5th International Workshop on Climate Informatics
September 24-25, 2015
Hosted by the National Center for Atmospheric Research in Boulder, CO

HURRICANE TRAJECTORY PREDICTION VIA A SPARSE RECURRENT NEURAL NETWORK

Mina Moradi Kordmahalleh[1], Mohammad Gorji Sefidmazgi[1], Abdollah Homaifar[1], Stefan Liess[2]

***Abstract*—A proposed sparse recurrent neural network with flexible topology is used for trajectory prediction of the Atlantic hurricanes. For prediction of the future trajectories of a target hurricane, the most similar hurricanes to the target hurricane are found by comparing directions of the hurricanes. Then, the first and second differences of their positions over their life time are used for training the proposed network. Comparison of the obtained predictions with actual trajectories of Sandy and Humberto hurricanes show that our approach is quite promising for this aim.**

I. MOTIVATION

Hurricanes are severe tropical cyclones with wind speeds exceeding 75 miles per hour. They originate in the tropical regions of the Atlantic Ocean or Caribbean Sea, usually travel North, Northwest or Northeast from their points of origin, and bring heavy rains and dangerous tides[1]. Potential destruction of hurricanes has created anxiety about the accuracy of tracking.

Hurricane track forecasting is a complex problem with significant societal importance [2]. Over past decades, major understanding of the issue has been achieved. Scientists have become interested to improve the capability of predictive models and track hurricanes for the next few hours or days.

In [3], an autoregressive model for increments in track of speed and direction has been presented. In[4], an autoregressive model has been used to model the latitudinal and longitudinal increments (rather than the velocity increment) of powerful tropical cyclones in the Coral Sea near northeastern Australia.

In this paper, tracking of the Atlantic tropical hurricanes is performed by our proposed sparse Recurrent Neural Network (RNN) with flexible topology [5]. RNNs are capable of identifying the temporal patterns within a system, and are appropriate tools for time series prediction. However for complex systems with few available observations, sparse networks (i.e. non-fully connected networks) are more appropriate to avoid over fitting that is an inherent problem. In the rest of this paper, the data and the developed technique for trajectory prediction are discussed. Prediction results for two case studies from the Atlantic tropical hurricanes are illustrated at the end.

Corresponding author: A. Homaifar, homaifar@ncat.edu, [1] North Carolina A&T State University, [2] University of Minnesota

II. DATA

Accuracy of such works is closely tied to data. NOAA has tracked the Atlantic tropical hurricanes and storms since 1851. During the life time of hurricanes, NOAA provides the information including the position in latitude and longitude, the maximum wind speed, and the central pressure which are recorded every six hours. In this work, hurricanes Sandy and Humberto that have occurred at 2012 and 2013 respectively are two target hurricanes predicted by the proposed method. For similarity analysis, the hurricanes which have occurred in period of 1990 to 2013 are used.

III. METHOD

In this paper, to find more accurate patterns of an ongoing target hurricane, the most similar hurricanes in the dataset are used for training the sparse RNN. Techniques for evaluating the similarity between time series have long been of interest to the database community. To capture the similarity between two time series, the choice of distance function is a critical issue [6].

For hurricane i with longitude $x^i(t)$ and latitude $y^i(t)$ at time t, the first difference of the position, $[\Delta x^i(t), \Delta y^i(t)]$, is equal to $[x^i(t) - x^i(t-1),\ y^i(t) - y^i(t-1)]$. Direction of the hurricane i called $[v_x^i(t), v_y^i(t)]$ is considered as:

$$[\frac{\Delta x^i(t)}{\sqrt{\Delta x^i(t)^2 + \Delta y^i(t)^2}}, \frac{\Delta y^i(t)}{\sqrt{\Delta x^i(t)^2 + \Delta y^i(t)^2}}] \quad (1)$$

The Euclidean measure ($L2$ norm) which sums the Euclidean distance between direction of two hurricanes over time can be considered as a distance function. However, considering time shifts or lags of two time series may provide more accurate interpretation of the distance. Distance between hurricanes i and j by considering their first K observations is:

$$D = \sum_{t=I}^{n} (v_x^i(t), v_x^j(t+r))^2 + (v_y^i(t), v_y^j(t+r))^2 \quad (2)$$

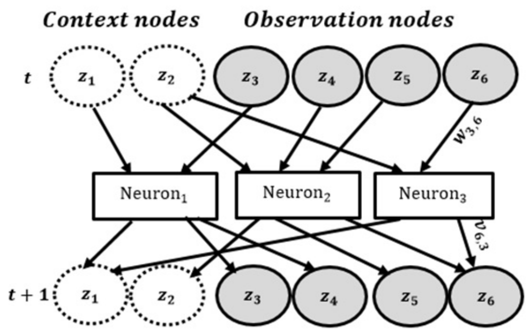

Fig. 1. Sparse Recurrent Neural Network with flexible topology.

where n, I, and r have to satisfy $I \geq 2, n + r \geq K$, and $n - I \geq K/2$. The similarity of the target hurricane i to j^{th} hurricane is equal to exponential of minus of the distance. For prediction of a target hurricane, the top 50 percent most similar hurricanes to the target hurricane are selected and used for training the sparse RNN.

