economic Dynamics and Control* 2:329-352.

Goldstone, J.A., R.H. Bates, D.L. Epstein, T.R. Gurr, M.B. Lustik, M.G. Marshall. J. Ulfelder and M. Woodward (2010) "A Global Model of Forecasting Political Instability," *American Journal of Political Science* 54:190-208.

Harvey, D. and P. Newbold 2000. "Tests for Multiple Forecast Encompassing," *Journal of Applied Econometrics* 15:471-482.

Hendrix, C and Brinkman, H 2013. Food Insecurity and Conflict Dynamics: Causal Linkages and Complex Feedbacks. *Stability: International Journal of Security and Development* 2(2):26, DOI: http://dx.doi.org/10.5334/sta.bm accessed August 2014.

Hsiao, C. (1979) "Autoregressive Modeling of Canadian Money and Income Data," *Journal of the American Statistical Association* 74:553-560.

Kling, J. and D.A. Bessler. 1989. "Calibration-Based Predictive Distributions: an Application of Prequential Analysis to Interest Rates, Money, Prices and Output." *Journal of Business* 62(4):447-500.

Lichtenstein, S., B. Fischhoff and L. Phillips, (1982) "Calibration of Probabilities: The State of the Art to 1980" in Editors: D. Kahneman, P. Slovic and A. Tversky, *Judgment under Uncertainty: Heuristics and Biases,* Cambridge University Press.

Lynn, M.R. "Riots and Rye: Bread and the French Revolution," www.ultimatehistoryproject.com/breadhtml   accessed August 15, 2014.

Mkumbwa, Solomon S. (2011)."Cereal food commodities in Eastern Africa: consumption-production gap trends and projections for 2020." MPRA paper number 42113, posted 22. October 2012. Online at http:// mpra.ub.uni-muenchen.de/42113.

Mostafa, Raja. Elagali Mohamed and Abulgasim, Shazy (2013). Trends in Wheat Production and Consumption in Sudan. *International Journal of Research in Management, Economics and Commerce*. Volume 3. Issue 4.

Pearl, J. (2000). *Causality: Models, reasoning, and inference.* Cambridge, MA: Cambridge University Press.

Scott, D.W. (1979) "On optimal and data-based histograms," *Biometrika* 66:605-610.

Spirtes, P., Glymour, C., and Scheines, R. 2000. *Causality, prediction, and search.* Cambridge, MA: MIT press.

The Economist "Food and the Arab Spring: Let them eat baklava*,"* January 2012, http://www.economist.com/node/21550328 ,accessed August 13, 2014.

United Nations (2012) "Security Council Report of the UN," New York, NY, http://www.securitycouncilreport.org/monthly-forecast/2012-05/lookup_c_glKWLeMTIsG_b_8075191.php , accessed August 13, 2014.

Wand, M.P. (1997) "Data-based choice of histogram bin width," *The American Statistician* 51:59-64.