Report of master 2 research intership in computer science

Fused deep learning for hurricane track forecast from reanalysis data

Author :
Mo Yang

Internship supervisor :
Claire Monteleoni
Guillaume Charpiat

Host organisation :
Laboratoire de l'Accelerateur Lineaire
Inria

Secretariat - tel : 01 69 15 81 58
e-mail : alexandre.verrecchia@u-psud.fr

Contents

Abstract

Hurricanes is one of the most severe natural disaster that cost tremendous damage and death in each year. The forecast of incoming hurricanes trajectory can be very important for protecting people who live in shoreline areas and reducing their economic loss. In this report, we propose deep learning method for hurricanes trajectory forecasting. We introduce a fusion architecture. The proposed fusion architecture of neural network allows for processing different sources of data in parallel, including past hurricanes movements and reanalysis atmospheric wind fields and geopotential fields images. This fused network is trained to estimate the longitude and latitude 24h-forecast of hurricanes and depressions from a large database from both hemispheres (more than 3000 storms since 1979). It is demonstrated that the fused network which integrate all sources of data have significantly lower error distance than network that use any single source of data. (add compare with state-of-art)

Mots clefs

Hurricanes, Trajectory forecasting, Deep learning, Fusion architecture, Neural networks.

CHAPTER 1

Introduction

1.1 Introduction To Laboratory

1.2 Introduction To Groups

1.3 Internship Subject

This internship is **(where is the internship comes from...)** The objective of the internship is to study for using deep learning as alternative methodology for hurricane trajectory forecasting. We propose a neural network architecture taking into account past trajectory data and reanalysis atmospheric physical fields images. This fused network estimates the longitude and latitude 6h-72h forecast of hurricanes and depressions from both hemispheres and different basins (more than 3000 storms since 1979). We then compare our forecast results with results of NHC's models, including a representative early model BCD5 and their official model OCFL.

1.4 Structure

introduction to the structure of the article, the first chapter is what, the second...

Chapter 2

Background

2.1 Hurricane Trajectory Forecasting

Cyclones, hurricanes or typhoons are words for the same phenomena: a rare and complex event characterized by strong winds surrounding a low-pressure area. Hurricane is one of the most severe natural disaster that cost tremendous damage and death each year. During its life cycle a hurricane can expend as much as 10000 nuclear bombs [2]. Since 2000, over 45000 people were killed by hurricanes. In 2005, hurricane Katrina has cost as much as 125 billion US Dollars worth of damage. The explosion of the population also raises the risk of potential damage made by hurricanes[17]. In 2010, it was estimated that 1.53 billion people lived in hurricanes prone areas in 81 different countries and territories. 133.7 millions of people were exposed to hurricanes, the number has increased three times compared to the year 1970. This makes the forecast of hurricane trajectory crucial, in order for people living in that area to be alarmed and get some time to protect their goods and evacuate. Improving the prediction of hurricane movements has a large impact on society.

Formation of Hurricanes

The evolution and path of hurricanes depend on many factors at different scales and altitudes. All hurricanes are formed over warm ocean water near the equator. Warm and moist air over the ocean rises upward, which cause low-pressure area below. The air from high-pressure surrounding areas flows over into low-pressure areas and new air is heated and rise upward again. As warm and moist air rise and cools off, the water in the air forms clouds. Figure **??** shows the wind flows inside storm. The storm grows up and rotates as the cycle continues to be fed by ocean's heat and water evaporating from the surface. When the winds in the rotating storm reach 39mph(miles per hour, equivalent to 63km/h), the storm is called a "tropical storm". When the wind speeds reach 74 mph, the storm is officially called a "tropical cyclone", or hurricane. The Saffir–Simpson hurricane scale (SSHS) classifies hurricanes into 5 categories distinguished by the intensities of their sustained winds. The highest classification in the scale, category 5, consists of storms with sustained wind speeds higher than 156mph, See Figure2.1
A hurricane can be very unpredictable. The hurricane movement can perform loops, hairpin turns,

Saffir-Simpson Hurricane Scale		
Category	Wind Speed mph	Wind Speed 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

