ConvLSTMs for High Resolution Conflict Forecasts

ViEWS Prediction Competition Entry

Benjamin J. Radford

April 2, 2022

University of North Carolina at Charlotte

This talk

Preprint PDF

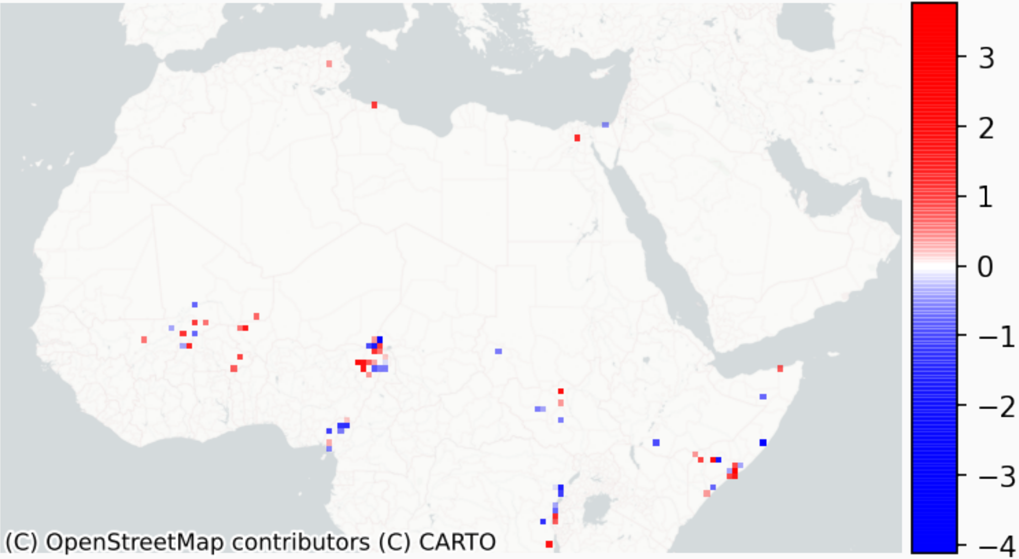
Introduction

The Goal

ViEWS Competition

- Make predictions of $\Delta ln(\text{battle deaths} + 1)_{t+s}$
- Resolution: monthly grid cells
- Grid cells: $\sim$2500.0 km^2 (one-half degree lat / lon)
- Time frame: 1990–2020

What Does the Target Look Like?



Features

	Variable	Description
1	*ln_ged_best_sb*	Current $\ln(\text{deaths} + 1)$
2	*pgd_bdist3*	Border distance (km)
3	*pgd_capdist*	Distance to capital (km)
4	*pgd_agri_ih*	Agricultural area %
5	*pgd_pop_gpw_sum*	Population
6	*pgd_ttime_mean*	Travel time to major city
7	*spdist_pgd_diamsec*	Diamond resources (spatial lag?)
8	*pgd_pasture_ih*	Pasture area %
9	*pgd_savanna_ih*	Savanna area %
10	*pgd_forest_ih*	Forest area %
11	*pgd_urban_ih*	Urban area %
12	*pgd_barren_ih*	Barren area %
13	*pgd_gcp_mer*	Gross cell product (USD)

* A subset of features from the benchmark model.

