Identifying Strategic Weather Forecast Bias: Case of Typhoon in Southeast Asia

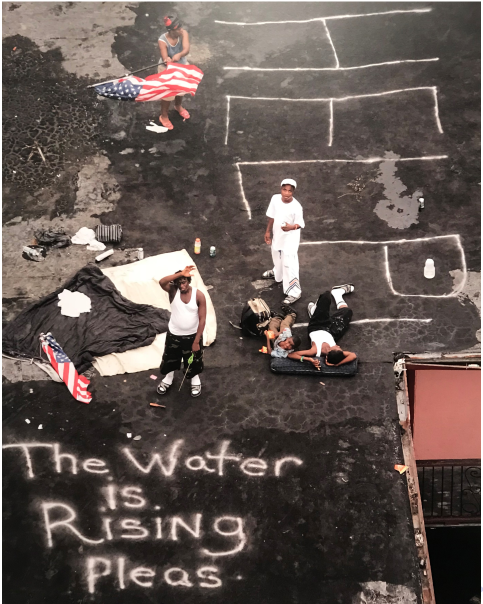
Yi Cheng[1] Chungsang Tom Lam[2]

Georgia Institute of Technology[1]

Clemson University[2]

November 20, 2018

Hurricane Katrina



Typhoon Forecasting Narratives

- NCEP delivers science-based environmental predictions to the Nation and the global community. We collaborate with partners and customers to produce reliable, timely, and accurate analyses , guidance, forecasts and warnings for the protection of life and property and the enhancement of the national economy.

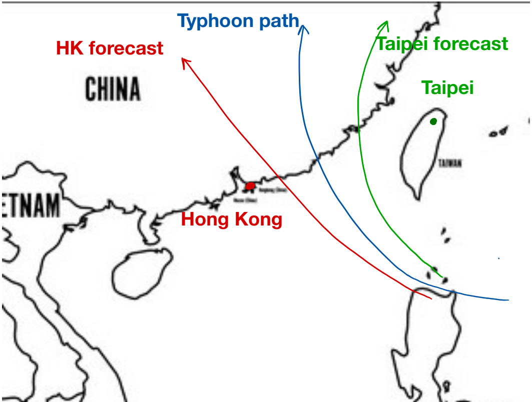
Source: National Centers for Environmental Prediction's Mission and Vision

- TV meteorologists weren't placing much emphasis on accuracy
- Aren't bothering to make accurate forecasts
- Presentation takes precedence over accuracy

Source: Silver, Nate "The Signal and the Noise - Why so many predictions fail but some don't"

Research Question

- Construct model to capture the mechanism behind biased typhoon forecasting
- Identify empirically the strategic bias in typhoon forecasting
- Analyze the factors that determine the bias



Model

Assumptions

- The observatory minimizes the expected cost of misreporting
- Costs incur when the weather forecasting is inaccurate
- Two types of inaccurate forecasting
  - Observatory forecasts that a typhoon hits but it does not
  - Observatory does not forecast a typhoon hits but it does
- The costs are different

Model

$$C^G(P) = \text{minimize}_{(F_x,F_y)}\ \alpha_1 \mathbb{E}\{[\sqrt{R_x^2+R_y^2} - \sqrt{F_x^2+F_y^2}]^+\} + \alpha_2 \mathbb{E}\{[\sqrt{F_x^2+F_y^2} - \sqrt{R_x^2+R_y^2}]^+\} \quad (1)$$

- 2-dimensional coordinate system S, with origin as the observatory's location
- R_x, R_y: random variable of real typhoon location in S (mean: μ_x, μ_y; variance: σ_x^2, σ_y^2)
- F_x, F_y: decision variable of forecast typhoon location in S
- α_1 and α_2 are the cost caused by the gap of aggressive and conservative biased forecasting, respectively. $\alpha_1 \neq \alpha_2$

Model

Optimal Forecast location:

$$F_x^* = \sqrt{\mu_x + \frac{\sigma_x \alpha}{\sqrt{1-\alpha^2}}}$$

$$\text{Similarly, } F_y^* = \sqrt{\mu_y + \frac{\sigma_y \alpha}{\sqrt{1-\alpha^2}}}$$

where

$$\alpha = \frac{\alpha_1 - \alpha_2}{\alpha_1 + \alpha_2}$$

Implications:

- The forecast is based on the mean μ_x, μ_y
- if σ_x or σ_y are larger, the forecast will be more distanced from the mean of the real location
- if $|\alpha|$ is larger, the forecast will be more distanced from the mean of the real location

