Climate-Net: Can AI Help Fight Climate Change?

Extreme precipitation forecasting using deep learning networks

Weichen Huang

St. Andrew's College

Contact Information:
St Andrew's College
8Q4X+55 Blackrock, County Dublin

Phone: +353 (085) 1191 391
Email: `w.huang@students.st-andrews.ie`

Abstract

This work presents an in-depth investigation of the use of deep neural networks for the prediction of short-term severe precipitations. We focus on attention-based network topologies for forecasting, and the scaling techniques for weather features. We have discovered that feature scaling is critical in the context of severe precipitation prediction. Attention-models with robust scalers outperform previous approaches and the uncertainty associated with future forecasts is reduced.

Introduction

Climate change is one of the most crucial challenges influencing humanity's future and consequently, severe weather events will rise in frequency and severity. Recent approaches to numerical weather prediction (NWP) are becoming increasingly inaccurate and unstable due to the unpredictability of weather extremes as a direct result of climate change. A widely applied neural network architecture for NWP is the Convolutional Neural Network (CNN). CNNs are not spatially invariant to the input data and do not encode the position of each value in the input matrix, which pose many challenges to detection.

This work addresses the space invariance issue by combining the convolutional layer with a multi-head self-attention module. Moreover, by applying Robust Scalers to a number of precipitation features, we successfully normalized temporal measurements and achieved significantly better detection rates of extreme events. Compared to previous methods, the proposed model achieves better accuracy in detecting extreme events and with much lower costs.

Main Objectives

1. We evaluated current methods for extreme precipitation prediction using a number of different datasets.
2. We have created a novel neural network: Self-Attention Convolutional Neural Networks.
3. We applied Robust Scalers to a number of precipitation features.
4. We repeated experiments and made evaluations with a multitude of model variatons on real-world datasets.

Methods

We build a self-attention augmented convolutional neural network to predict severe precipitation days using daily SLP and 500-hPa GPH anomalies. Each day's input data is processed via various layers to produce an output categorization of extreme precipitation or non extreme precipitation. Our model receives a three-dimensional matrix with dimensions of $15 \cdot 35 \cdot 2$ for each day (i.e., latitude · longitude · 2 input variables). Precipitation data is used to produce ground-truth labels for model training.

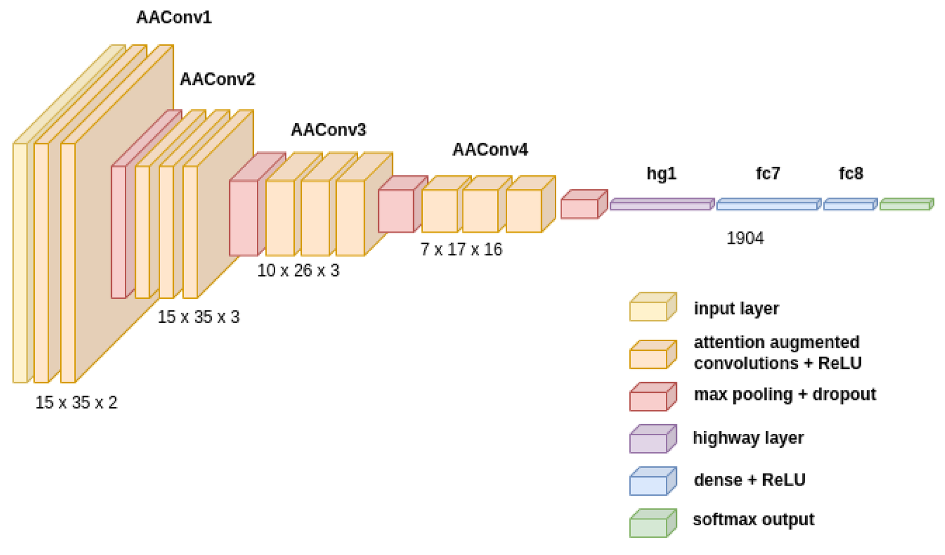

Figure 1: Model Architecture

Attention-Augmented Convolutions

We propose concatenating convolutional feature maps with a set of feature maps obtained by self-attention to enhance convolutional operators with this self-attention mechanism.

Given an input tensor of shape (H, W, D), which represents the height, width, and the feature depth of the anomalies respectively, we flatten it to a matrix $A(1, H \cdot W \cdot D)$ and perform multihead attention. The output of the self-attention mechanism for a single head h can be formulated as below:

$$Attn(X) = softmax(\frac{QK^T}{\sqrt{d_k}})V \quad (1)$$

where: $Q = (X * W_q)$, $K = (X * W_k)$, $V = (X * W_v)$, W_q, W_k, W_v are the learned linear transformations that map the input X to queries (Q), keys (K), and values (V) respectively.