The proposed network, as shown in Fig.1, consists of a set of processing units (i.e. neurons) which are sparsely connected to observation nodes and context nodes. Context nodes are without experimental measurements and play the role of memory or internal states of the model. Observation nodes consist of the first difference of the position, $[\Delta x^i(t), \Delta y^i(t)]$, and the second difference of the position, $[\Delta v_x^i(t), \Delta v_y^i(t)] = [\Delta x^i(t) - \Delta x^i(t-1), \Delta y^i(t) - \Delta y^i(t-1)]$. These four measurements $[\Delta x^i, \Delta y^i, \Delta v_x^i, \Delta v_y^i]$ are computed for all the similar hurricanes and feed to the network during the training process.

A customized Genetic Algorithm (GA) is developed to train the network. Through the evolutionary process of the GA, the topology of the network including the incoming and outgoing connections of the neurons along with their corresponding connection weights are trained using guidance of a fitness function. At each generation of the GA, the observation and context nodes (z_i) which are outgoing connections of at least one neuron are updated by Eq. 3, and context nodes without any connections hold their previous values.

$$z_i\,(t+1) = \sum_k v_{i,k}.g\left(\sum_j w_{k,j}.z_j\,(t)\right) \qquad (3)$$

where $w_{k,j}$ is the incoming connection weight from $z_j(t)$ to k^{th} neuron. The $v_{i,k}$ is the outgoing connection weight from k^{th} neuron to $z_i(t+1)$. To induce non-linearity in the model, g is considered as a hyperbolic tangent activation function.

IV. EVALUATION

In this work, the first 10 observations of the hurricanes are used for similarity analysis. Hurricane Sandy was one

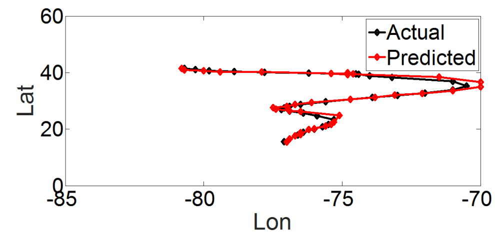

Fig. 2. Trajectory tracking of Sandy hurricane at 2012.

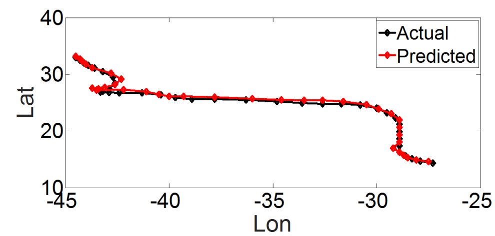

Fig. 3. Trajectory tracking of Humberto hurricane at 2013.

of the costliest natural disasters in U.S. history due to its excessive wind field, large storm surge, and unusual track into the population-dense Northeast corridor. Results of the prediction over test sets (unseen data in the training process) for both hurricanes Sandy and Humberto are shown in Figs. 2 and 3 respectively. In our future works, we expect that by completing the data archive with hurricanes that have occurred before 1990, considering more historical observations for the similarity analysis, and applying Euclidean measure to the spherical coordinates instead of lat-lon coordinates, accuracy of the prediction will be improved.

ACKNOWLEDGMENTS

This work is supported by the National Science Foundation under cooperative agreement No. CCF-1029731.

REFERENCES

[1] B. Taranath, *Reinforced Concrete Design of Tall Buildings*. CRC Press, 2009.

[2] T. M. Hall and S. Jewson, "Statistical modelling of north atlantic tropical cyclone tracks," *Tellus A*, vol. 59, no. 4, pp. 486–498, 2007.

[3] P. Vickery, P. Skerlj, and L. Twisdale, "Simulation of hurricane risk in the us using empirical track model," *Journal of structural engineering*, vol. 126, no. 10, pp. 1222–1237, 2000.

[4] M. James and L. Mason, "Synthetic tropical cyclone database," *Journal of waterway, port, coastal, and ocean engineering*, vol. 131, no. 4, pp. 181–192, 2005.

[5] M. Moradi Kordmahalleh, M. Gorji Sefidmazgi, A. Homaifar, D. B. KC, and A. Guiseppi-Elie, "Time-series forecasting with evolvable partially connected artificial neural network," in *GECCO2014*, pp. 79–80, ACM, 2014.

[6] M. D. Morse and J. M. Patel, "An efficient and accurate method for evaluating time series similarity," in *Proceedings of the 2007 ACM SIGMOD international conference on Management of data*, pp. 569–580, ACM, 2007.