Fused Deep Learning For Hurricane Track Forecast From Reanalysis Data

Mo YANG

Supervisor:

Claire Monteleoni
Guillaume Charpiat
Sophie Giffard-Roison

September 3, 2018

Outline

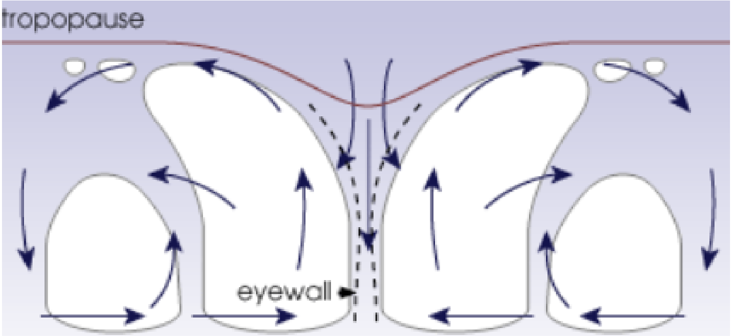
1. Background
   - Hurricane Track Forecast
   - Hurricane Track Forecasting Models for Meteorologists
2. Preliminaries
   - Problem Setting
   - Previous Works
3. Method
   - Data Description
   - Feature Selection
   - The proposed method
   - Fusion Network Training Framework
4. Experiments
   - Selecting Network Configurations
   - Comparison with the Existing Forecasting Models
5. Discussion
6. Discussion

Hurricane

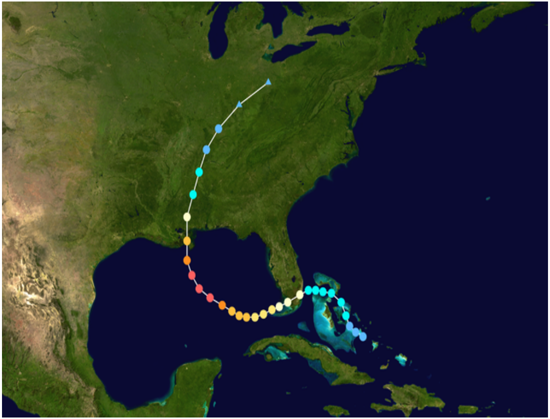
- Cyclones, hurricanes or typhoons: words for the same phenomena
- Cost tremendous damage and death each year
- Can be very unpredictable.
- **Tracks** and **Intensity** : Two main goals of the forecast

Hurricane

- How hurricane is formed



- The evolution and path of the hurricane Katrina (2005)



Saffir-Simpson Hurricane Scale		
Category	Wind Speed	
	mph	knots
5	>=156	>=135
4	131-155	114-134
3	111-130	96-113
2	96-110	84-95
1	74-95	65-83
Non-Hurricane Scale		
Tropical Storm	39-73	34-64
Tropical Depression	0-38	0-33

Guidance Models

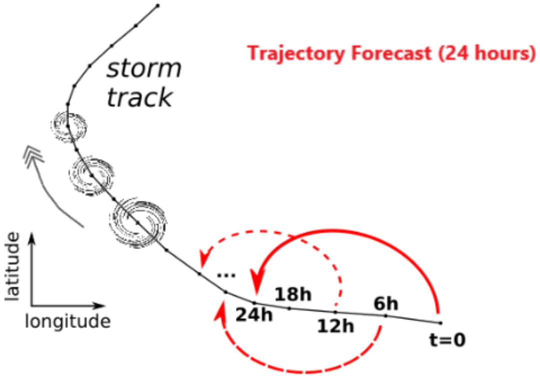
- Statistical models: do not consider physics, generate forecast in seconds. e.g. Best Track Decay (BCD5)
- Dynamic models: solve the physical equations governing the motions in the atmosphere, more accurate but computationally demanding.
- Official NHC forecast (OFCL) : consensus models which are created by combining the forecasts from a collection of other models.

