LONG-LEAD PREDICTION OF EXTREME PRECIPITATION CLUSTER VIA A SPATIO-TEMPORAL CONVOLUTIONAL NEURAL NETWORK

Yong Zhuang[1], Wei Ding[1]

***Abstract*—A reliable long-lead (5-15 days ahead) prediction of extreme precipitation cluster is vitally important for regional flooding forecasting. A significant research effort is to develop methods for making long-lead flood forecasts using machine learning techniques,as current physics-based numerical simulation models can be extremely complex to account for compounding uncertainty in measurements and modeling. Accurate precipitation forecasts by numerical weather prediction models are limited to a few days lead-time, because non-linearity in the governing equations of the atmosphere creates a sensitive dependence on initial conditions. We design a novel Spatio-Temporal Convolutional Neural Network (ST-CNN) to fully utilize the spatial and temporal information and automatically learn underlying patterns of precipitation precursors from data for extreme precipitation cluster prediction. We validate the ST-CNN model using 62 years historical precipitation data collected in the State of Iowa, USA, from 1948-2010.**

I. INTRODUCTION

According to the U.S. Geological Survey [1] , floods were the number-one natural disaster in the United States in terms of number of lives lost and property damage during the 20th century. Regional flooding is often produced by long sequences of slowly moving, low-pressure or frontal storm systems including decaying hurricanes or tropical storms accruing over periods of several days to several weeks. A reliable long-lead (5-15 days ahead) prediction of extreme precipitation event is vitally important for mitigating flood damage. Accurate precipitation forecasts by numerical weather prediction models are limited to a few days lead-time because the non-linearity in the governing equations of the atmosphere creates a sensitive dependence on initial conditions that causes an effort in the initial conditions to double after just a few days, thus making long-range forecasts (longer than 7 days) practically impossible. Understanding the future trend of climate requires accurately identifying the precipitation precursors. Long-lead predictions have to consider variables in a long time period and large spatial neighborhoods, which involves an enormous amount of potentially influencing variables.

The goal of this study is to integrate machine learning and data mining methods with hydrological science and atmospheric science to detect interesting spatio-temporal patterns from this huge feature space to improve long-lead forecasting of extreme precipitation events. We design and implement a new Spatio-Temporal Convolutional Neural Network (ST-CNN) model to automatically learn the dependency of meteorological variables on spatio-temporal neighborhoods and summarize the patterns of local neighboring groups of neurons, to predict heavy precipitation cluster in 10 days ahead. We evaluate ST-CNN using 62 years historical meteorological data collected in the State of Iowa, USA.

II. RELATED WORK

Over the last three decades, a great deal of attention in statistics and machine learning has been directed toward extreme weather prediction[2] [3]. Most of them rely on meteorological inputs that usually come from observation networks and radar [4], and require a complex and meticulous simulation of the physical equations in the atmosphere model.

In recent years, machine learning feature selection methods, which aim to select a subset of relevant features from an original feature set, often at a scale of millions of features, have become popular in climate research for constructing forecasting models. For instance, Wu et al. used Online Streaming Feature

{yong.zhuang001, wei.ding}@umb.edu [1]Department of Computer Science, University of Massachusetts Boston, Boston, MA

Selection (OSFS) for heavy precipitation prediction [5][6][7]; Wang et al. applied the fast-OSFS algorithm for extreme flood forecasting [8][9]. Although feature selection methods can simplify forecasting models for easier interpretation and time efficiency, the approach usually requires domain scientists to provide initial feature sets that are closely related with the problem domain.

Neural-network-based machine learning approaches have been very successfully used for detecting high level patterns from raw low-level features without much intervention with prior domain knowledge [10]. Anctil et al. used artificial neural network (ANN) technique to forecast rainfall [11]. The results show that ANN forecasting models can get superior results to those obtained by linear regression models. More recently, Shi et al. implemented a new convolutional long short term memory (LSTM) deep neural network for precipitation nowcasting [12]. This model is trained on two dimensional radar map time series data. Their study showed that deep networks reveal a great potential on various climate problems. In our study, we explore a new Convolutional Neural Network architecture to learn patterns from meteorological variables in spatio-temporal grids for the long-lead prediction of extreme precipitation clusters.

