ENSO Forecasting Over Multiple Time Horizons Using Deep Convolutional LSTM Network and Rolling Mechanism

Bin Mu

*School of Software Engineering*

*Tongji University*

Shanghai, China

binmu@tongji.edu.cn

Cheng Peng

*School of Software Engineering*

*Tongji University*

Shanghai, China

pengcheng@tongji.edu.cn

Shijin Yuan\*

*School of Software Engineering*

*Tongji University*

Shanghai, China

yuanshijin2003@163.com

Lei Chen

*Shanghai Central Meteorological*

*Observatory*

Shanghai, China

qqydss@163.com

*Abstract*—TBD

Keywords—Convolutional LSTM, Rolling Mechanism, Spatiotemporal Sequence Forecasting, ENSO

# Introduction

"fully coupled models are computationally expensive and only just starting to become publicly available. From a practical point of view, it remains challenging for the typhoon community to replace existing operational forecast systems with coupled models at this stage."

Approximately every 4 years, the sea surface temperature (SST) is higher than average in the eastern equatorial Pacific. This phenomenon is called El Niño-Southern Oscillation (ENSO) and is considered as the dominant mode of interannual climate variability observed globally (Wunsch 1990). ENSO is associated with many climate changes (Fraedrich 1994; Wilkinson et al.1999), affecting climate of much of the tropics and subtropics, then cause enormous damage worldwide, so a skillful forecasting of ENSO is strongly needed.

1. What is ENSO and its extremely impacts (why a skillful forecasting is very important);

2. Methods to study ENSO currently (Climate method and it should be improved with longer prediction ahead) predict result not well enough + computationally expensive --> remain room for further study of this problem.

3. ENSO is a spatiotemporal sequence forecasting problem and DL (machine learning) methods have great potential to handle this problem. exist work have been attempt to this problem. However, little work have been done with this problem for exploring the spatial and temporal information of SST pattern.

2 complex factors:

\* Spatial dependencies

\* Temporal dependencies

Why Rolling? Direct is not good and rolling is a better choice for multi-step forecasting.

We formulize the grid SST pattern problem and use the Convolutional LSTM model to capture the spatial and temporal information of ENSO development simultaneously,

In summary, the contributions of our work are 3-fold:

1. We formulize the ENSO SST pattern forecasting problem, a I \* J grid map based on longitude and latitude where a grid donates a region of NINO3.4, which can be converted as a multi-channels physical parameters setting spatiotemporal sequence forecasting problem;

2. We apply Convolutional LSTM network, which can capture the spatial and temporal information of SST data effectively, to predict ENSO with -6, -9, -12 monthly ahead respectively, the result show that our model outperform other neural network models and conventional statistical model.

3. Visualization of ENSO Pattern

The remainder of the paper is structured as follow:

1.

# Related Work

## Machine Learning for ENSO Forecasting

TBD

## Spatiotemporal Sequence Forecasting with Deep Learning

TBD

# Methodology

## Formulation of ENSO Forecasting Problem

TBD

## Convolutional LSTM and Rolling Mechanism

TBD

# Experiments

## Experiment Settings

We conduct experiments on the real-world monthly (1850.01~2015.12) SST grid dataset, which covers the Niño 3.4 (5N-5S, 170W-120W) region with 1° latitude multiplied by 1° (<https://www.esrl.noaa.gov/psd/data/gridded/data.cobe2.html>). To evaluate the prediction performance over different time horizons, we apply different sliding windows with 6, 9 and 12 month ahead to construct the input sequence respectively. 80% of data is used for training, 10% are used for testing while the remaining 10% for validation. To evaluate the performance of our model, we adopt 3 commonly used metrics: Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE), all are the lower the better (formulations as follow).

In the following experiments, we first compare different forecasting methods over different prediction horizons (6- month, 9-month and 12-month) to validate the effect of our model, then discuss the influence of different hyperparameters settings to the final result. Finally, we take the ENSO during 2015/16 as the case, investigate the model interpretation with the generated SST patterns. All neural network based approaches are implemented using Keras, and all the codes are available on GitHub at <https://github.com/KrisCheng/Deep4Cli>. We run all the experiments on a computer with two NVIDIA 1080Ti GPUs.