Outline

1. Background
   - Hurricane Track Forecast
   - Hurricane Track Forecasting Models for Meteorologists
2. Preliminaries
   - Problem Setting
   - Previous Works
3. Method
   - Data Description
   - Feature Selection
   - The proposed method
   - Fusion Network Training Framework
4. Experiments
   - Selecting Network Configurations
   - Comparison with the Existing Forecasting Models
5. Discussion
6. Discussion

Problem Setting

- Goal: estimating the 24h-forecast trajectory of all tropical storms.



Previous Works

- A large family of previous methods use spatio-temporal features
  - Convolutional LSTM [Xingjian et al., 2015]
  - Temporal Convolutional Networks [Bai et al., 2018]
- Only two preliminary studies have tackled the hurricane forecast tracking using machine learning
  - Use random forests on local reanalysis histograms [Liberge et al., 2011]
  - Train a sparse recurrent neural network from trajectory data [Moradi Kordmahalleh et al., 2016]

Outline

1. Background
   - Hurricane Track Forecast
   - Hurricane Track Forecasting Models for Meteorologists
2. Preliminaries
   - Problem Setting
   - Previous Works
3. Method
   - Data Description
   - Feature Selection
   - The proposed method
   - Fusion Network Training Framework
4. Experiments
   - Selecting Network Configurations
   - Comparison with the Existing Forecasting Models
5. Discussion
6. Discussion

- **Storm track data**: composed of tropical/extra-tropical storm tracks since 1979 extracted from the NOAA database IBTrACS.

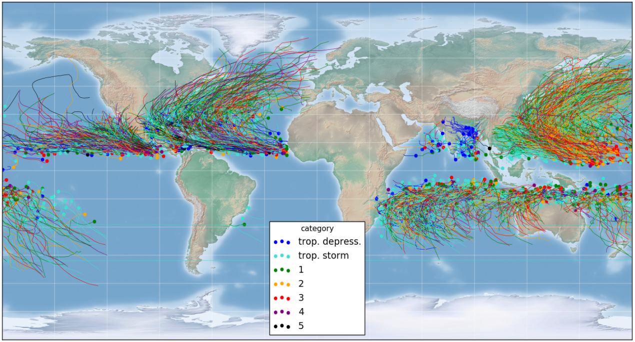


Figure: Database: tropical/extra-tropical storm tracks since 1979. Dots = initial position, colors = maximal storm strength according to the Saffir-Simpson scale.

Data

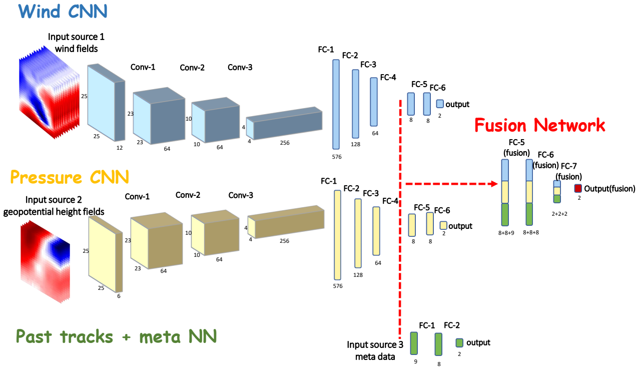
- **Reanalysis**: dynamically consistent estimates of the climate atmospheric fields or state. Created via data assimilation scheme and model(s) which take all available observations (radiosonde, satellite, buoy, aircraft and ship reports ...) as inputs every 6 hours.
- ERA-Interim: the latest global atmospheric reanalysis produced by The European Centre for Medium-Range Weather Forecasts (ECMWF).
    - spectral resolution: about 80km on 60 vertical levels from the surface up to 0.1 hPa
    - atmospheric fields: u-wind, v-wind, pressure, temporature, humiity, vorticity ......