Empirical Challenges

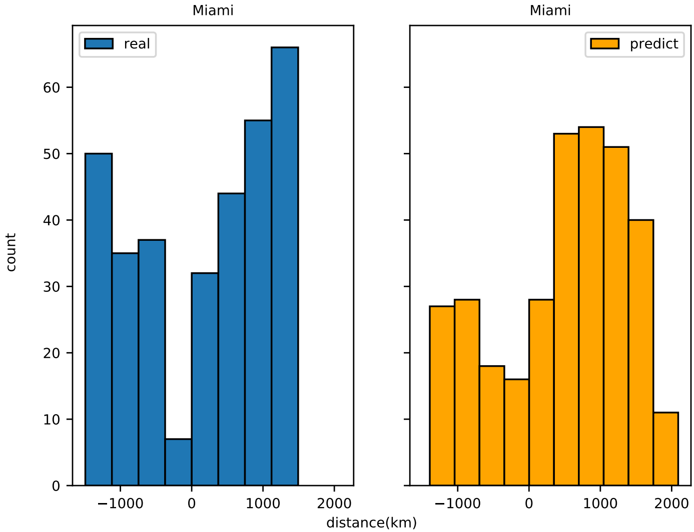
- Data
  - Historical typhoon paths are super easily available (long history of records)
  - However, we need more than historical typhoon paths. Since we are analyzing the observatories' behaviour, not the typhoon's behaviour
  - Historical forecast paths is needed (This limits the sample size)

Empirical Challenges

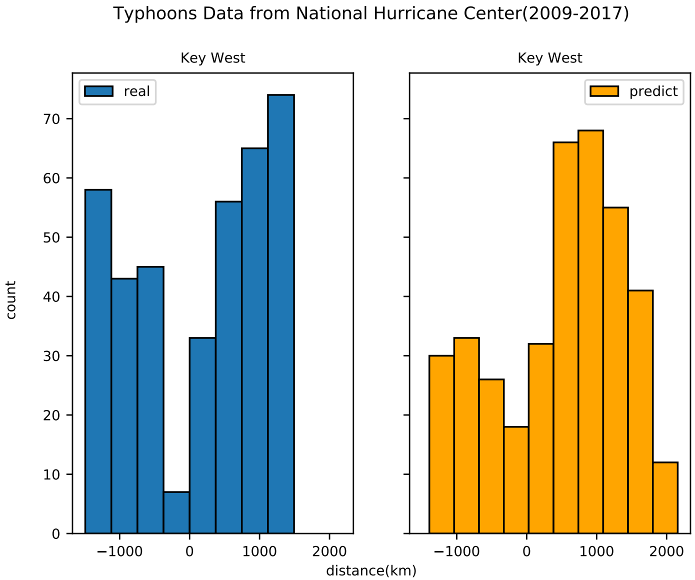
- Identification
  - Identifying bias (statistically significant) is not difficult
  - The difficulty is about the “strategic” bias.
  - Typhoon forecasting is a very difficult task, even with modern technology
  - Inaccurate reports from the observatories do not necessarily indicate intentionally biased forecasts

Some NHC Examples

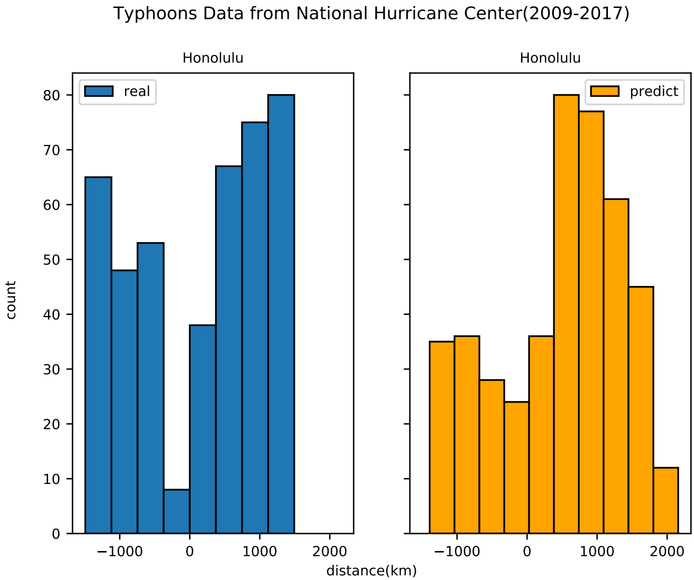
Typhoons Data from National Hurricane Center(2009-2017)