1. Performance comparison of different models

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| T | Metric | HA | ARIMA | SVR | FNN | CNN | FC-LSTM | ConvLSTM | ConvLSTM-RM |
|  | RMSE | 1.555 | 1.300 | 2.056 | 1.261 | 1.048 | 1.341 | 0.947 | **0.729** |
| 6-month | MAE | 1.271 | 1.053 | 1.767 | 0.860 | 0.777 | 1.004 | 0.749 | **0.555** |
|  | MAPE | 4.78% | 3.95% | 6.44% | 3.15% | 2.96% | 3.85% | 2.72% | **1.45%** |
|  | RMSE | 1.506 | 1.314 | 2.056 | 1.248 | 1.147 | 1.313 | 0.976 | **0.807** |
| 9-month | MAE | 1.224 | 1.051 | 1.791 | 0.997 | 0.920 | 0.981 | 0.769 | **0.605** |
|  | MAPE | 4.59% | 3.95% | 6.54% | 3.73% | 3.38% | 3.75% | 2.86% | **2.27%** |
|  | RMSE | 1.251 | 1.158 | 2.119 | 1.295 | 1.039 | 1.079 | 1.033 | **0.789** |
| 12-month | MAE | 0.969 | 0.905 | 1.882 | 1.034 | 0.801 | 0.814 | 0.805 | **0.607** |
|  | MAPE | 3.64% | 3.39% | 6.86% | 3.86% | 3.00% | 3.06% | 2.98% | **2.25%** |

## Effect of Spatiotemporal Modeling

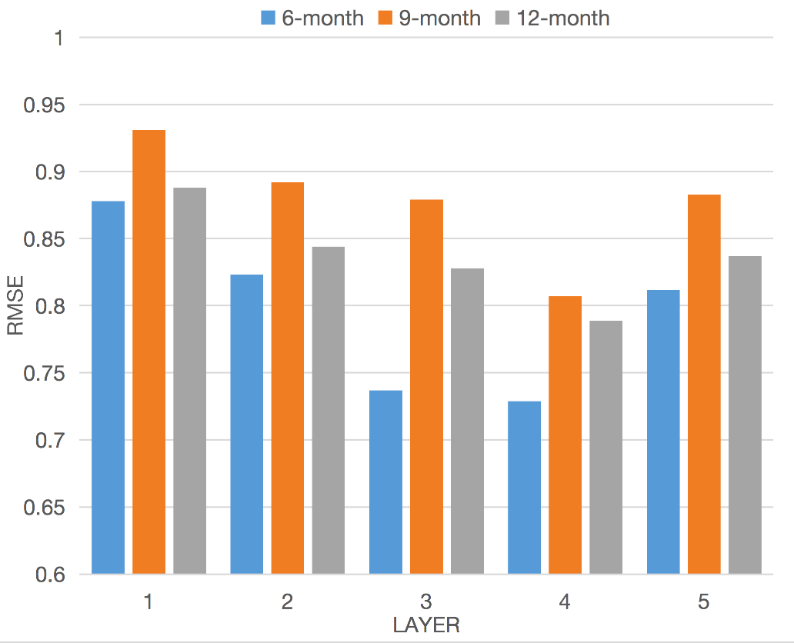
We compare ConvLSTM-RM network with other widely used time series regression models, including (1) HA: Historical Average, which models the development of SST as a periodic variation, and uses the historical observation as the prediction; (2) ARIMA: Auto-Regression Integrated Moving Average, which is a well-known model for understanding and forecasting future; (3) SVR: Support Vector Regression, which applies linear support vector machine for regression task. The following deep neural network based models are also included: (4) Feed forward Neural Network (FNN): Feed forward neural network; (5) Convolutional Neural Network (CNN); (6) Recurrent Neural Network with fully connected LSTM hidden units (FC-LSTM); (7) Convolutional LSTM Network without Rolling Mechanism (ConvLSTM). Those deep neural networks all with 3 hidden layers and roughly same amount of parameters, and they are fully trained with a fixed number of epochs (e.g., 10000 epochs) for fair comparison. Adam optimizer is applied during training process to accelerate learning process, and batch normalization layer is also applied to reduce over-fitting.