Feature Selection

- **The wind fields.** u-wind and v-wind fields on a 25x25 degree grid centered on the current storm location, at three atmospheric pressure levels (700 hPa, 500 hPa, and 225 hPa).
- **The pressure fields.** geopotential height fields on a 25x25 degree grid centered on the current storm location, at three atmospheric pressure levels (700 hPa, 500 hPa, and 225 hPa)
- **Displacement in history.**
- **Other hand-crafted features:** current lat. & lon., windspeed, Jday predictor(Gaussian function of "Julian day of storm init - peak day of the hurricane season" [DeMaria et al., 2005]), and current distance to land.

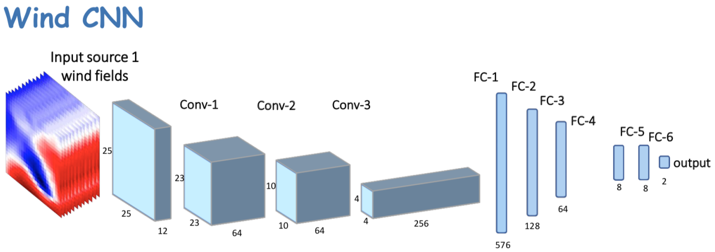
Overview of the proposed method

- Three-stream fusion network



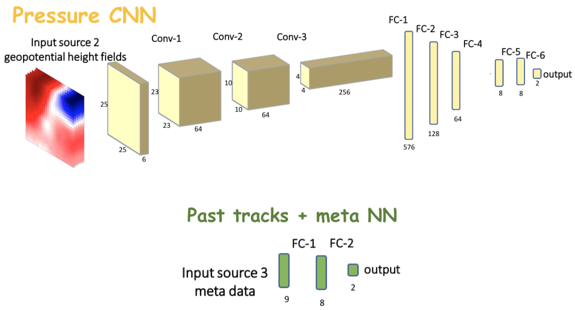
Train Stream Networks

- Stage I: Train stream networks
  - Wind CNN
    - pretty standard (popular) CNN architecture, like VGG Net [Simonyan and Zisserman, 2014]
    - trying from shallow to deep
    - **Details:** ReLU, batch norm, adam, batch size=32, L2 *coeff*. = 0.01, initial learning rate = 0.001, *He* initialization [He et al., 2015], everything fine tuned... **No magic!**



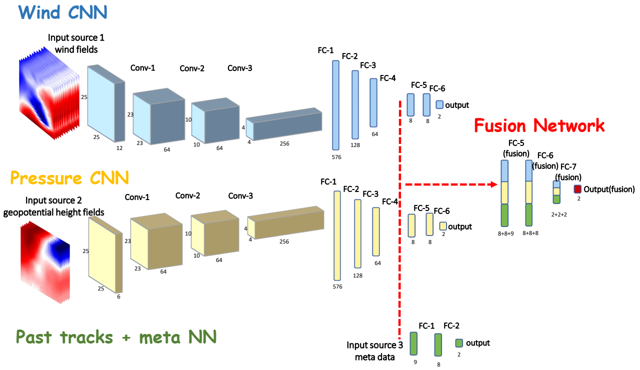
Train Stream Networks

- Stage I: Train stream networks
  - Pressure CNN
    - same principle...
  - Past tracks + meta NN
    - a simple multi-layer perception (MLP)



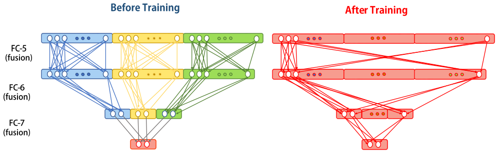
Train Fusion Network

- Stage II: Train the fusion network



Train Fusion Network

- Stage II: Train the fusion network
  - zoom in fusion layers



  - two-phase optimization:
    - first phase: optimize weights only in fusion layers
    - seconc phase: optimize weights in all layerts

Outline

1. Background
   - Hurricane Track Forecast
   - Hurricane Track Forecasting Models for Meteorologists
2. Preliminaries
   - Problem Setting
   - Previous Works
3. Method
   - Data Description
   - Feature Selection
   - The proposed method
   - Fusion Network Training Framework
4. Experiments
   - Selecting Network Configurations
   - Comparison with the Existing Forecasting Models
5. Discussion
6. Discussion