Table 2.1: The Saffir-Simpson Hurricane wind scale and Non-Hurricane Scale

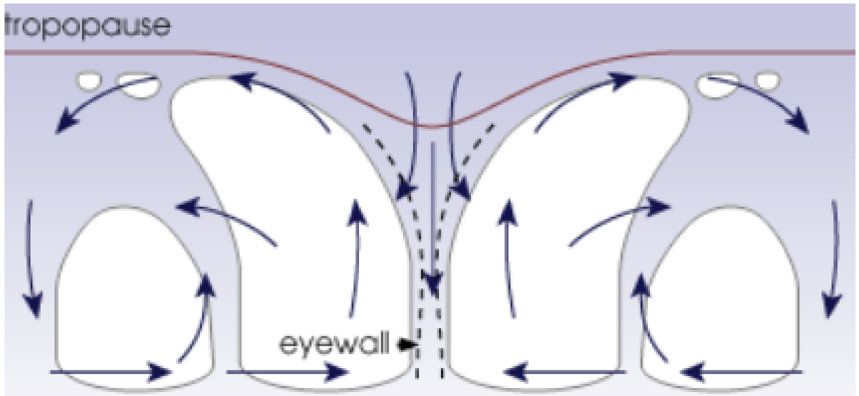

Figure 2.1: Graphic of wind flow inside storm [2]

and sharp curves. Its intensity can also change dramatically in a very short time. In 2005, hurricane Katrina was only classified as a tropical storm when it made its first landfall in Florida . After emerged into the Gulf of Mexico, it rapidly strengthened into a Category 5 hurricane in two days. Its second landfall finally caused the largest damage in history. The track and evolution of Hurricane Katrina is shown in Figure2.2 Global warming is also considered to influence hurricane activities. It is estimated that global warming will cause mode intense hurricanes globally in the coming century and have higher rainfall rates[12], which may lead to both more representative and more consistent error statistics for forecasting.

2.2 Forecast Method for Meteorologists

Tracks and intensity are the two main goals of the forecast. Today, the forecasts (track and intensity) are provided by a numerous number of guidance models [3]. In the 70's and 80's the models were statistical models based on historical relationships between storm behavior and various other parameters[6]. Today statistical models are mostly used for testing and comparing new models. With the increase of computing capacity and data assimilation techniques, dynamic models gradually replaced statistical models. Dynamical models solve the physical equations governing the motions in the atmosphere, but they are very complex and computationally demanding. These models normally run on high-speed supercomputers.

Guidance models are characterized as either early or late, depending on whether or not they are available to the Hurricane Specialist during the forecast cycle. [3] The late model can take hours to run. For example, if the NWS/Global Forecast System (GFS) runs at 12UTC, the result will not be available until 16UTC. The Dynamic models, in general, are late models. Although they can have a very precise forecast, the bottleneck is the computing speed. If the computation lasts for days, it doesn't make sense to make a prediction since the prediction will already be outdated. But fortunately, there is a technique that allows to take the latest available run of a late model and adjust its forecast to apply to the current synoptic time and initial conditions[3]. Current national forecasts are typically driven by consensus methods able to combine different dynamical models. The forecasts are usually calculated in order to have from 12h-predictions, 24h-predictions, 36h-predictions and up to 120h-predictions.