A single attention score calculation uses $O(A^2 \cdot h)$ space, with the multi-head attention calculation using $O(A^2 \cdot h^2)$ space. This is much more efficient than many other previous attention-based augmentations to the convolutional layer due to omitting relative position embeddings. The multi-head attention calculation is performed on a single head at a time, and the output of the multi-head attention is concatenated to the output of the convolutional layer.

$$MHA(X) = Concat[Attn_1(X), Attn_2(X), ..., Attn_N h(X)] * W_m h \quad (2)$$

where:

$W_m h$ is the learned linear transformation that maps the concatenated attention scores to the output of the multi-head attention.

Finally, the multi-head attention scores are concatenated to the output of the convolutional layer:

$$AAConv(X) = Concat[Conv(X), MHA(X)] \quad (3)$$

where:

$Conv(X)$ is the output of the convolutional layer, and $MHA(X)$ is the output of the multi-head attention.

Results

This section shows the impact of feature scaling mechnisms on model performance. Due to length limitations, we only showcase the performance of the proposed self-attention augmented convolution model, with Standard Scaler and Robust Scaler. The case of other models is very similar with the proposed model. And these two scalers significantly outperform other scalers.

Scalers	Accuracy	Recall	Precision
Raw Scaler	0.9499	0.028	0.5000
Standard Scaler	0.9052	0.9369	0.3384
Robust Scaler	0.8930	0.9265	0.3119

Table 1: Feature Scaling on Model Performance (SAConvNet)

As is shown in the table, without feature scaling, the model predictions generate predominantly arbitrary results. This is caused by the seasonal variability of the weather data, which hence must be removed in feature engineering.

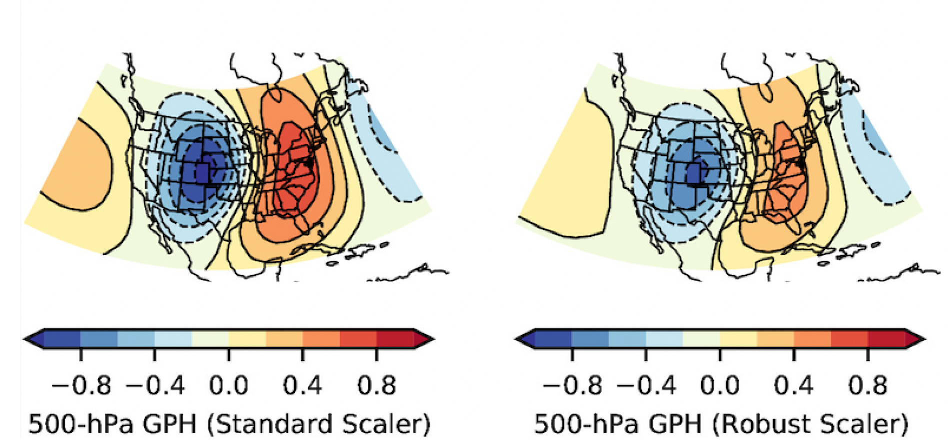

Figure 2: Feature Scaling Visualization (GPH)

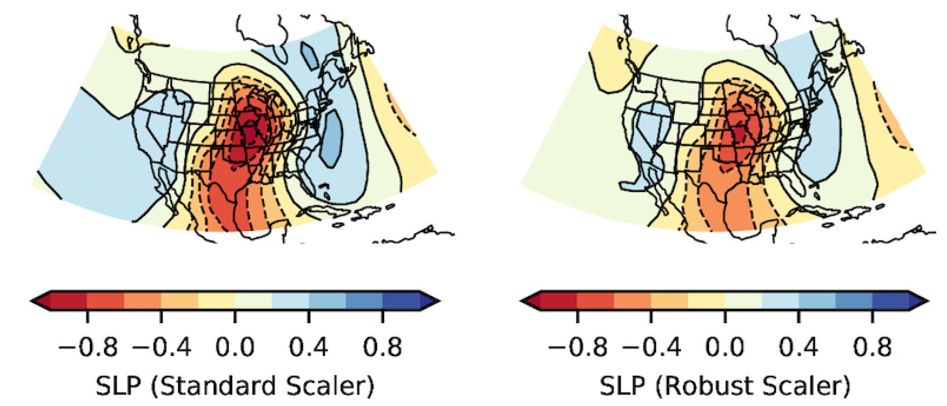

Figure 3: Feature Scaling Visualization (SLP)

Models	Accuracy	Recall	Precision
CNN	0.8671	0.9060	0.2613
SAConvNet	0.9052	0.9369	0.3119
SAConvNet + Highway	0.9616	0.8598	0.5983

Table 2: Model Performance (on Standard Scaler)

As is shown in the table above, the proposed SAConvnet model + highway networks achieves the best overall accuracy of 97%, with a 12% improvement as compared to the CNN model result. It is capable of accurately identifying more than 88% of severe precipitation days as extreme precipitation days (EPD), similar to the classical CNN model. Extreme precipitation occurred on 64% of days classified as EPD patterns, 1.5x better than the CNN model. Only fewer than 12% of days, classified as non-EPD patterns, resulted in severe precipitation.