Table 1 shows the result of different models for 6-, 9- and 12-month lead time forecasting. We observe the following phenomena on this process: (1) ConvLSTM-RM outperforms all other baselines regarding all the metrics for all forecasting horizon, which suggests the effectiveness of handling spatiotemporal dependencies. (2) Deep neural network based methods, especially CNN, ConvLSTM and ConvLSTM-RM, tend to have a better performance than other baselines. One intuitive reason is that the development of SST is irregular and highly spatial-correlated, so it is hard for a model to give accurate predictions on test set without learning the inner dynamics development of the climate system. (3) The performance of different models does not show consistently tendency with the growth of forecasting horizon, and the performance of CNN is better than FC-LSTM. The intuition is that the temporal dependencies of ENSO is hard to capture than spatial dependencies in this experiment.

## Comparison of Different ConvLSTM-RM Structures

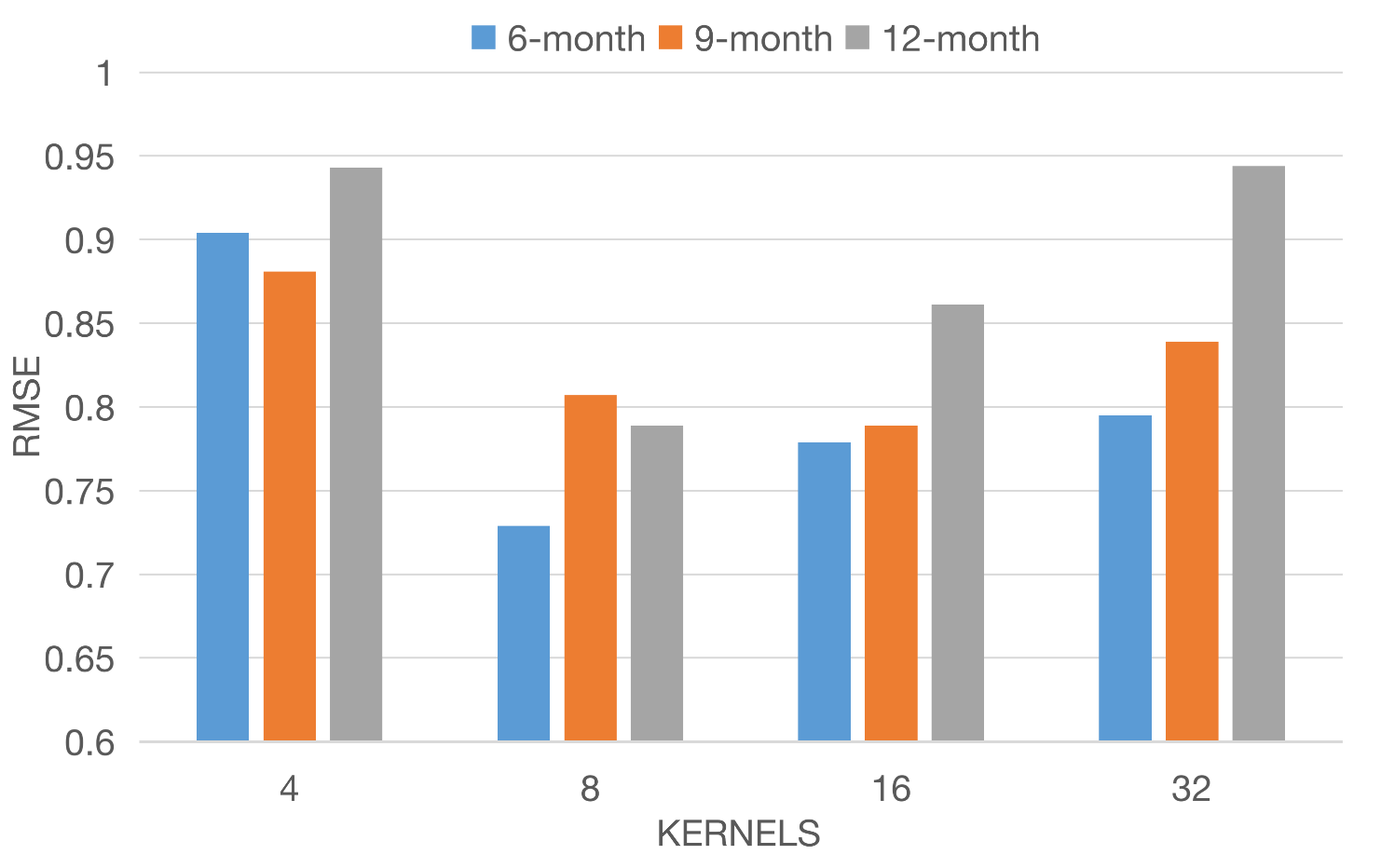
To further investigate the impact of different network structures to the final result and figure out the best network structure, we compare ConvLSTM-RM with the following two aspects: (1) Number of Layers, which is a fundamental setting for deep neural network structure. (2) Number of kernels and kernel size, which can stand for the collectors of spatial correlation between SST grid data. We apply grid search strategy in this process.