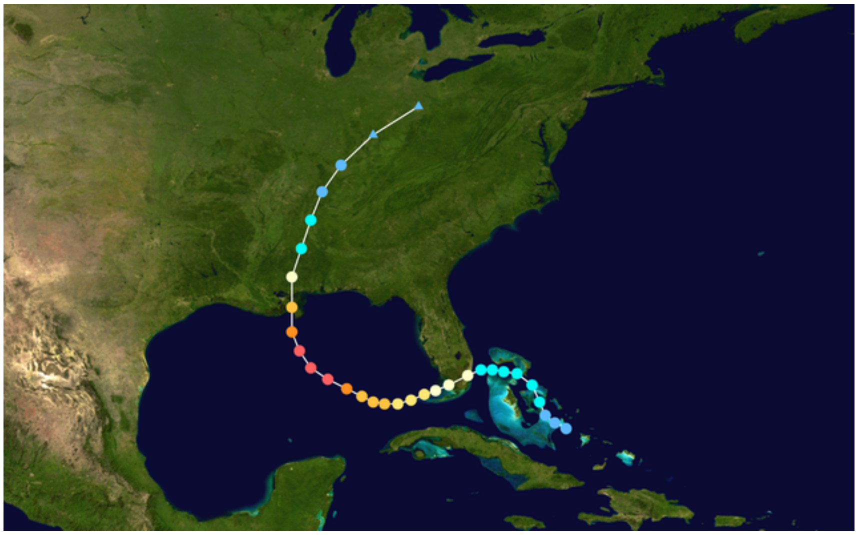

Figure 2.2: Hurricane Katrina track. Uses the color scheme from the Saffir-Simpson Hurricane Scale. The points show the location of each storm at six-hour intervals. The color represents the storm's maximum sustained wind speeds as classified in the Saffir-Simpson Hurricane Scale (see 2.1), and the shape of the data points represent the nature of the storm [1]

2.3 Convolutional Neural Network

The hurricane statistical forecasting models perform poorly with respect to the dynamical models, even if the database made of past hurricane is constantly growing. The machine learning methods, able for example to capture non-linearity and complex relations, have only been scarcely tested. Recent advances in deep learning have shown promising results on enormous pattern recognition tasks, such as ImageNet Large Scale Visual Recognition Challenge (ILSVRC) [19] [13] [22] and natural language processing [8] [21]. Convolutional Neural Network (CNN) is one of the most important advance in deep learning. CNN is a type of deep, feed-forward artificial neural networks. An example of CNN is presented in Figure 2.3 It is usually made of several conv layers followed by some fully connected layers. conv layers is inspired by animal's cortex visual system, where each neuron only process data for its receptive field. Between two successive conv layers there is usually also sub-sampling layers also known as pooling layers. The Advantage of CNN is that it learns both spatial and temporal differences in different scales and thus could extract features automatically without learning huge amount of parameters. The CNNs are widely adopted as a very effective model for analyzing images or images-like data for pattern recognition.

The CNN architecture has evolved over time. Modern CNNs tend to be deep with a large number of hidden layers. AlexNet is seen as a breakthrough when it won the 2012 ImageNet LSVRC-2012 competition by a large margin (reached 15.3% compared to the former best 26.2% in error rates on top-5 classification problem). AlexNet has 5 conv layers followed by 3 fully-connected layers. Researchers continually introduce new architectures of CNNs that are deeper, more complex and have better precision on ImageNet. The CNNs, as deep learning in general, is an empirical construction of a learning algorithm. Some works in CNNs show that CNNs should have more layers in order for the hierarchical representation of visual data to work[10]. VGG Net[20] is one of the most influential work that gives a guideline to design a CNN architecture. A VGG Net applies

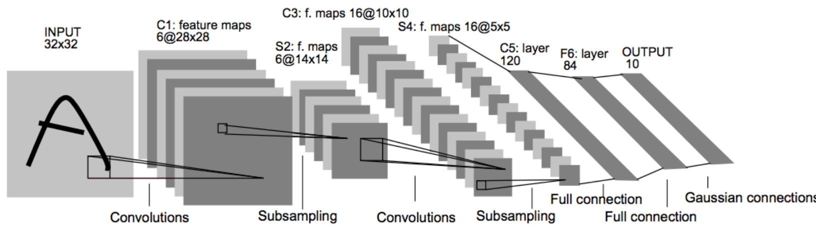

Figure 2.3: Architecture of LeNet, shows a series of layers in CNN that learn hierarchical features representation [14]

alternatively conv layers and Maxpooling layers through the whole network, and put several fully connected layers at the end of the network to generate output. Rather than using relatively large receptive fields in the first conv layers (e.g. 11x11 with stride 4 in AlexNet[13]), the author suggests to use very small 3x3 receptive fields throughout the whole network. The reason is that there could be more non-linear layers and smaller number of parameters, and effectively to have scopes from small to large receptive fields. It turns out that the more important for VGG Net becomes tunning the depth of the network without worrying about other hyper-parameters like kernel size. It is the simplicity and effectiveness of VGG Net that makes it among the most popular CNN architectures.

