Fused Deep Learning For Hurricane Track Forecast From Reanalysis Data

Mo YANG

Supervisor:

Claire Monteleoni
Guillaume Charpiat
Sophie Giffard-Roisin

September 6, 2018

Outline

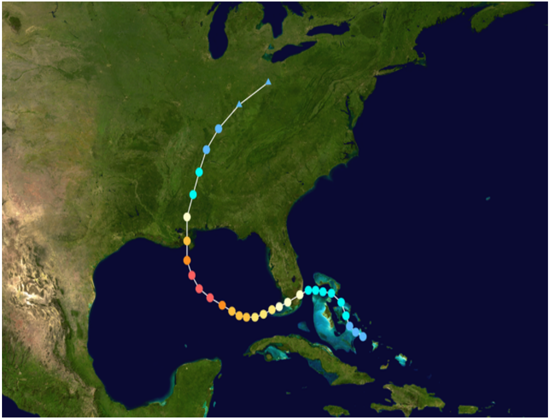
1. Background
   - Hurricane Track Forecast
   - Hurricane Track Forecasting Models for Meteorologists
2. Preliminaries
   - Problem Setting
   - Previous Works
3. Method
   - Data Description
   - The proposed method
4. Experiments
   - Selecting Network Configurations
   - Comparison with the Existing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

- 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

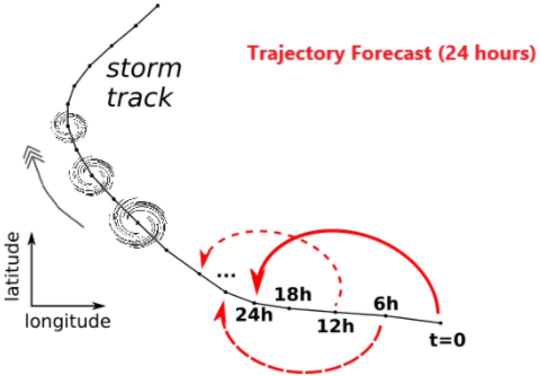
3 Method
- Data Description
- The proposed method

4 Experiments
- Selecting Network Configurations
- Comparison with the Existing Forecasting Models

5 Discussion

Problem Setting

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



Previous Works

- A large family of previous methods used spatio-temporal features
  - Two-stream Convolutional Network [Simonyan and Zisserman, 2014a]
  - Convolutional LSTM [Xingjian et al., 2015]
  - Temporal Convolutional Networks [Lea et al., 2016]
- 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
- The proposed method

4 Experiments
- Selecting Network Configurations
- Comparison with the Existing Forecasting Models

5 Discussion

Data sources

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

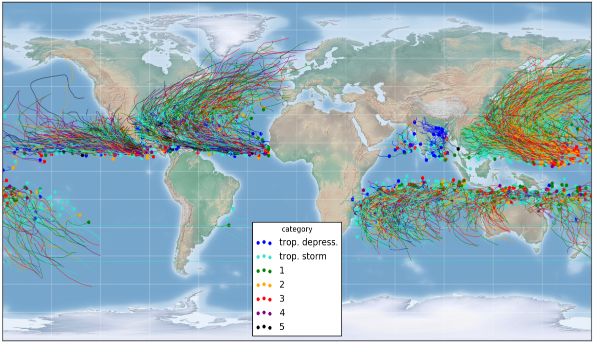


Figure: Database: tropical/extra-tropical storm tracks since 1979. Dots = initial position, colors = maximal storm strength according to the Saffir-Simpson scale.