Fig. 1 shows the comparison between different ConvLSTM layers. We find that: (1) deeper models can produce better results with fewer parameters; (2) More layer does not always reach better performance, the RMSE of result is first decrease gradually and then increase over all different time horizons. The metric of MAE and RMSE also show the same tendency. Actually, experiments in previous studies show that the layer is not the more the better, and the training data are not sufficient enough to learn so many parameters. 4 layer can be considered an optimal choice in this case.



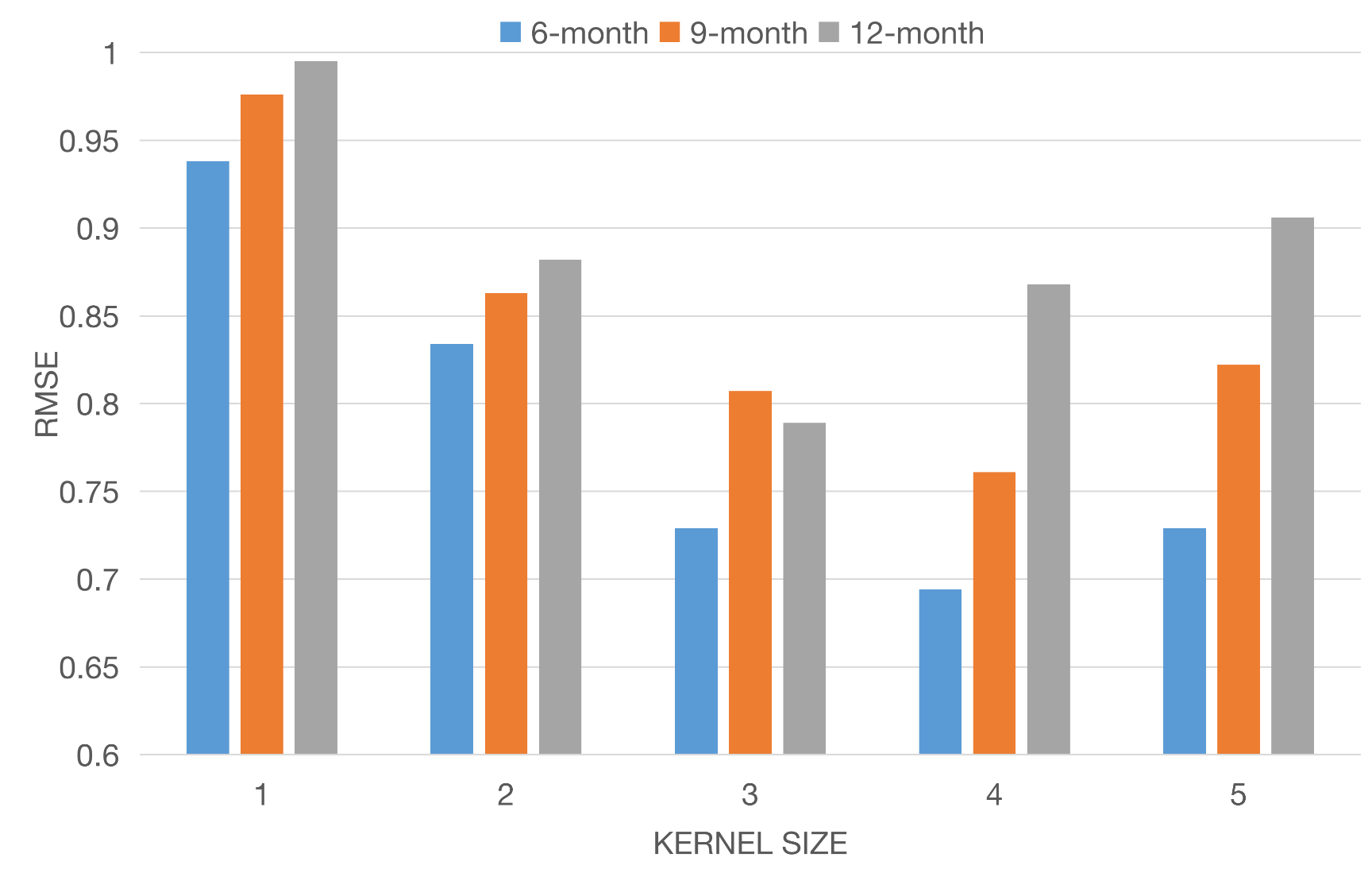
1. RMSE of different layers with 6-, 9- and 12-month ahead.