CNN is known to be tedious to train. To build an excellent CNN, researchers have proposed different methods. Some of them are crucial to successfully train a CNN. Training a CNN with huge number of learnable parameters usually require of large number of examples. The data Augmentation methods such as flipping or randomly cropping can effectively increase the network's performance when there are only limited number of examples. Proper input data normalization and weights initialization are important steps to accelerate convergence during the training. To realize it, the input data can be zero-centered and normalized at each channel to have a similar distribution. Weights are usually initialized by random weights drawn by Gaussian distributions[13], but it can lead to a very slow convergence when training very deep CNN. A paper[7] shows that the reason is because the norm of the outputs in each layer does not initially sum to 1, if the model have too many layers the norm will be exploded. The author propose a method called 'Xavier' initialization to ensure the outputs in each layer are initially standard normally distributed. 'Xavier' initialization is based on an assumption that activations are linear, which is not valid on rectifier nonlinearities(ReLU). Another paper[9] suggests a more robust initialization that deals with the rectifier nonlinearities. Batch Normalization(BN) layers are also recommended to be used to maintain the distribution of the outputs in each layer during training and thus to accelerate convergence[13].

2.4 CNN in climate science

Artificial Neural Network have recently shown their efficacy in a various number of forecasting tasks. Particularly, convolutional neural networks (CNN) have raised attention as they are suited for large imaging data. A convolutional LSTM model has been used for precipitation forecast in a promising study [23]. Another recent study predicts the evolution of sea surface temperature maps by combining CNN with physical knowledge [4]. The CNN has also been used for the detection of extreme weather like hurricanes from meteorological variables patches. [18]. These studies show the strong potential on various climate problems.

To our knowledge, only two preliminary studies have tackled the hurricane forecast tracking using machine learning: the first one uses random forests on local reanalysis histograms [15], however it was tested only on tropical storms in 2015 in North Atlantic, and the mean error of 6h-forecasts seem to indicate poor results (more than 60km). The second uses a sparse recurrent neural network from trajectory data [16], it was tested on only 4 hurricanes and seem also to give large distance errors (mean 6h-forecast error is 72km). We have noticed that Deep Learning as a generative method has not been tested before. In this study, we expect to explore Deep Convolutional Neural Network to be a new Hurricane Forecasting method, and hopefully could improve Hurricane Trajectory Forecasting results.

2.5 Fusion

CHAPTER 3

Method

3.1 Problem Setting

The goal of the research is to predict the trajectory of Hurricane (**forecast cycle not yet decided**) using the information from the past storms since 1979. The task is shown in Figure **??**. We aim at building an end-to-end model using two types of data (reanalysis and hand-crafted features), for each time step of each storm, we want to independently predict its displacement in the future(depending on forecast cycle). In this chapter, first we are going to describe the sources of data, then we will show how we design the neural network for each source of the data, last we show that how we improve results with a fusion architecture of Neural Network.

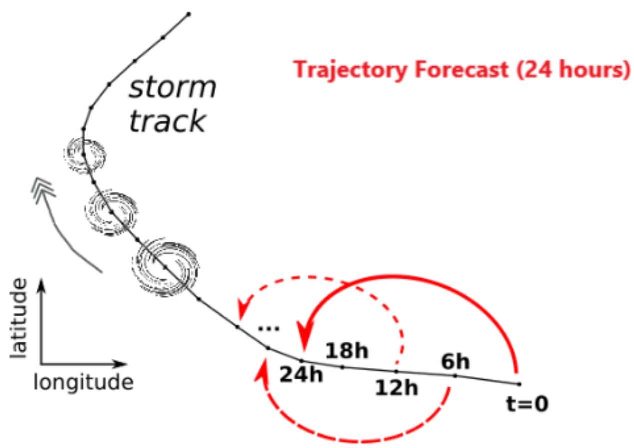

Figure 3.1: General architecture: the two types of data are feeding two neural networks trained separately. The final fused network is re-trained before predicting the forecast track.

3.2 Data Description

The raw storm track data used in this study is composed of more than 3000 extra-tropical and tropical storm tracks since 1979 extracted from the NOAA database IBTrACS[11], see Figure 3.2. The tracks are defined by the 6-hourly center locations (latitude and longitude). They come from both hemispheres and the number of records per storm varies from 2 to 120 time steps. In total, the database counts more than 90 000 time steps.