Features

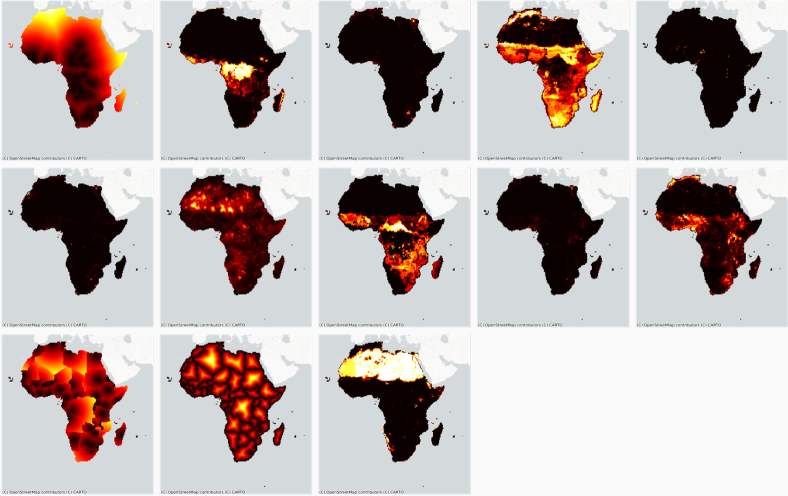
	Variable	Description
1	*ln_ged_best_sb*	Current $\ln(\text{deaths} + 1)$
2	*pgd_bdist3*	Border distance (km)
3	*pgd_capdist*	Distance to capital (km)
4	*pgd_agri_ih*	Agricultural area %
5	*pgd_pop_gpw_sum*	Population
6	*pgd_ttime_mean*	Travel time to major city
7	*spdist_pgd_diamsec*	Diamond resources (spatial lag?)
8	*pgd_pasture_ih*	Pasture area %
9	*pgd_savanna_ih*	Savanna area %
10	*pgd_forest_ih*	Forest area %
11	*pgd_urban_ih*	Urban area %
12	*pgd_barren_ih*	Barren area %
13	*pgd_gcp_mer*	Gross cell product (USD)

* A subset of features from the benchmark model.

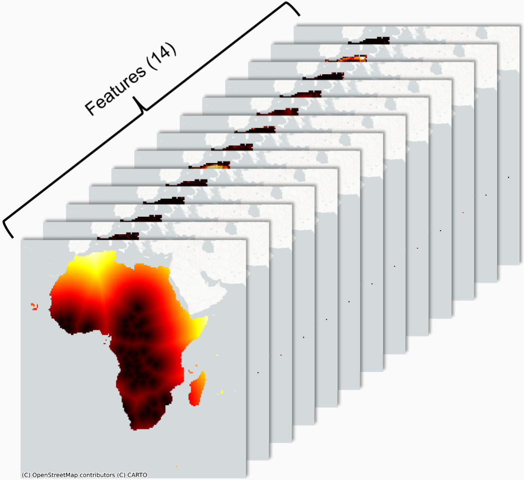
What's the Data Look Like?

Let's consider what our data "look like," to motivate our modeling choices.

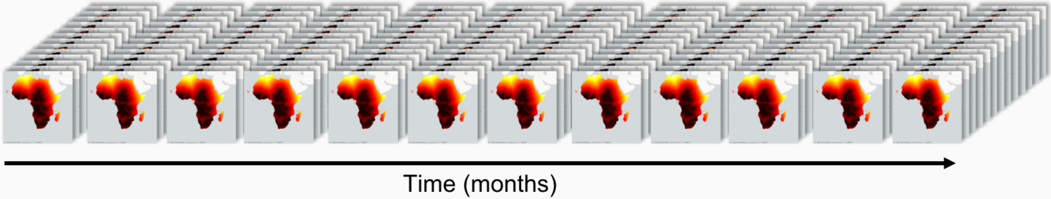
Feature Maps



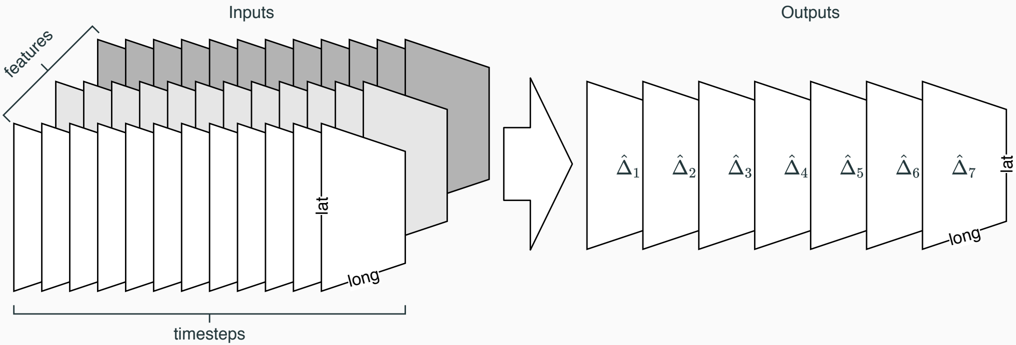
Feature per Month



Features Over Time



Reshape


Inputs
Outputs
features
timesteps
lat
long
$\hat{\Delta}_1$
$\hat{\Delta}_2$
$\hat{\Delta}_3$
$\hat{\Delta}_4$
$\hat{\Delta}_5$
$\hat{\Delta}_6$
$\hat{\Delta}_7$
lat
long