Some NHC Examples



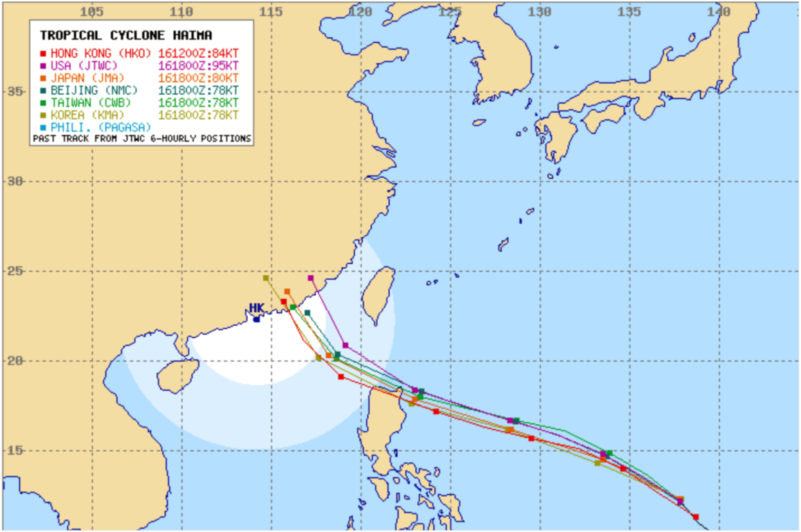
Some NHC Examples



The case we need

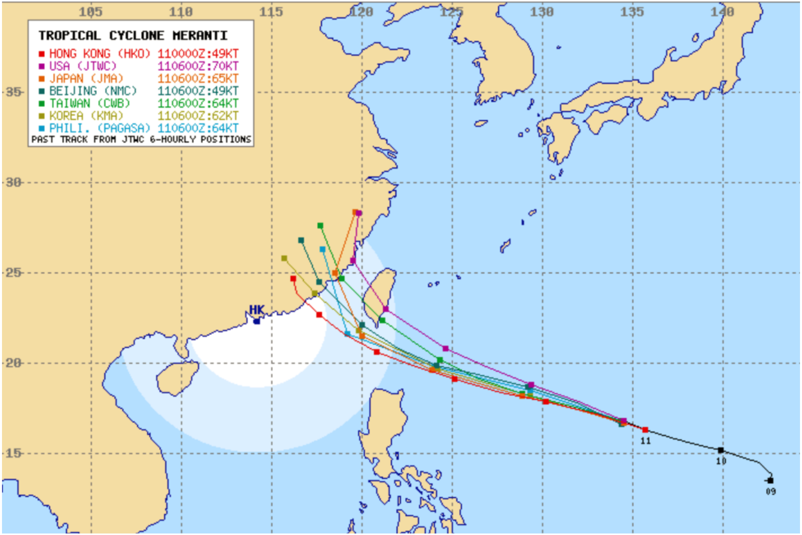
- Two independently operated observatories
- Observatories in small cities
- Similar technology
- Many typhoons paths between the two cities
- Typhoons with relatively simple paths

Cherry Picking Example



Source: https://typhoon2000.ph

Cherry Picking Example



Source: https://typhoon2000.ph

Data

- Data Source: https://typhoon2000.ph
- Historical typhoon data recorded of HKO(Hong kong Observatory) and CWB (Central Weather Bureau of Taiwan)
- From 2009-2017, with 204 typhoons and 14701 forecast records
- The data include forecast location records and real location records
- The data include typhoon properties such as geological location, wind intensity and forecast time.