Data sources

- **Reanalysis:** the latest global atmospheric reanalysis ERA-Interim produced by The European Centre for Medium-Range Weather Forecasts (ECMWF).
  - global atmospheric fields : u-wind, v-wind, pressure, temperature, humidity, vorticity ... every 6 hours
  - spectral resolution: about 80km on 60 vertical levels from the surface up to 0.1 hPa

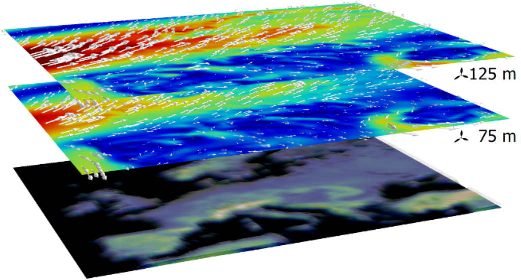


Figure: Reanalysis wind fields (Source: from website)

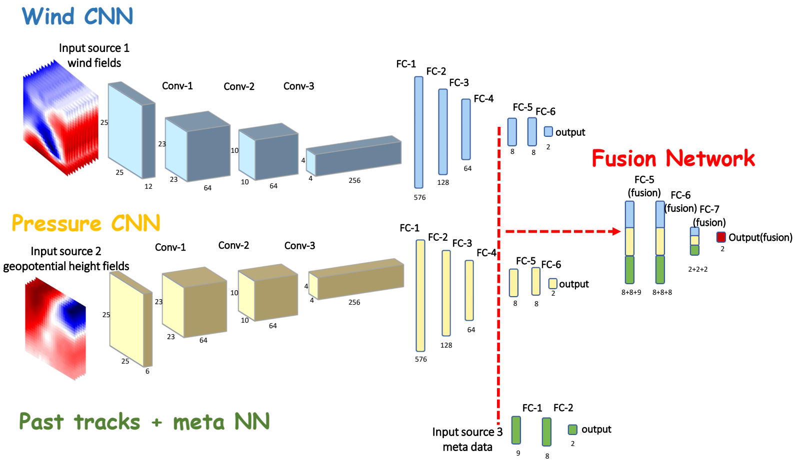
Input features

- **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), from consecutive previous - current time point **(3D tensor)**
- **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), from consecutive previous - current time point **(3D tensor)**

Input features

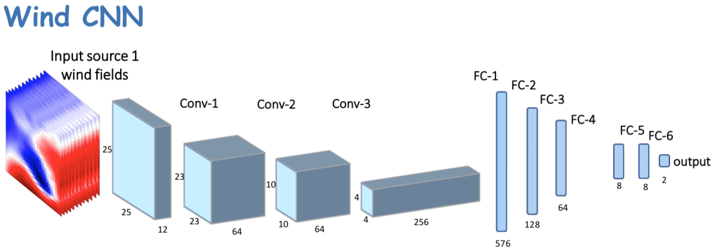
- **Storm track in history.** The past 6-hourly displacement in lat. and lon. **(1D tensor)**
- **Other hand-crafted features:** current lat. and lon., windspeed, Jday predictor(Gaussian function of "Julian day of storm init - peak day of the hurricane season" [DeMaria et al., 2005]), and current distance to land. **(1D tensor)**

- Three-stream fusion network



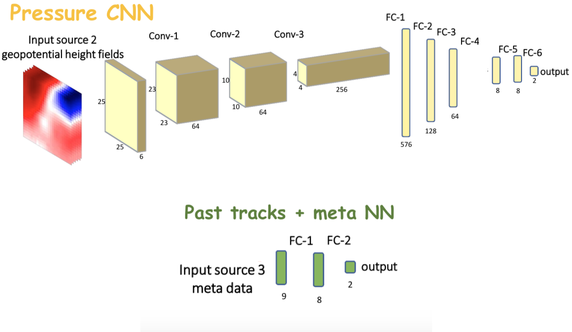
Train Stream Networks

- Stage I: Train stream networks
  - Wind CNN
    - pretty standard (popular) CNN architecture, like VGG Net [Simonyan and Zisserman, 2014b]
    - trying from shallow to deep