Missing Cells

What about missing PGM cells?

- Some cells are missing (because ocean...).
- Let's just add them!

But what about the feature values?

- We could cleverly mask these cells.
- Instead, let's call them 0...
- ...and add a "missing" feature to indicate them.

Input Size

Sample Size

$(12 \times 178 \times 169 \times 14) = \mathbf{5,053,776}$ values per observation.

Training Set Size

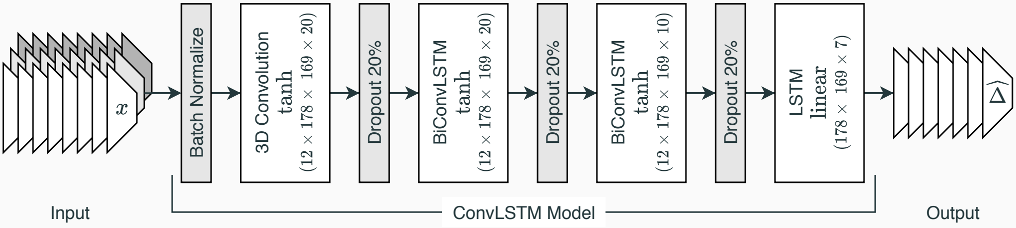
$5,053,776/12 \times 270 = \mathbf{113.7M}$.

Output Size

$(178 \times 169 \times 7) = \mathbf{210,574}$ values per output observation.

Approach

Model


x
Batch Normalize
3D Convolution
tanh
$(12 \times 178 \times 169 \times 20)$
Dropout 20%
BiConvLSTM
tanh
$(12 \times 178 \times 169 \times 20)$
Dropout 20%
BiConvLSTM
tanh
$(12 \times 178 \times 169 \times 10)$
Dropout 20%
LSTM
linear
$(178 \times 169 \times 7)$
$\widehat{\Delta}$
Input
ConvLSTM Model
Output

- 281,016 parameters
- Loss: MSE
- Optimizer: RMSprop
- Batch Size: 8
- Epochs: 75

Training time: about 1.5 hours

Results

Table 1: Validation Set

Steps	MSE	TADDA
$s = 1$	0.020001	0.013797
$s = 2$	0.021097	0.014095
$s = 3$	0.020870	0.013470
$s = 4$	0.021124	0.013904
$s = 5$	0.021368	0.013742
$s = 6$	0.021357	0.014156
$s = 7$	0.021576	0.014696

Table 2: Test Set

Steps	MSE	TADDA
$s = 1$	0.021483	0.016579
$s = 2$	0.022296	0.016795
$s = 3$	0.022141	0.016235
$s = 4$	0.022344	0.016404
$s = 5$	0.022486	0.016198
$s = 6$	0.022962	0.016912
$s = 7$	0.022581	0.017468

Max Predictions in Test Set (+2 Months)