III. METHOD

We formulate the long-lead prediction of extreme precipitation cluster as a classification problem with multiple spatio-temporal tensor data as inputs.

A. Spatio-Temporal Tensor Features

If we use the positions of cells in a data matrix to represent spatio grids, then one observed variable over a spatial region of m by n can be listed in a $m \times n$ matrix, which consists m rows and n columns. Then the matrices of k variables, which are collected at the same time t, can be stacked as a $m \times n \times k$ cuboid Q_t. If the observations are recorded periodically, we get a sequence of cuboids $Q_{t_1}, Q_{t_2}, ..., Q_{t_q}$ (in this study, we use $q = 10$ days records of 9 meteorological variables over a 32 by 32 region, and the size of cuboid is $32 \times 32 \times 9$). Thus, the multiple spatio-temporal sequences can be represented by a tensor $\chi \in \mathbb{R}^{m \times n \times k \times q}$. Then the long-lead ($x$ time-stamps ahead) prediction of extreme precipitation cluster problem can be formulated as follows:

$$P(C_{t_{q+x}}) = \operatorname{argmax} P(C_{t_{q+x}} | Q_{t_1}, Q_{t_2}, ..., Q_{t_q}) \quad (1)$$

Here $C_{t_{q+x}}$ denotes the class label of the $(t_{q+x})^{th}$ time-stamp, and the objective function selects the mostlikely outcome class (extreme precipitation cluster vs. non-extreme precipitation cluster) in a x time-stamps lead time given previously known spatio-temporal tensor data sequences.

B. Spatio-temporal Data Analysis

Our goal is let a CNN model automatically identify interesting and physically meaningful spatio-temporal patterns from data for precipitation cluster precursor identification. Certain pre-cursor patterns in the synoptic domain may indicate the development and movement of strong storms, including the location of fronts or strong horizontal temperature gradients, the presence of an upstream trough / ridge axis or a strong jet streak or a change in low-level winds.

C. The Spatio-Temporal Convolutional Neural Network

A Convolutional Neural Network (CNN) architecture is usually formed by a stack of distinct layers that transform the input volume into an output volume through a differentiable function. In this study, we build our ST-CNN architecture with six layers, including two convolutional layers, two max polling layers, and two fully connected layers. The ST-CNN is designed to use those layers of neurons (learning units) to automatically detect very local and detailed representations of a broad class of patterns from tensor data at the convolutional layers, and then summarize those local patterns to build high level features in the max pooling layers. The configuration of our architecture is depicted in Figure 1. The output volume is a vector that includes the class scores of the binary class labels of the extreme precipitation clusters after the calculation of the fully connected layers. We use back propagation algorithm to search for minimum of loss function in weight space and apply L_2 regularization to prevent overfitting.

Convolutional Layer: In ST-CNN architecture, there are two convolutional layers, which consist of multiple 3D filters(kernels) with the size of 5×5 (in this project, we choose to detect local patterns in a 5 by 5 neighborhood in the convolutionary layer). The first convolutional layer has 10 filters and the second one has 15 filters. Each filter takes inputs from a cuboid section of the previous layer. The weights for this cuboid section are the same for each filter in the convolutional layer to reduce the number of neural network parameters to be learned. Thus, the convolutional layer is just a feature map convolution of the previous layer. The task of the convolutional layer is to automatically

6th International Workshop on Climate Informatics
September 22-23, 2016
CI 2016
Hosted by the National Center for Atmospheric Research in Boulder, CO