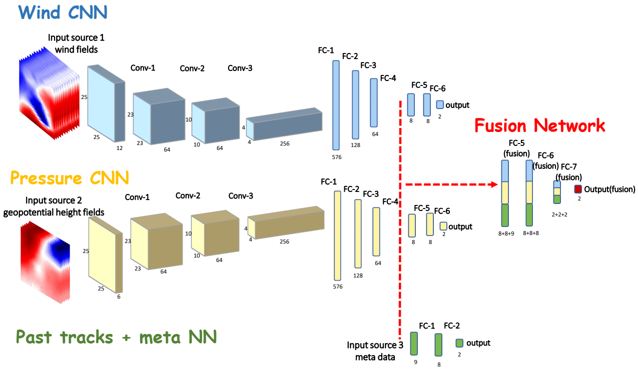
Train Stream Networks

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



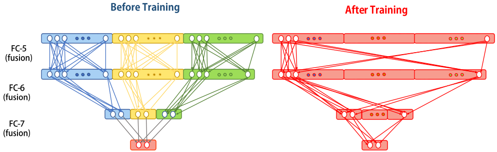
Train Fusion Network

- Stage II: Train the fusion network



Train Fusion Network

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



  - connections between different streams in fusion layers are added
  - two-phase optimization:
    - first phase: optimize weights only in fusion layers
    - second phase: optimize weights in all layers

Input preprocessing

- the entire dataset into train (60%) / validation (20%) / test (20%)
- separate all the storms into the three bins to avoid data leakage
- standardization

Loss function

- Stage I: train independently three stream networks

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

  - above showed Wind CNN loss function, analogical for other two stream networks.
  - δt: the prediction interval
  - $(x_w^n, x_p^n, x_{1d}^n, loc_t^n, loc_{t+\delta t}^n)\}$: labeled data
  - θ_w: Wind CNN parameters
  - F_{wind}: Wind CNN (input: wind fields $\mapsto$ output: expected displacement)
  - $M_{transform}$: transformation matrix from differences in latitude and longitude to distance in kilometers

Loss function

- Stage II: train the fusion network (2 phases)

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

  - θ_f: parameters of fusion layers
  - θ_s: parameters of stream layers
  - F_f: Fusion network (input: wind fields + pressure fields + 1d features (past tracks + meta) $\mapsto$ output: expected displacement)
- first phase: optimize only θ_f
- second phase: optimize on θ_f and θ_s

Algorithmic Details

Algorithmic details: ReLU, batch norm, adam, batch size=256, L2 $coeff. = 0.01$, fine-tuned initial learning rate, *He* initialization [He et al., 2015]... **No magic!**

Outline

1 Background
- Hurricane Track Forecast
- Hurricane Track Forecasting Models for Meteorologists

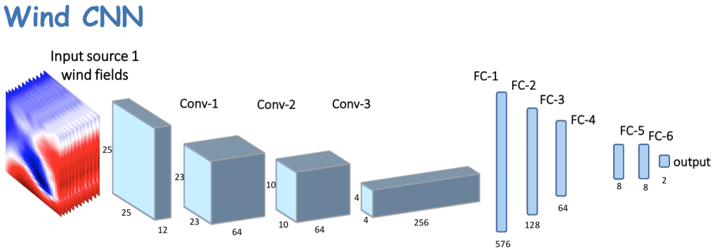
2 Preliminaries
- Problem Setting
- Previous Works