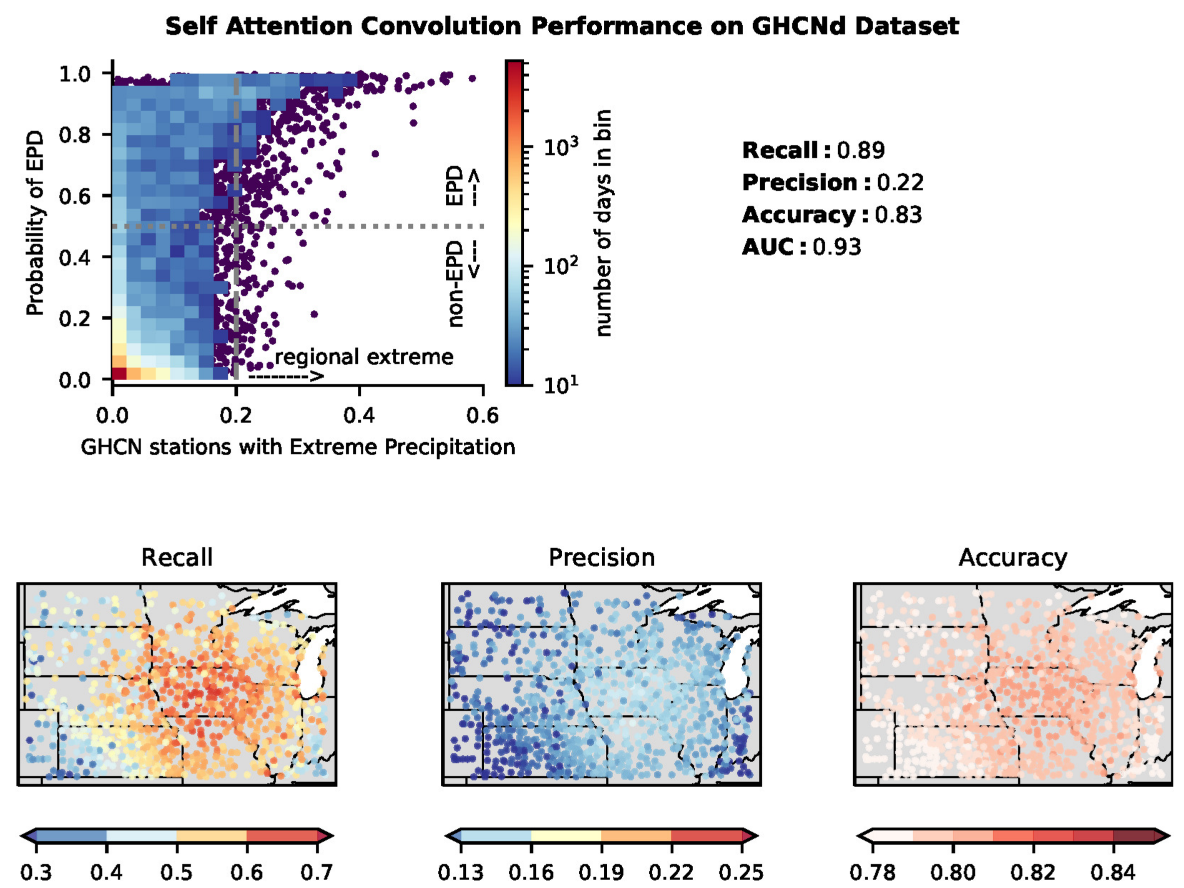

Figure 4: Model Performance Visualization

Conclusions

This work successfully creates feature scaling techniques and network architectural designs that improve the accuracy and applicability of extreme precipitation forecasting. However, some extremes might still be missed by the algorithms due to extreme precipitation often being controlled by localized processes not reflected in the regional-mean daily precipitation. Additionally, by incorporating geographical and climate data, the models could provide insight into different causes of extreme precipitation and extreme weather events. Because the models are trained on variables that are available globally, transfer learning can be used to analyse extreme precipitation and additional types of extreme weather in other regions. The results demonstrate that deep learning can provide critical insight into the physical processes underlying changes in climate extremes.

Acknowledgements

I wish to extend my thanks and appreciation to:

- My parents for helping me with the technical aspects of my project,
- Ms Buckley and St Andrew's College for the chance to partake in the event,
- BT Young Scientist for hosting the event.