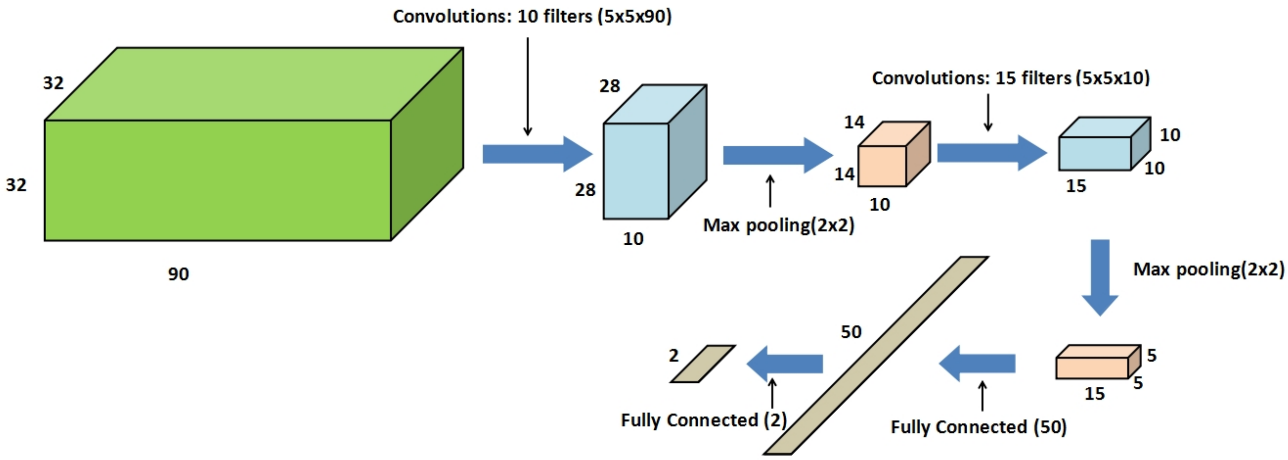

Fig. 1. The ST-CNN Architecture and its layer parameters. The ST-CNN contains two convolutional layers with filter size (5×5), two pooling layer with filter size (2×2), and two fully connected layer with 50 neurons and 2 neurons respectively. The input is a $32 \times 32 \times 90$ tensor which is associated with nine meteorological variables of 10 days over a 32 by 32 region. Nine meteorological variables are PW, T850, U300, U850, V300, V850, Z300, Z500, and Z1000 [13]. The final output is a vector that includes the class scores of the binary class labels (extreme precipitation cluster vs. non-extreme precipitation cluster).

learn local meaningful patterns that are associated with class labels.

Pooling Layer: Each convolutional layer is followed by a pooling layer which takes small cuboid blocks from the convolutional layer and sub-samples it to produce a single output from that block. In other words, each pooling layer is the summary of the patterns learned by convolutional layer. Here, pooling layers are max-pooling layers with 2×2 filters, which let the patterns of previous convolutional layers be reduced at half size and the outputs of adjacent pooling units do not overlap. The function of the pooling layer is to progressively reduce the spatial size of the representation to reduce the amount of parameters and computation in the network, and also to control over-fitting.

Fully Connected Layer: ST-CNN has two fully connected layers with 50 neurons and 2 neurons respectively. Each one takes all neurons in the previous layer and connects it to every single neuron it has. The final layer outputs the class score vector of the binary class label (extreme precipitation cluster vs. non-extreme precipitation cluster).

IV. EXPERIMENTS

A. Data

The dataset we used for experiments is the historical meteorological data collected in the State of Iowa, USA from January 1st, 1948 to December 31^{st}, 2010 [13]. A total of 23,011 samples over 63 years. We chose nine meteorological variables from the dataset (Figure 1, Table I), which are collected at different pressure surfaces and typically used by meteorologists for making forecasts, as meteorological predictor variables. To enhance efficacy, we only chose the samples collected during the rainy season (April to October) every year, which might have correlation with extreme precipitation events. The samples in (1948-2000) are used as training set to learn the prediction model, and the remaining 10 years data are used as test set to evaluate the prediction model.

TABLE I
METEOROLOGICAL VARIABLES.

PW	Precipitable Water
T850	850hPa Temperature
U300	300hPa Zonal Wind
U850	850hPa Zonal Wind
V300	300hPa Meridional Wind
V850	850hPa Meridional Wind
Z300	300hPa Geopotential Height
Z850	850hPa Geopotential Height
Z1000	1000hPa Geopotential Height

B. Spatio-temporal Feature Space Construction

In order to build a feature space with the spatial and temporal information of the meteorological variables, we create the raw spatio-temporal feature input space for ST-CNN as following steps:

Step 1. Choose 1,024 locations, which are uniformly distributed wrapped around the State Iowa (32 latitudes and 32 longitudes).