3 Method
- Data Description
- The proposed method

4 Experiments
- Selecting Network Configurations
- Comparison with the Existing Forecasting Models

5 Discussion

The Stream Networks



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

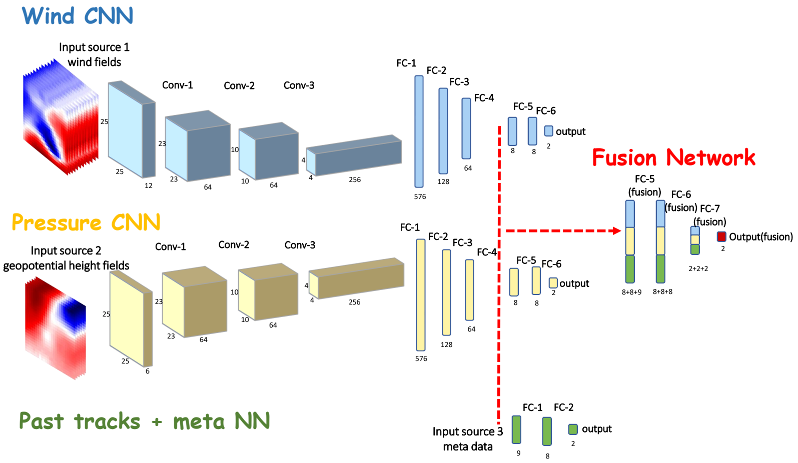
ConvNet Configurations			
A	B	C	D
7 layers	8 layers	9 layers	10 layers
input (25*25, 12 channels image)			
conv3-32 maxpool	conv3-32 conv3-32 maxpool	conv3-64 conv3-64 maxpool conv3-256	conv3-64 conv3-64 maxpool conv3-128 conv3-256 maxpool
FC-576			
FC-128			
FC-64			
FC-8			
FC-8			
FC-2			

- Evaluation results on 24 hours of storm track prediction on the validation set (**A,B,C,D from shallow to deep**)

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

- Three-stream fusion network



Three scenarios:

- how many layers should be fused?
- Does fusing three streams outperforms fusing two streams or using single stream?
- Do we need to pretrain the 3 stream networks before training the fusion networks?

- 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**.

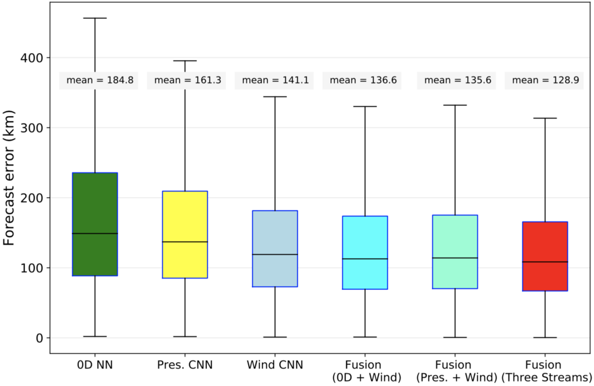
Model	Mean Square Error(km^2)	Mean Absolute Error(km)
Fus network, 3 streams, fuse 2 FC	25453.48	130.04
Fus network, 3 streams, fuse 4 FC	25628.22	130.27
Fus network, 3 streams, fuse 6 FC	25846.74	130.95
Fus network, 3 streams, fuse 6 FC + 1 Conv layers	25928.81	131.17

- Does fusing the 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.
  - compare also with naive fusion (no pretrain)
  - Result shown in the next slide

Experiments

Selecting Fusion Network Configuration

Figure: 24h-forecast results, in distance between predicted and real location.



naive fusion's forecasting error in distance: **142.63 km**, worse than best single stream alone!

Outline

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

Experiment

Comparison with the Existing Forecasting Models

- Compare with: A standard statistical model (BCD5)
- Also compare with: Official NHC forecast (OCFL).Ensemble methods, imporved over years

Experiments

Comparison with the Existing Forecasting Models

Quantitative:

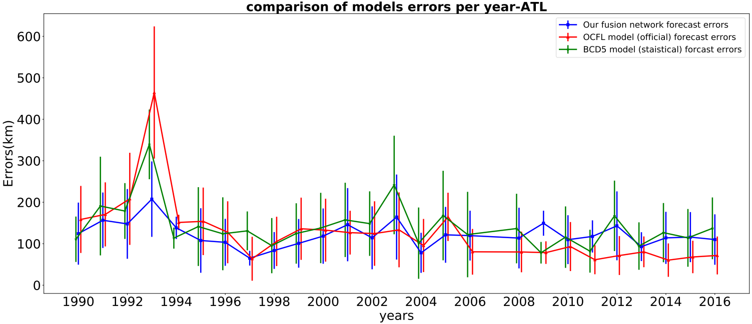


Figure: The yearly average 24-hours storm track forecasting errors (km) and standard deviation on the test set in **Atlantic** for our fusion network forecasts(blue), the BCD5 (green) and the OFCL (red), 1989-2016.