Data

Table: DOKSURI and HATO Data Sample

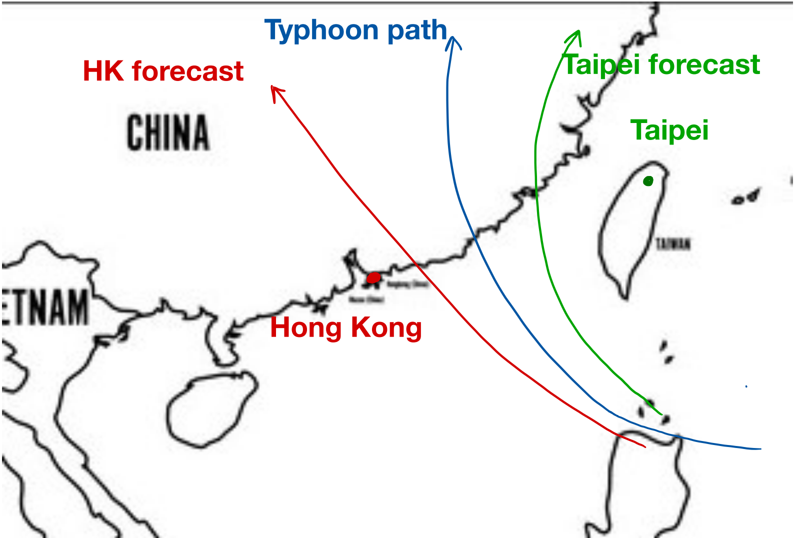
Name	Year	Anchor	(Latitude,Longitude)	Knot	Observatory
DOKSURI	2017	+000H	(14.9N,119.6E)	30	HKO
DOKSURI	2017	+024H	(16.1N,116.4E)	49	HKO
DOKSURI	2017	+048H	(17.2N,113.0E)	70	HKO
DOKSURI	2017	+072H	(18.6N,109.7E)	70	HKO
DOKSURI	2017	+096H	(20.4N,104.9E)	40	HKO
DOKSURI	2017	+120H	(22.8N,101.6E)	22	HKO
DOKSURI	2017	+000H	(14.0N,119.0E)	35	CWB
DOKSURI	2017	+024H	(14.9N,115.6E)	54	CWB
DOKSURI	2017	+048H	(16.7N,112.1E)	58	CWB
DOKSURI	2017	+072H	(18.3N,108.8E)	68	CWB
DOKSURI	2017	+096H	(20.5N,104.8E)	58	CWB
DOKSURI	2017	+120H	(21.6N,102.6E)	45	CWB
HATO	2017	+000H	(19.6N, 128.0E)	30	HKO
HATO	2017	+024H	(20.4N, 124.0E)	40	HKO
HATO	2017	+048H	(22.2N,119,9E)	49	HKO
HATO	2017	+072H	(23.1N,114.9E)	35	HKO
HATO	2017	+096H	(25.4N,109.4E)	22	HKO
HATO	2017	+000H	(20.0N,128.0E)	35	CWB
HATO	2017	+024H	(20.8N,124.2E)	45	CWB
HATO	2017	+048H	(21.9N,120.0E)	45	CWB
HATO	2017	+072H	(23.0N, 115.6E)	39	CWB
HATO	2017	+096H	(24.0N, 109.9E)	29	CWB

Data

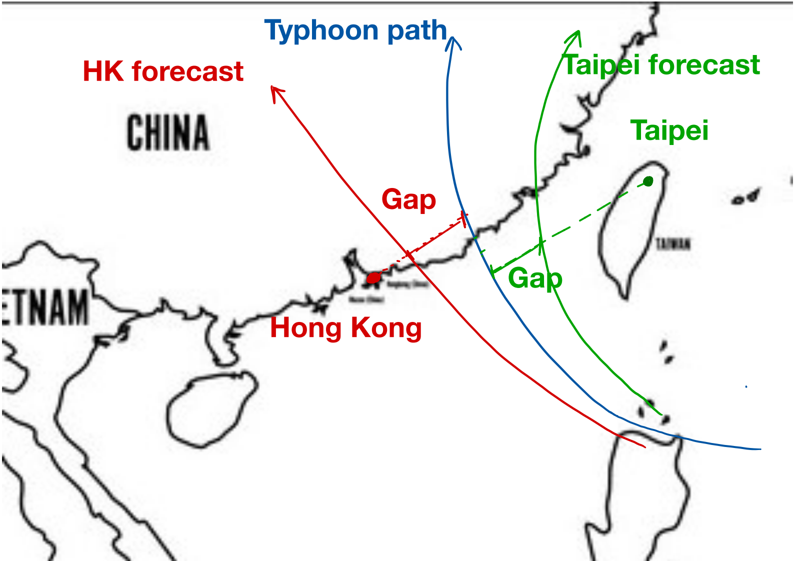
Two key variables

- Gap
  - The difference between
    "the distance between forecast location and the observatory"
    and
    "the distance between real location and the observatory"
- Forecast comparison(Hong Kong)
  - the difference between
    "HKO's forecast distance to Hong Kong"
    and
    "CWB's forecast distance to Hong Kong"
- Similarly we can calculate the Forecast comparison(Taiwan)