Reanalysis

The trajectory of a storm depends on large scale atmospheric flows. Thus, we extracted the wind fields of the neighborhood of the storm at every time step t from the ERA-interim reanalysis database [5]. Specifically, we extracted the u-wind and v-wind fields on a $25x25$ degree grid centered on the current storm location, at 3 atmospheric pressure levels (700 hPa, 500 hPa and 225 hPa). Because we wanted to capture the dynamics, we also extracted the wind fields measured at $t - 6h$ at the same locations.

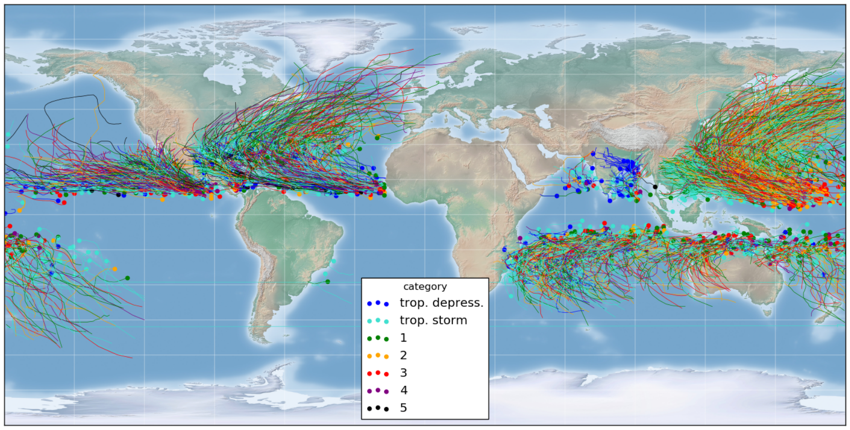

Figure 3.2: Database: more than 3000 tropical/extra-tropical storm tracks since 1979. Dots = initial position, colors = maximal storm strength according to the Saffir-Simpson scale.

The choice of the 3 pressure levels was driven by statistical forecast models [6]. The reason why we focused on the wind parameter is that we applied a sparse feature selection technique (Automatic Relevance Determination, based on linear regression) over all available reanalysis fields, which highlighted the usefulness of wind.

3.3 CNN Configuration for Wind Fields and Geopotential Fields

Convolutional neural networks (CNN) are suited for non-linear learning with image-like data. They have already shown their efficiency in the climate informatics field [23, 4, 18]. The centered wind fields at different pressure levels at t and $t - 6h$ can be seen as 12 images of size 25x25. We used as a guideline a typical CNN architecture alternating convolutional layers and max-pooling layers and added several fully connected layers at the end of the network [20]. To measure the improvements brought by increasing the CNN depth, we have designed 4 CNNs with the same number of neurons and varying depths.

3.4 Neural Network for Past Tracks

Another important source of information is the previous displacements (latitude and longitude for $t - 12h$ and $t - 6h$). We designed a small neural network (two small fully connected layers) able to learn the future track from this past track.

3.5 Fusion Architecture for All Sources of Data

Because of the different nature of the wind field image and of the past track data, it is not straightforward to mix them as a common input to a bigger network. Instead, we first train separately the wind field CNN and the small past track neural network (NN) previously mentioned, and then we fuse their two last layers, and re-train them together (see Figure ??).

Table 3.1: **ConvNet configurations**(shown in columns). The depth of the configurations increases from left(A) to the right(D), as more layers are added. The convolutional layer parametres are denoted as "conv(kernel size)-(number of output channels)". The ReLU activation layer and Batch normalization layers are not shown in figure

ConvNet Configurations			
A	B	C	D
7 layers	8 layers	9 layers	10 layers
input (25*25, 12 channels image)			
conv4-32 maxpool	conv4-32 conv3-32 maxpool	conv4-64 conv3-64 maxpool conv3-256	conv4-64 conv3-64 maxpool conv3-128 conv3-256 maxpool
FC-576			
FC-128			
FC-64			
FC-8			
FC-8			
FC-2			

Table 3.2: Number of parameters (in millions)

Network	A	B	C	D
Number of parameters	2.27	2.33	2.75	2.67

CHAPTER 4

Experiments and Result Analysis

4.1 Evaluation Settings

4.2 Training Details

The storms were randomly separated in 3 sets as follows: train (60%) / valid (20%) / test (20%). Then, within each set, all time instants were treated independently. As a loss function (quantity to optimise), we used the mean square error (MSE) in kilometers between the forecast and the true storm location at $t+6h$. We added an L2 penalty on the weights of the model ($coef. = 0.01$). The training was performed by the Adam optimizer.