Input preprocessing

- the entire dataset into train (60%) / validation (20%) / test (20%)
- separate all the storms into the three bins to avoid data leakage
- input data standardized by subtracting mean value then dividing by the standard variation of each layer, computed on the training set, from each feature

Training

Let $D = \{(x_w^1, x_p^1, x_{0d}^1, loc_t^1, loc_{t+\delta t}^1), ..., (x_w^n, x_p^n, x_{0d}^n, loc_t^n, loc_{t+\delta t}^n)\}$ be the labeled data; with x_w^i, x_p^i, x_{0d}^i denoting wind fields, pressure fields, 0d features respectively and loc_t^i, $loc_{t+\delta t}^i$ corresponding to the ground truth location at current time and after δt time. δt can be 6 hours, 12 hours, 18 hours, 24 hours... depending on the specific forecasting task.

Training

- Stage I: Train the stream networks
  let F_{wind} be the mapping function of Wind CNN, the network parameters are marked as θ_w, $F_{transform}$ is the transformation function from differences in latitude and longitude to distance. The training of Wind CNN is carried out by optimizing on the loss (MSE) between the forecast storm location and the corresponding ground truth location

$$L(\theta_w) = \frac{1}{n}\sum_{i=1}^{n} \left\| F_{transform}((F_{wind}(x_w^i;\theta_w) + loc_t^i) - loc_{t+\delta t}^i) \right\|^2 \quad (1)$$

  Where n is the total number of instances. Training the Pressure CNN and Past tracks + meta NN can be performed analogously.

Training

- two-phase manner
  - Let F_f be the function of fusion network, θ_f, θ_s be parameters of fusion layers and of stream layers. In the first phase, optimize only the weights in the fusion layers (keeping the weights of the stream layers intact). The first phase's loss function would be:

$$L(\theta_f) = \frac{1}{n}\sum_{i=1}^{n} \left\| F_{transform}((F_f(x_w^i, x_p^i, x_{0d}^i; \theta_f) + loc_t^i) - loc_{t+\delta t}^i) \right\|^2 \quad (2)$$

  - In the second phase, let the weights in the whole network be equally optimized. The second phase's loss function would be:

$$L(\theta_f, \theta_s) = \frac{1}{n}\sum_{i=1}^{n} \left\| F_{transform}((F_f(x_w^i, x_p^i, x_{0d}^i; \theta_f, \theta_s) + loc_t^i) - loc_{t+\delta t}^i) \right\|^2 \quad (3)$$

Outline

1. Background
   - Hurricane Track Forecast
   - Hurricane Track Forecasting Models for Meteorologists
2. Preliminaries
   - Problem Setting
   - Previous Works
3. Method
   - Data Description
   - Feature Selection
   - The proposed method
   - Fusion Network Training Framework
4. Experiments
   - Selecting Network Configurations
   - Comparison with the Existing Forecasting Models
5. Discussion
6. Discussion

- Four candidates' network configurations for Wind CNN
  - followed the same generic design
  - from shalow to deep

ConvNet Configurations			
A	B	C	D
7 layers	8 layers	9 layers	10 layers
input (25*25, 12 channels image)			
conv3-32 maxpool	conv3-32 conv3-32 maxpool	conv3-64 conv3-64 maxpool conv3-256	conv3-64 conv3-64 maxpool conv3-128 conv3-256 maxpool
FC-576			
FC-128			
FC-64			
FC-8			
FC-8			
FC-2			

- Evaluation results on 24 hours of storm track prediction on the validation set

Model	Mean Square Error (km^2)	Mean Absolute Error(km)
A	31430.08	145.43
B	31761.95	146.62
C	31552.91	145.5997
D	31772.62	146.73

- Adding more temporal parts (storm fields data at the same location at more consecutive time steps to input data)
  - No noticeable improvement

Two senarios:

- how many layers should be fused?
- Does fusing three streams outperforms fusing two streams or using single stream?