Quantitative:

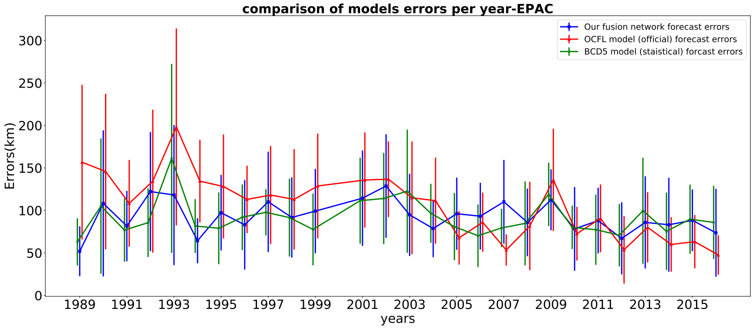


Figure: The yearly average 24-hours storm track forecasting errors (km) and standard deviation on the test set in **East Pacific** for our fusion network forecasts (blue), the BCD5 (green) and the OFCL (red), 1989-2016.

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 (best !)	69.95	94.03 (best !)	58.77

Qualitative:

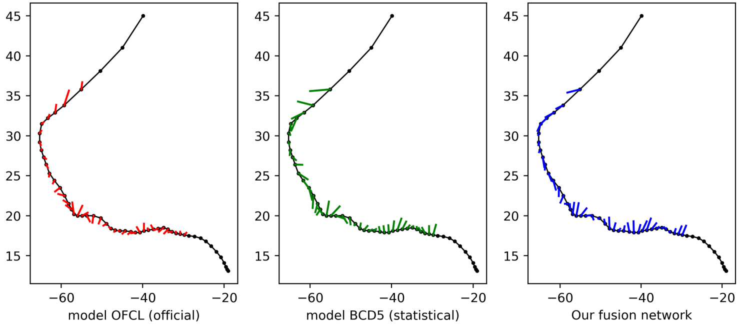


Figure: The forecast errors of the three models (left: the OCFL, middle: the BCD5, right: our fusion network) on the Hurricane Ian in 2016

Qualitative:

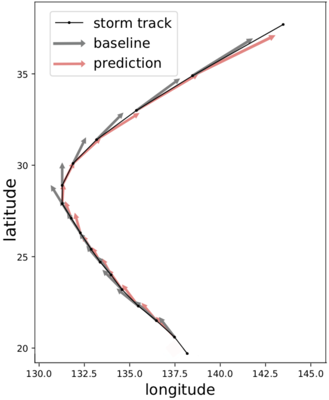


Figure: Example of 6h-forecasts on one storm track. The baseline prediction is equal to the last 6h-displacement (going straight).

Discussion

Conclusion:

- Proposed a promising deep learning framework for storm track forecasting
- Designed a fusion network consists of a three-stream convolutional neural network that can learn to fuse information from both atmopheric fields and history track
- 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

- work published in Climate Informatics Workshop Proceedings 2018, Sep 2018, Boulder, United States: 'Fused Deep Learning For Hurricane Track Forecast From Reanalysis Data' (same title)
- **advertisement:** Data challenge (Hackson) of Saclay Datascience Center about hurricane intensity forecast is going to be hold in Boulder, United States, Oct.

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.

Lea, C., Vidal, R., Reiter, A., and Hager, G. D. (2016).
Temporal convolutional networks: A unified approach to action segmentation.
In *European Conference on Computer Vision*, pages 47–54. Springer.

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. (2014a).
Two-stream convolutional networks for action recognition in videos.
In *Advances in neural information processing systems*, pages 568–576.

Simonyan, K. and Zisserman, A. (2014b).
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.