Our implementation uses PyTorch 4.0. The training and testing took less than 1 hour on 4 TitanX GPUs with data parallelism [**?**].

4.3 Results

Figure 4.1 shows the 6h-forecast results on the test set in absolute distance error. We define the baseline prediction as equal to the last displacement (from $t-6h$ to t). We can see the improvement of fusing networks (mean error $\bar{e} = 32.9km$) with respect to the wind field CNN alone ($\bar{e} = 40.7km$) or the track neural network alone ($\bar{e} = 35km$). We have plotted in Figure 4.2 an example of 6h-forecasts on one storm track for the baseline and for our prediction (fusion networks). Our forecast predicts well, even in the case of change of direction or speed.

If these results are promising, some more long-term predictions are needed for a practical use. Moreover, current forecast models do not provide less than 24h-forecasts, which prevents us from comparing the results. With respect to the existing machine learning studies predicting 6h-forecasts [15, 16], we tend to perform better (error larger than 60km for both studies) and on a larger/more diverse dataset. Moreover, if we only look at hurricane time steps (without depressions), our mean prediction error drops to 25.8km. Depressions seem to be more difficult to predict: an explanation can be that they are smaller and more subject to local perturbations.

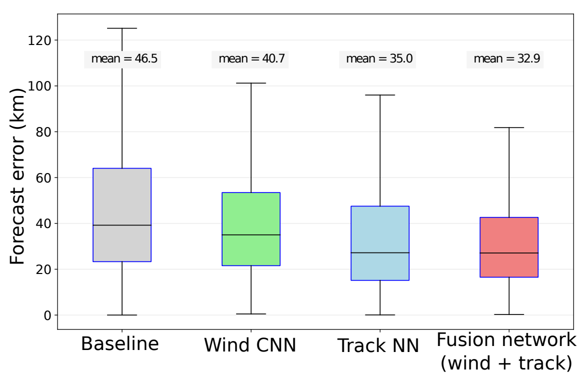

Figure 4.1: 6h-forecast results on the test set (storms coming from all oceanic basins), in distance between predicted and real location. Baseline = previous displacement (going straight).

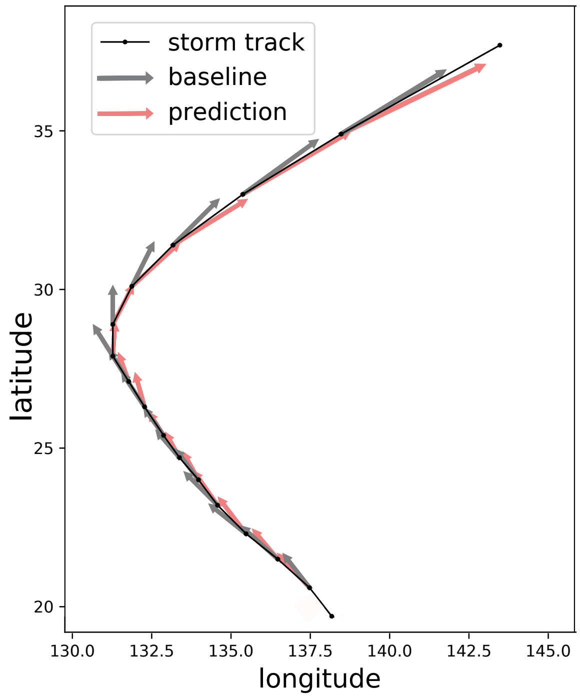

Figure 4.2: Example of 6h-forecasts on one storm track. The baseline prediction is equal to the last 6h-displacement (going straight).