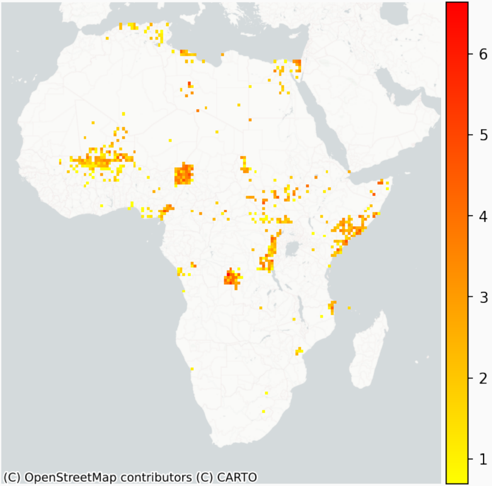


Figure 2: Observed Max

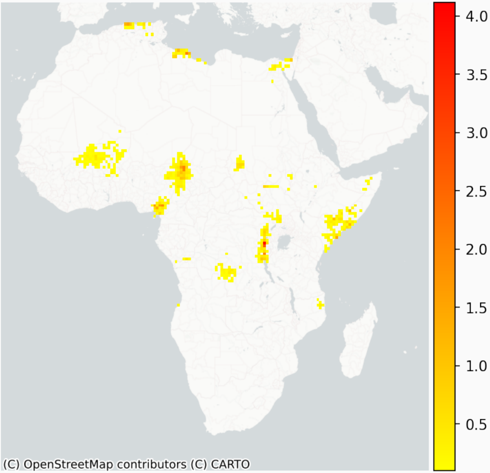


Figure 3: Predicted Max

Min Predictions in Test Set (+2 Months)