Step 2. Sample 9 meteorological variables from 1,024 locations in the same day as the feature cuboid (32 rows and 32 columns) of one day.

Step 3. Repeat Step 2 until the feature cuboids of 10 continuous days are accumulated, then stack these cuboids as a tensor (32 x 32 x 90) which is the experimental sample of the last day in the 10 continuous days.

Step 4. Repeat Step 1-3 until all experimental samples are created .

C. Class Label Creation:

Here, we use historical spatial average precipitation data (the mean of daily precipitation totals from 22 observation stations divided by the standard deviation) of the State Iowa from the same time period to create the class labels. We define any 14 days periods as extreme precipitation clusters and label them as a positive sample if the total amount of precipitations of the 14 days reaches a historical high level (i.e., above the 95% percentile of the historical records). Otherwise, we label it as a negative sample.

V. PRELIMINARY RESULTS AND CONCLUSION

Here we compare our model with the streaming feature selection method OSFS [5], and use Accuracy and F-measure for evaluation. Particular, Accuracy ($\frac{TP+FP}{TP+TN+FP+FN}$, where TP is true positive, TN is true negative, FP is false positive, FN is false negative) refers to the closeness of a predicted class label to a known class label. And F-measure ($\frac{2*TP}{2*TP+FP+FN}$) conveys the balance between the exactness and the completeness.

TABLE II
EXPERIMENTS RESULTS.

Event	Metrics	OSFS	CNN
Extreme precipitation	Accuracy	0.712	0.708
Extreme precipitation	F-measure	0.748	0.743

Table II summarizes the performance of our CNN model predicting extreme precipitation clusters. Four learnable layers (two convolutional layers and two pooling layers) were able to produce comparable results. Our next research work will focus on improving the architecture of our model in convolutional and pooling layers and understand how to interpret the features learned by the ST-CNN with physically meaningful patterns among these meteorological variables.

ACKNOWLEDGMENTS

We thank the team of Dr. Shafiqul Islam and Dr. David Small from Tufts University for their help on historical meteorological data collection.

REFERENCES

[1] H. Jr and et al., "US geological survey natural hazards science strategy: Promoting the safety, security, and economic well-being of the nation," *US Geological Survey*, 2013.
[2] Johnson and et al., "Development of a $European$ flood forecasting system," *International Journal of River Basin Management*, pp. 49–59, 2003.
[3] H. Kaixun and et al., "Long-lead term precipitation forecasting by hierarchical clustering-based bayesian structural vector autoregression.," *IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC)*, 2016.
[4] Reyniers and et al., "Quantitative precipitation forecasts based on radar observations: Principles, algorithms and operational systems," *Institut Royal Mtorologique de Belgique*, 2008.
[5] X. Wu and et al., "Online feature selection with streaming features," *IEEE transactions on pattern analysis and machine intelligence*, pp. 1178–1192., 2013.
[6] Y. Di and et al., "Developing machine learning tools for long-lead heavy precipitation prediction with multi-sensor data," *Networking, Sensing and Control*, pp. 63–68, 2015.
[7] Z. Yong and et al., "An evaluation of big data analytics in feature selection for long-lead extreme floods forecasting.," *IEEE 13th International Conference on Networking, Sensing, and Control (ICNSC)*, 2016.
[8] D. Wang and et al., "Towards long-lead forecasting of extreme flood events: a data mining framework for precipitation cluster precursors identification," *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 1285–1293, 2013.
[9] K. Yu and et al., "Classification with streaming features: An emerging-pattern mining approach," *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 2015.
[10] W. Tong and et al., "Text simplification using neural machine translation," *Thirtieth AAAI Conference on Artificial Intelligence*, 2016.
[11] Anctil and et al., "Impact of the length of observed records on the performance of ANN and of conceptual parsimonious rainfall-runoff forecasting models," *Environmental Modelling Software*, pp. 357–368, 2004.
[12] Xingjian and et al., "Convolutional $LSTM$ network: A machine learning approach for precipitation nowcasting," *Advances in Neural Information Processing Systems*, pp. 802–810, 2015.
[13] Kalnay and et al., "The ncep/ncar 40-year reanalysis project," *Bulletin of the American meteorological Society*, pp. 437–471, 1996.