Data

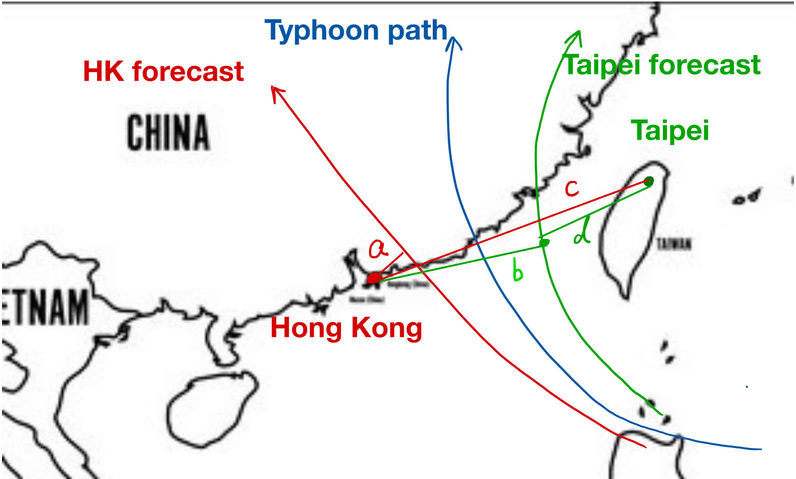

Typhoon path
HK forecast
Taipei forecast
CHINA
Taipei
TAIWAN
ETNAM
Hong Kong

Data


Typhoon path
HK forecast
Taipei forecast
CHINA
Taipei
Gap
Gap
ETNAM
Hong Kong
TAIWAN

Data

- Forecast comparison(Hong Kong):= a - b
- Forecast comparison(Taiwan):= d - c



Data

Table: Summary Statistics

	Knot	Forecast	Gap	Forecast comparison (Hong Kong)	Real
count	14701	14701	14701	14701	14701
mean	65.303	1536.320	-2.400	1714.627	1538.660
std	22.780	913.939	167.991	1122.884	905.078
min	0.000	12.207	-1260.352	2.625	32.271
25%	46.000	857.517	-77.220	887.421	867.053
50%	64.000	1392.739	-2.281	1450.060	1410.708
75%	84.000	2034.294	68.792	2296.329	2036.452
max	140.000	5781.257	1341.319	6594.057	5805.630

Gap

Table: Analyzing the Gap

Observatory	Location type	Scope(km)	Mean(km)	t-stat
CWB	Between	less than 400	-75.527	10.992
CWB	East of Taiwan	less than 400	-45.171	6.177
HKO	Between	less than 400	-42.563	6.118
HKO	West of Hongkong	less than 400	-52.332	6.177

- Both observatories forecast typhoons to land closer to their own location than the (mean of the) real situation
- This is just replicating the histograms

Forecast Comparison

Table: Analyzing the Forecast comparison

Observatory	Counterpart	Scope(km)	Mean(km)	t-stat
CWB	HKO	less than 400	-154.177	10.248
HKO	CWB	less than 400	-162.123	10.647

- Both observatories forecast typhoons to land closer to their own location than their counterpart.
- This is true even if after controling for typhoon fixed effects

Regression

Table: Regression results on Location type (HKO)

Dep. Variable:	gap
Intercept	139.739(22.661)***
Between HK and TW	-51.332(8.195)***
West of HK	-48.735(9.294)***
wind type	-79.580(31.317)*
Anchor	-7.942(2.605)**
Other Controls	Yes
Fixed Effects	Yes

Regression

Table: Regression Results on Wind Type (HKO and CWB)

Dep. Variable:	gap
Intercept	45.164(16.308)**
Wind Type	-104.595(19.851)***
East of TW	17.760(5.513)**
West of HK	40.054(8.212)***
Anchor	-4.528(1.390)***
knot	0.751(0.261)**
real	-0.070(0.012)***
Other controls	Yes
Observatory Fixed Effects	Yes
Typhoon Fixed Effects	Yes

Others

- The bias increases when there are or more typhoons approaching (σ)
- The bias direction changes depending on the direction of error of the previous typhoon (α)

Conclusion

- Accuracy is the best policy for a forecaster. It is forecasting's original sin to put politics, personal glory or economic benefit before the truth of the forecasts

Source: Silver, Nate "The Signal and the Noise - Why so many predictions fail but some don't"