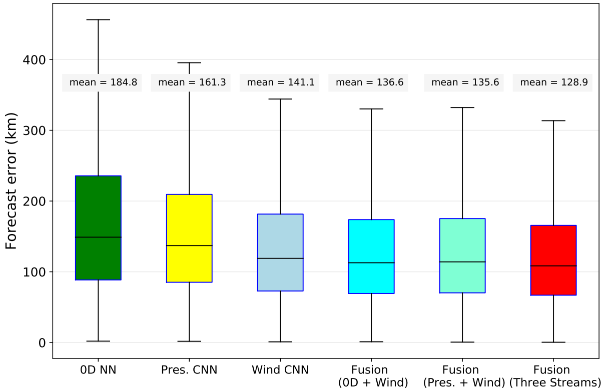
- how many layers should be fused?
  - Comparison of fusion networks that fuse different number of layers on 24 hours storm track prediction on validation set.

Model	Mean Square Error(km^2)	Mean Absolute Error(km)
Fus network, 3 streams, fuse 2 FC	25453.48	130.04
Fus network, 3 streams, fuse 4 FC	25628.22	130.27
Fus network, 3 streams, fuse 6 FC	25846.74	130.95
Fus network, 3 streams, fuse 6 FC + 1 Conv layers	25928.81	131.17

- Does fusing three streams outperforms fusing two streams or using single stream?
  - Comparison between the fusion network fusing all three streams with networks fusing two streams and single stream networks.

Model	Mean Sqaure Error(km^2)	Mean Abosolute Error(km)
0D NN	51367.42	180.00
Pres. CNN,	39387.76	164.91
Wind CNN	30426.02	142.71
Fus network, 0D + Wind, fuse 2 FC	28723.56	136.64
Fus network, Pres. + Wind, fuse 2 FC	28045.30	137.26
Fus network, Three streams, fuse 2 FC	25453.48	130.04

Figure: 24h-forecast results on the test set (storms coming from all oceanic basins), in distance between predicted and real location.



Outline

1. Background
   - Hurricane Track Forecast
   - Hurricane Track Forecasting Models for Meteorologists
2. Preliminaries
   - Problem Setting
   - Previous Works
3. Method
   - Data Description
   - Feature Selection
   - The proposed method
   - Fusion Network Training Framework
4. Experiments
   - Selecting Network Configurations
   - Comparison with the Existing Forecasting Models
5. Discussion
6. Discussion

- **Baselines:** A commenly used statisticall model for benchmarking: Best Track Decay (BCD5)
- Also compare with: Official NHC forecast (OCFL). With increased computational power, the performance of OFCL is constantly improving

Quantitative:

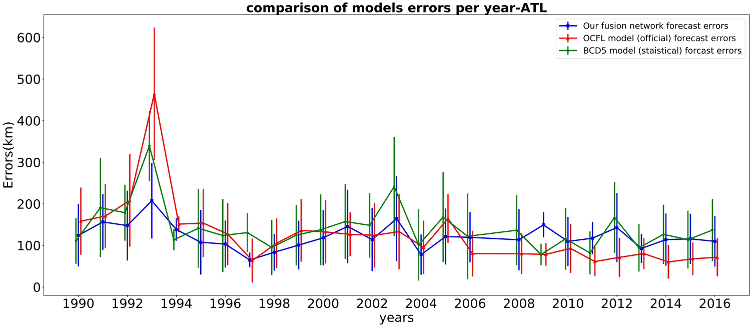


Figure: The yearly average 24-hours storm track forecasting errors (km) and standard deviation on the test set in Atlantic for our fusion network forecasts(blue), the BCD5 (green) and the OFCL (red), 1989-2016.