Next, we explore the choice of different number of kernel and kernel size. Fig. 2 shows the performance of different kernel number with a four layer network. Larger kernel number enables the model to capture different representation of spatial dependencies at the cost of increasing learning complexity, which can be considered that different abstract of spatial information are captured by different kernels. We observe that with the increase of kernel number, the error on the testing dataset first quickly decrease, and then slightly increase. 8 kernels of each layer works best on 6- and 12-month while 16 kernels of each layer works best on 9-month.



1. RMSE of different kernels with 6-, 9- and 12-month ahead.

With the fixed layer and kernel number, we go deep to explore the influence of different kernel size, which can be view as the volume of collector. Larger kernel size enables the model to capture broader spatial dependency over the Niño region. However, the experiment result shows that not the larger the better. The kernel size of 4 performs best with 12-month lead, and the kernel size of 4 works best with 6- and 9-month lead.



1. RMSE of different kernel size with 6-, 9- and 12-month ahead.

## Model Interpretation with Generated SST Patterns

To better understand the behavior of our model, In this part, we take the ENSO occured during 2015/2016 as the case, which is considered as the most extreme ENSO since records began, and the fluctuation of SST is a good candidate for study. We visualize the generated SST patterns from 2015.01 to 2016.01 with our model. Fig 4 shows the visualization of 12-month lead forecasting. We have the following observations: (1) ... (2) we calculate the NINO3.4 index during this period, and compared it with Climate Model Prediction From IRI CPC. (3) .... which can be considered as a successful prediction of ENSO.

Outline:

Pattern of ENSO during 2015/2016 --> our model capture the development of SST pattern --> regression result, predict the peak value and can be considered as a successful prediction of ENSO 15/16.

(https://iri.columbia.edu/our-expertise/climate/forecasts/enso/2015-January-quick-look/)

(https://iri.columbia.edu/our-expertise/climate/forecasts/enso/2015-January-quick-look/?enso\_tab=enso-sst\_table)

# Conclusion and Future work

In this paper, we

##### References

1. TBD