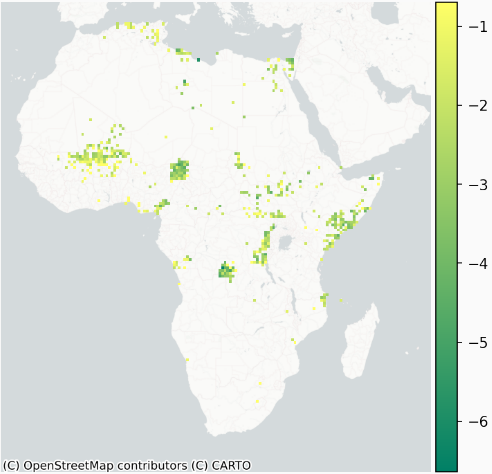


Figure 4: Observed Min

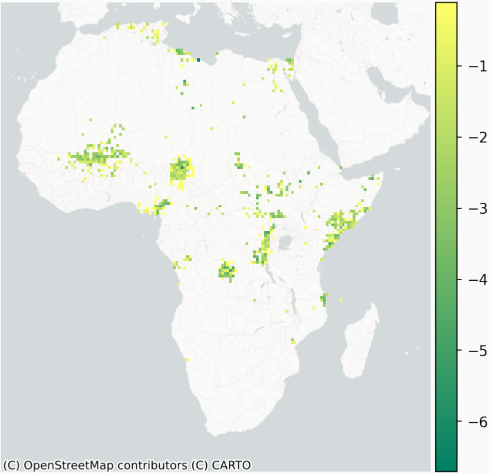
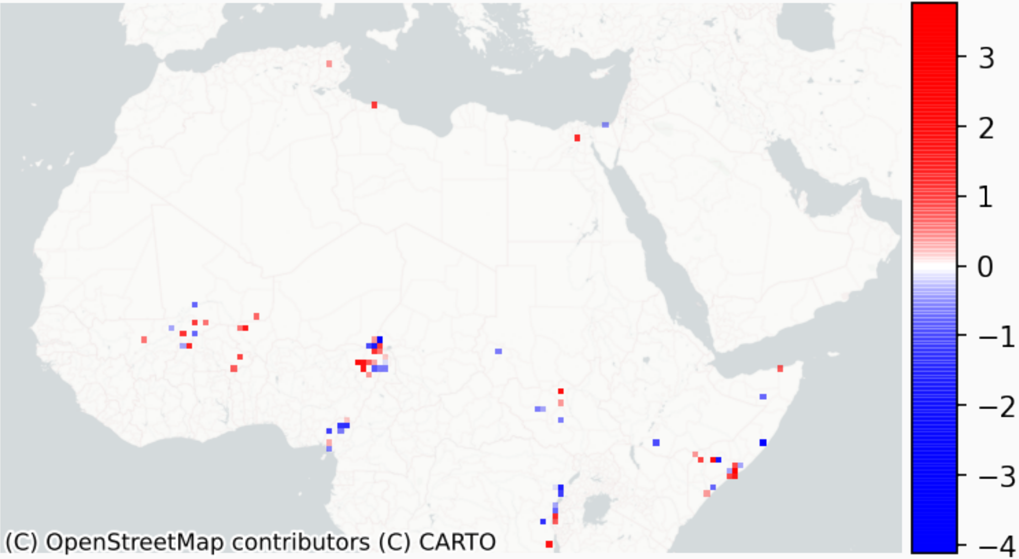
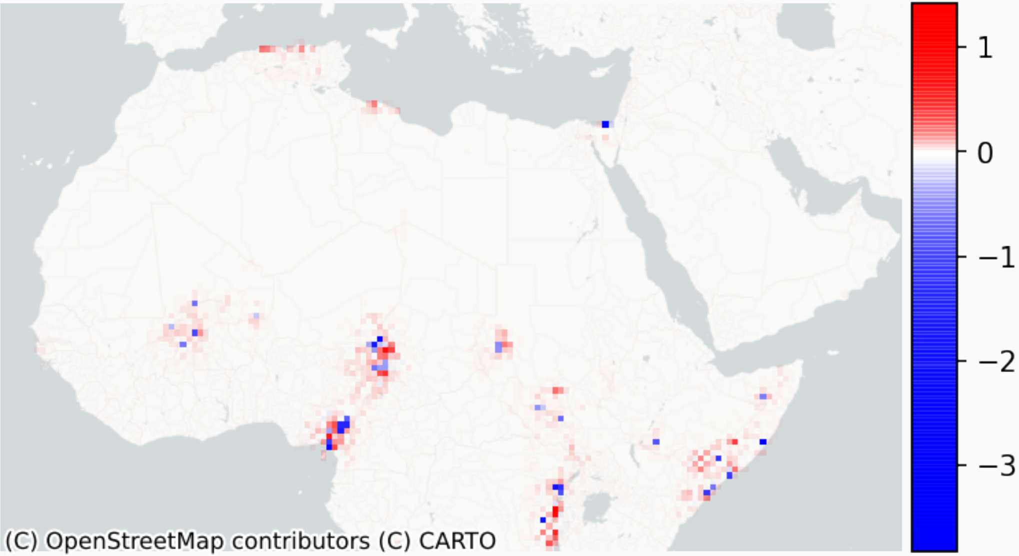


Figure 5: Predicted Min

Actual: December 2018 (+2 Months)



Predicted: December 2018 (+2 Months)



What's the Model Learning?

What if...

- ... the model is only learning a reversion to the mode (0) when the current death count is greater than 0,
- ... and, when the current death count is 0, it just predicts something like the mean increase in deaths from the training set?

$$\hat{\Delta}_{s=X} = \begin{cases} -\ln(\text{deaths}+1)_{t=0} & \text{if } \ln(\text{deaths}+1)_{t=0} > 0 \\ \bar{\Delta}_{s\neq X} & \text{else} \end{cases}$$

