Modeling Short-term Wildfire Spread with Convolutional LSTM: Challenges and Opportunities

Jianing Fang, Caroline Juang

1 Introduction

Climate change is expected to alter the global wildfire distributions by changing both fuel availabilities and the conditions conducive to burning (Krawchuk et al., 2009). Current generation climate models produce divergent predictions of future wildfire frequency and severity in different climate regimes, with fire probabilities predicted to increase in some regions but to decrease in others (Parks et al., 2016; Westerling & Bryant, 2007). The changing pyrogeography will likely pose new challenges for coupled natural-human communities, urging the development for more coordinated efforts to manage fire risks and land use patterns (Moritz et al., 2014). While knowledge is needed to understand both changes in long-term fire probabilities and short-term fire dynamics, at the forefront of wildfire management, firefighters and wildfire management agencies need timely predictions of wildfire behaviors to plan for fire control and evacuations.

Advancements in computational power and data availability have vastly expanded our capability to simulate fire behaviors using computer models that can be categories into physical, empirical, and mathematical analog models (Sullivan, 2009a). Physical fire models faithfully represent fire physics and chemistry of fire by solving conservation laws, yet these models are often prohibitively expensive to be solved for real-time fire predictions (Sullivan, 2009a). Empirical models, on the other hand, are not theory-based but are instead based on the statistical relations between wildfire and various environmental variables(Sullivan, 2009b). Simulation and mathematical analog models attempt to simulate fire spread over 2D lattice by propagating the fire front across the landscape (Sullivan, 2009c).

More recently, researchers have been using machine learning (ML) models as an alternative to the prescribed models to model and forecast fire behaviors. Although training models require a comparatively large amount of historical data, once trained ML models can make time-critical fire forecasts in a very efficient manner that is suitable for the task of predicting short-time fire dynamics.

In this study we formulate fire spread prediction as a spatiotemporal image forecasting task. Sequence-to-sequence learning algorithms such as recurrent neural network (RNN) or long-short-term memory (LSTM)(Hochreiter & Urgen Schmidhuber, 1997) have been proved successful in extracting the temporal dynamic properties and memory effects in earth system studies (Reichstein et al., 2019). Convolutional neural networks (CNN), on the other hand, have the property of translation invariance that can be leveraged to learn the proximate spatial context in the dataset (Reichstein et al., 2019). Recent efforts to blend the convolutional and recurrent approaches have led to the development of convolutional LSTM (convLSTM), a model first developed for precipitation nowcasting with radar imaging (Shi et al., 2015). The original LSTM framework was designed to tackle the inefficient recurrent backpropagation by using memory cells and gate units to control whether gradient information will flow to the next timestep, thereby learning to truncate the temporal dependency and alleviate the vanishing gradient problem. Briefly, an LSTM cell is formulated as (Hochreiter & Urgen Schmidhuber, 1997)

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci} \circ c_{t-1} + b_i)$$
$$f_t = \sigma\left(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf} \circ c_{t-1} + b_f\right)$$
$$c_t = f_t \circ c_{t-1} + i_t \circ tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c)$$
$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co} \circ c_t + b_o)$$
$$h_t = o_t \circ tanh(c_t)$$

In which i_t is the input gate controlling how much new input information will be accumulated to the cell, f_t is the forget gate that allows information from the past cell c_{t-1} to be forgotten when activated, and the output gate o_t controls how much cell information will be propagated to the will the hidden state output h_t. The ∘ notation in the equations denote Hadamard product. The novelty of the convolutional LSTM is that it replaces both input-to-state and state-to-state transitions by a convolutional operator (replacing matrix multiplication with the weight tensor with convolutional operator in the LSTM formulations equations) (Shi et al., 2015).

$$i_t = \sigma(W_{xi} * x_t + W_{hi} * h_{t-1} + W_{ci} \circ c_{t-1} + b_i)$$
$$f_t = \sigma\left(W_{xf} * x_t + W_{hf} * h_{t-1} + W_{cf} \circ c_{t-1} + b_f\right)$$
$$c_t = f_t \circ c_{t-1} + i_t \circ tanh(W_{xc} * x_t + W_{hc} * h_{t-1} + b_c)$$

$$o_t = \sigma(W_{xo} * x_t + W_{ho} * h_{t-1} + W_{co} \circ c_t + b_o)$$
$$h_t = o_t \circ tanh(c_t)$$