4.4 Result Analysis

Quantitative

Qualitative

Chapter 5

Conclusion and Discussion

We showed a promising deep learning framework for storm track forecasting. We demonstrated the benefit of coupling two types of data (past tracks and wind fields) in an efficient fusion model. Our results on a large database (90 000 time steps from 3000 storms) from different oceanic basins are showing 6h predictions with less than 33km error. Moreover, the error on only hurricane data points (without depressions) drops to 25.8km. We think that the use of such deep learning methods can help the current forecast modellers by providing a complementary prediction that could be integrated in some consensus methods.

Acknowledgment

many thanks to ...

Bibliography

[1] Hurricane katrina track. https://en.wikipedia.org/wiki/Meteorological_history_of_Hurricane_Katrina#/media/File:Katrina_2005_track.png. Accessed: 2018-07-31.

[2] Hurricanes: The greatest storms on earth. https://earthobservatory.nasa.gov/Features/Hurricanes. Accessed: 2018-07-31.

[3] Nhc track and intensity models. https://www.nhc.noaa.gov/modelsummary.shtml. Accessed: 2018-07-04.

[4] Manu de Bezenac, Arthur Pajot, and Patrick Gallinari. Deep learning for physical processes: Incorporating prior scientific knowledge. *arXiv preprint arXiv:1711.07970*, 2017.

[5] Dick P Dee, S M Uppala, AJ Simmons, Paul Berrisford, P Poli, S Kobayashi, U Andrae, MA Balmaseda, G Balsamo, d P Bauer, et al. The era-interim reanalysis: Configuration and performance of the data assimilation system. *Quarterly Journal of the royal meteorological society*, 137(656):553–597, 2011.

[6] Mark DeMaria, Michelle Mainelli, Lynn K Shay, John A Knaff, and John Kaplan. Further improvements to the statistical hurricane intensity prediction scheme (ships). *Weather and Forecasting*, 20(4):531–543, 2005.

[7] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256, 2010.

[8] Yoav Goldberg and Omer Levy. word2vec explained: deriving mikolov et al.'s negative-sampling word-embedding method. *arXiv preprint arXiv:1402.3722*, 2014.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1026–1034, 2015.

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[11] Kenneth R Knapp, Michael C Kruk, David H Levinson, Howard J Diamond, and Charles J Neumann. The international best track archive for climate stewardship (ibtracs) unifying tropical cyclone data. *Bulletin of the American Meteorological Society*, 91(3):363–376, 2010.

[12] Thomas R Knutson, Joseph J Sirutis, Gabriel A Vecchi, Stephen Garner, Ming Zhao, Hyeong-Seog Kim, Morris Bender, Robert E Tuleya, Isaac M Held, and Gabriele Villarini. Dynamical downscaling projections of twenty-first-century atlantic hurricane activity: Cmip3 and cmip5 model-based scenarios. *Journal of Climate*, 26(17):6591–6617, 2013.

[13] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.

[14] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, 1998.

[15] Sterenn Marie Liberge, Sileye Ba, Philippe Lenca, and Ronan Fablet. 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, 2011.

[16] Mina Moradi Kordmahalleh, Mohammad Gorji Sefidmazgi, and Abdollah Homaifar. 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, 2016.

[17] Pascal Peduzzi, Bruno Chatenoux, H Dao, Andréa De Bono, Christian Herold, James Kossin, Frédéric Mouton, and Ola Nordbeck. Global trends in tropical cyclone risk. *Nature climate change*, 2(4):289, 2012.

[18] Evan Racah, Christopher Beckham, Tegan Maharaj, Samira Ebrahimi Kahou, Mr Prabhat, and Chris Pal. Extremeweather: A large-scale climate dataset for semi-supervised detection, localization, and understanding of extreme weather events. In *Advances in Neural Information Processing Systems*, pages 3402–3413, 2017.

[19] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015.

[20] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[21] Ilya Sutskever, Oriol Vinyals, and Quoc V Le. Sequence to sequence learning with neural networks. In *Advances in neural information processing systems*, pages 3104–3112, 2014.

[22] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015.

[23] SHI Xingjian, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems*, pages 802–810, 2015.