Actual versus Predicted

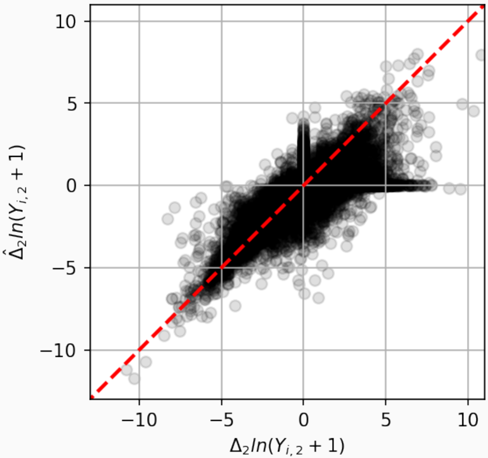


Figure 6: Observed v. Predicted

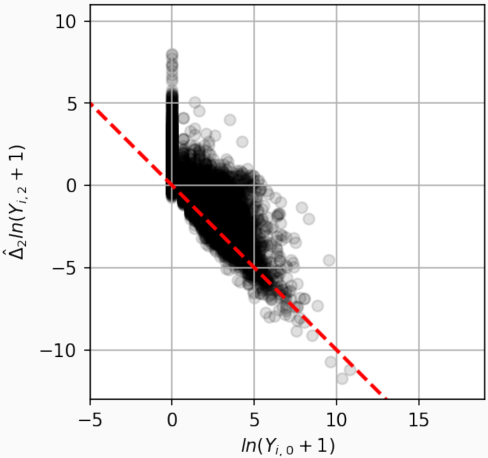


Figure 7: Count v. Prediction

Actual versus Predicted

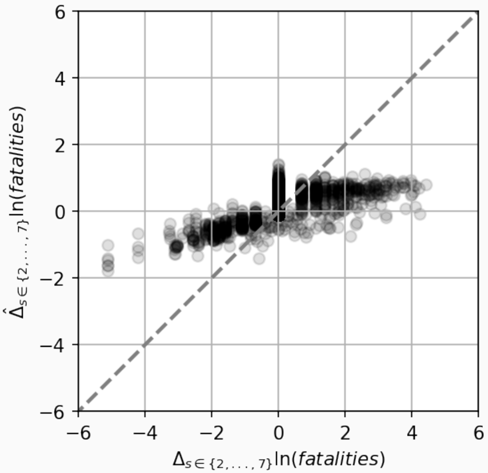


Figure 8: Benchmark Model

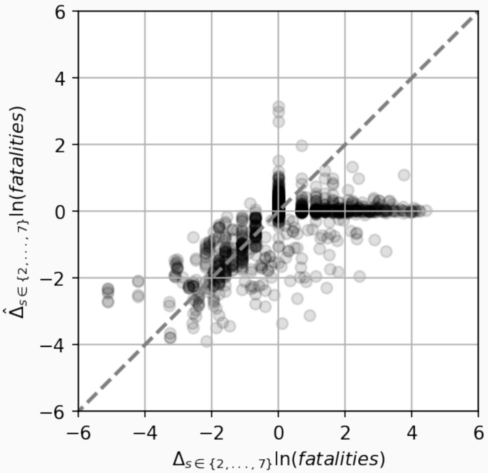


Figure 9: ConvLSTM

It's a Black Box

Not quite...

Methods to Inspect Model

- Shapley values
- LIME
- Occlusion Sensitivity (Zeiler & Fergus, 2014)
- Attention Layer (Bahdanau, Cho, & Bengio, 2015)
- Alternative Models

Attention Layer

Feature	Importance
ln_ged_best_sb	0.284
pgd_pop_gpw_sum	0.271
pgd_urban_ih	0.207
pgd_ttime_mean	0.051
pgd_agri_ih	0.040
pgd_gcp_mer	0.035
pgd_forest_ih	0.029
spdist_pgd_diamsec	0.017
pgd_barren_ih	0.016
pgd_bdist3	0.014
pgd_savanna_ih	0.012
pgd_pasture_ih	0.011
missing_indicator	0.010
pgd_capdist	0.010

Let's try the same ConvLSTM model with *only* one feature:

$$ln(\text{battle deaths} + 1).$$

Single Feature Model

	Competition Model		Single Feature	
Steps	MSE	TADDA	MSE	TADDA
$s = 2$	0.022	0.017	0.022	0.013
$s = 3$	0.022	0.016	0.022	0.013
$s = 4$	0.022	0.016	0.022	0.014
$s = 5$	0.022	0.016	0.022	0.013
$s = 6$	0.023	0.017	0.022	0.013
$s = 7$	0.023	0.017	0.022	0.014

Next

Let's do it better!

Upcoming Work

- What's the right resolution?
  - (Dis)aggregation probably degrades signal.
  - Spatio-temporal point processes.
- What are leading signals of violence?
  - Event data
  - Mobilization
  - Social media

Other Issues

- The "pixels" of our map images aren't equal area!
  - Rewrite the internals of the convolutional layer.
  - Interpolate $\longrightarrow$ Model $\longrightarrow$ Aggregate.
  - Use a graph convolutional network.

Thank you

Benjamin J. Radford
benjamin.radford@uncc.edu