In a 2D grid, the current state of each grid cell is computed as a function of the input state and the previous states in its neighboring cells. This property makes the convLSTM amenable to the task of fire spread prediction, as the probability of whether a grid cell will be burned on current ay t is influenced (1) whether its neighboring cells were burned in recent days (t-1, t-2…); (2) properties within the grid cell such as land cover, vegetation cover, the amount of available fuel; (3) near term meteorological variables in the region such as wind velocity and surface temperature; (4) topographic context within the larger landscape. The convLSTM allows us to effectively integrate these four types of information to make fire forecasts.

2 Related Works

Previous efforts to combine convolutional and recurrent neural networks for environmental modeling found the hybrid convolutional-recurrent models often outperform standalone neural networks (Barzegar et al., 2020; Burge et al., 2020). One approach to combine these two deep learning models is to first pass the input layers into a CNN, using the outputs of CNN as the input to LSTM cells, and finally passing the LSTM outputs to fully connected layers. For example, Barzeger et al. used a hybrid CNN-LSTM model to predict short-term water quality variable (Barzegar et al., 2020), Yusuft et al. adopted this strategy for predicting extreme temperature changes in enclosed compartments(Yusuf et al., 2021), and Kim & Cho used this method for housing energy consumption (Kim & Cho, 2019). A second strategy is the convLSTM method that embeds convolutional operations into the input-to-state and state-to-state transitions of LSTM cells as mentioned above. Burge et al. applied convLSTM to simulate the spread and dynamics of fires generated by a mathematical analog fire model (Burge et al., 2020). They tested ConvLSTMs on generated fire with varying degrees of complexity by iteratively adding confounding factors such as wind direction, terrain, and spatially varying moisture and vegetation distributions, and found ConvLSTMs outperform other models to reliably capture fire transmission pattern and the rate of fire spread (Burge et al., 2020). In this study, we adopt the second strategy to simulate actual fire recorded in MODIS burned area products in hope that the ConvLSTMs can reliably capture fire dynamics within spatially and temporally heterogeneous landscapes.

3 Method

3.1 Data Acquisition and Preprocessing

We used event and daily level polygons from the Fire event delineation product (FIRED) derived from MODIS MCD64A1 burned area product. The FIRED product uses an algorithm to identify individual fire events in a spatial-temporal data cube. The dataset contains 278568 fire events in the coterminous United States and Alaska from November 2001 and March 2021. However, the dataset is highly skewed towards small and short fires with 57.8% of the events lasting only 1 day and 61.6% of the events covering less or equal than 2 pixels (Figure 1). To construct a dataset with fire events with reasonable homogenous temporal and spatial dynamics, we filtered our dataset to use only fire events in the coterminous United States with a total burned area between 10-30 km^2, bounded by a 32X32 grid cells box in the MODIS sinusoidal grid, and with event duration greater than 3 days.