Quantitative:

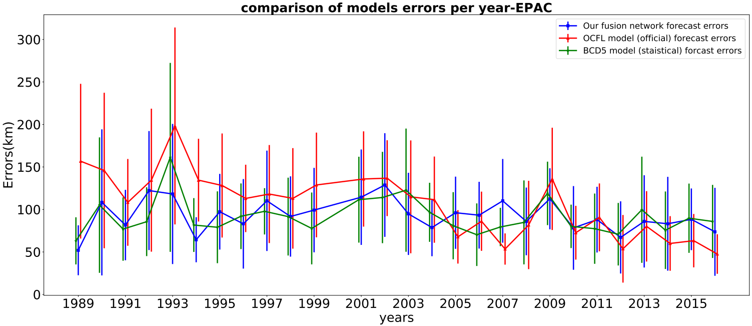


Figure: The yearly average 24-hours storm track forecasting errors (km) and standard deviation on the test set in East Pacific for our fusion network forecasts (blue), the BCD5 (green) and the OFCL (red), 1989-2016.

Quantitative:

- Mean storm track forecast errors of all years in the two basins (Atlantic and Pacific) on the test set for our fusion network and BCD5 model:

Model	Atlantic		East Pacific	
	Mean forecast errors(km)	Standard deviation	Mean forecast errors(km)	Standard deviation
BCD5	124.99	90.06	112.36	78.46
Our fusion network	114.93	69.95	94.03	58.77

Qualitative:

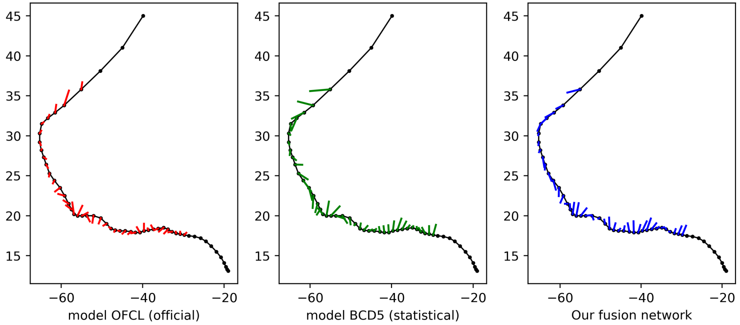


Figure: The forecast errors of the three models (left: the OCFL, middle: the BCD5, right: our fusion network) on the Hurricane Ian in 2016

Discussion

Conclusion:

- Proposed a promising deep learning framework for storm track forecasting
- Coupled different types of data (wind fields, pressure fields, past tracks and other handcrafted data) into our fusion model
- our model outperforms the BCD5 model
- our model can help to enhance the official forecast, even its forecast errors are larger than the OFCL model.

Discussion

What's next?

- Design an algorithm that could effectively learn information from high-dimensional tensor data.
- Try unsupervised learning algorithms (e.g. clustering)
- Apply multi-target learning

THANK YOU

Bai, S., Kolter, J. Z., and Koltun, V. (2018).
An empirical evaluation of generic convolutional and recurrent networks for sequence modeling.
arXiv preprint arXiv:1803.01271.

DeMaria, M., Mainelli, M., Shay, L. K., Knaff, J. A., and Kaplan, J. (2005).
Further improvements to the statistical hurricane intensity prediction scheme (ships).
Weather and Forecasting, 20(4):531–543.

He, K., Zhang, X., Ren, S., and Sun, J. (2015).
Delving deep into rectifiers: Surpassing human-level performance on imagenet classification.
In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034.

Liberge, S. M., Ba, S., Lenca, P., and Fablet, R. (2011).
Prévision de trajectoires de cyclones à l'aide de forêts aléatoires avec arbres de régression.

In *Conférence internationale francophone sur l'extraction et la gestion des connaissances*, pages 623–634. Hermann.

Moradi Kordmahalleh, M., Gorji Sefidmazgi, M., and Homaifar, A. (2016).
A sparse recurrent neural network for trajectory prediction of atlantic hurricanes.
In *Proceedings of the Genetic and Evolutionary Computation Conference 2016*, pages 957–964. ACM.

Simonyan, K. and Zisserman, A. (2014).
Very deep convolutional networks for large-scale image recognition.
arXiv preprint arXiv:1409.1556.

Xingjian, S., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., and Woo, W.-c. (2015).
Convolutional lstm network: A machine learning approach for precipitation nowcasting.
In *Advances in neural information processing systems*, pages 802–810.