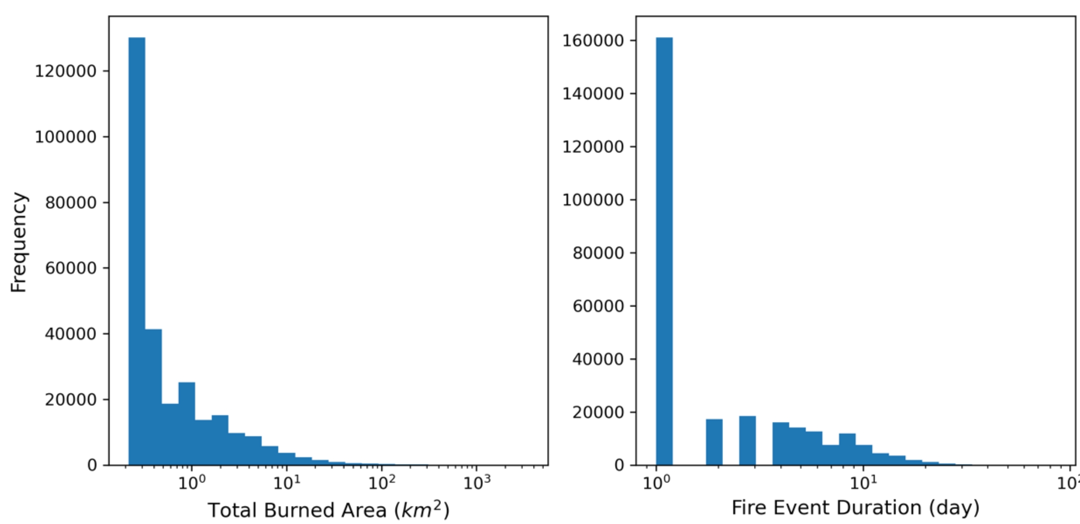


Figure 1. Distributions of fire event area and duration in the FIRED dataset

Next, we retrieved MODIS NDVI (MOD13Q1) product on the last observation date before the date of ignition and the MOD12Q1 land cover classification product as a proxy of vegetation cover and fuel availability. We used the 15-arc second Global Multi-resolution Terrain Elevation Data 2010 (GMTED2010) dataset as model input for elevation. Four meteorological variables, we sampled the 10m u-component of wind, 10-m v-component of wind, and 2m temperature fields at 2 am, 8 am, 2 pm, 10 pm local time for each date at the event location centroid, and evaluated daily-mean values from these four values. All input variables are rasterized into 32X32 matrices centered around the fire event centroid in 500m MODIS sinusoidal projection. We assume that the vegetation, land cover, and topography layers are only varying in the spatial dimension and that the meteorological variables are only varying in the temporal dimension for each fire event. The GMTED2010 dataset is locally normalized at each fire event site using a min-max scaler, whereas the meteorological variables are normalized based on the global min and max values of the entire dataset. The processed input dataset is of dimension N_cxTxDxD, where N_c (7) is the number of variables, T ($\geq$4) is the number of observation days in a fire event, and D (32) is the dimension of a single frame (32). The filtered dataset contains a total of 5074 fire events with an average number of 9 timesteps for each event. Despite gaps exist between adjacent timesteps due to missing data in the MOD64A1 dataset, we did not interpolate the gaps due to the highly stochastic nature of the fire spread pattern (Figure 3). The impact of missing data and potential methods to remediate their effects will be addressed in the later sections.

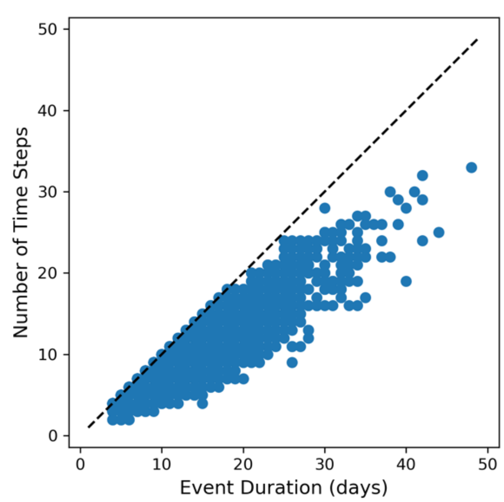


Figure 2. The number of timesteps in the processed sequence is less or equal than the event duration due to presence of missing data in the observation.

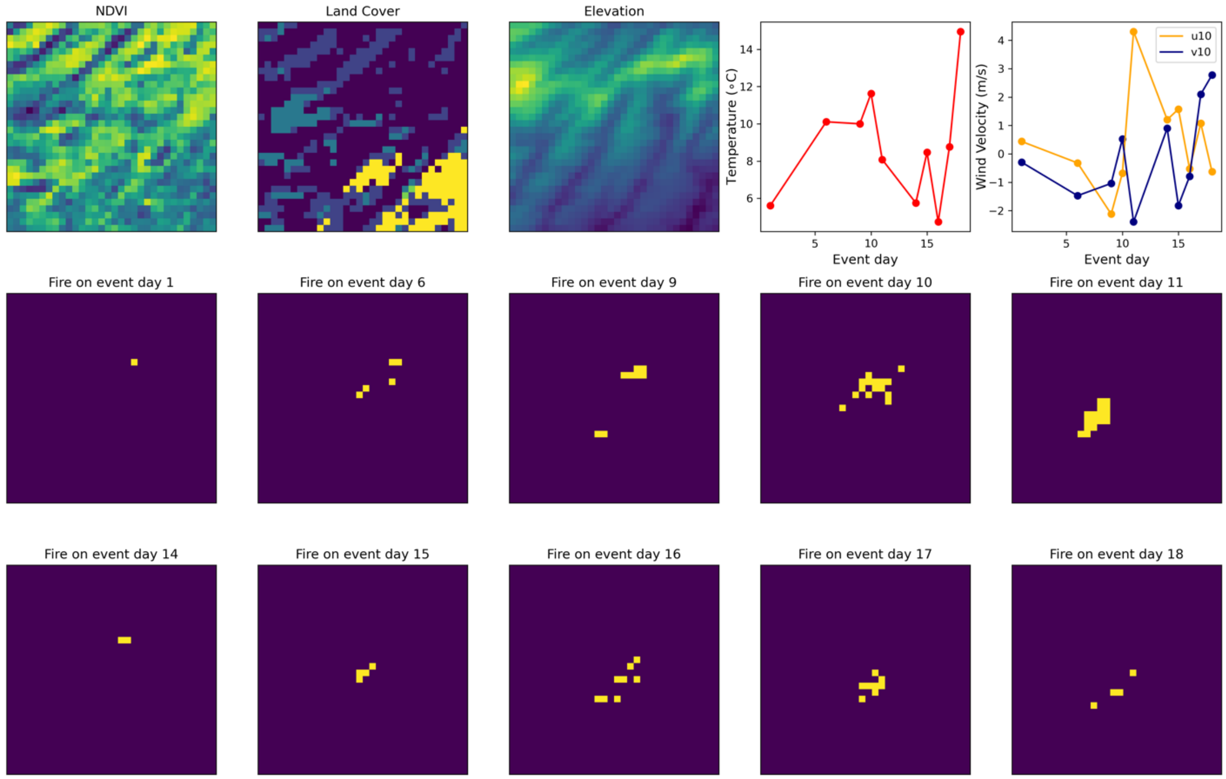


Figure 3. A fire event in the training sample. The top row shows the input variables of NDVI, Land Cover, Elevation, Temperature, and Wind Speed. The next rows show the fire dynamics from day 1 to day 18. Notice the gaps in the timeseries due to missing data in the MODIS burned area product.

3.2 Training Procedure

We trained a total of 54 ConvLSTM models using 80% of the dataset as the training dataset, 16% of the data as the validation dataset, and 4% of the data as test dataset. All variables from timestep 0 to timestep T were used as the input variable, and the target variable was the fire raster at timestep T+1 (Figure 4). Binary Cross Entropy (BCE) was used as the loss function between the predicted fire location and the observed fire location at T+1. To extract more complex non-linear dynamics between variables, we stacked multiple convLSTM on top of each other and apply batch normalization between the layers to build deeper networks. Given that the fire event samples within each minibatch may differ in length (T), we padded the samples so that all fire event samples have the same length as the longest event in the minibatch. This operation has no effect on the result but allows faster training with parallelized operations. We adjusted the kernel size between 3, 5, 7 to accommodate to fire with different spreading rates. We varied the number of kernels (N_h) at 32, 64, 96, the number of convLSTM layers at 1, 3, 5, and the learning rate at either 0.0001 or 0.00001. We adopted the ReLU activation function within the convLSTM, used the Adam optimizer for parameter update, and set the batch size at 20 fire event samples. All models were trained for a maximum of 500 epochs. To prevent overfitting, we applied early stopping by saving model parameters with the lowest validation loss and stopped training when the validation loss did not improve for 10 epochs. We also trained our model on the synthetic moving MINST dataset as a sanity check before training model over the actual fire dataset. While the original convLSTM implementation outputs an entire sequence shifted one timestep forward, our experiments on the moving MINST the loss function defined over the sequence required significantly more training epochs to converge on the moving MINST dataset (Shi et al., 2015). As we are only interested in the predicting the first timestep after the input sequence, we adapted the model structure so that the model only outputs the last timestep. The models were thoroughly tested on the MINST dataset to ensure reliable results can be obtained before we moved onto the actual fire dataset. The model code was adapted from a tutorial by Rohit Panda on GitHub[1] and implemented in the popular machine learning framework PyTorch.

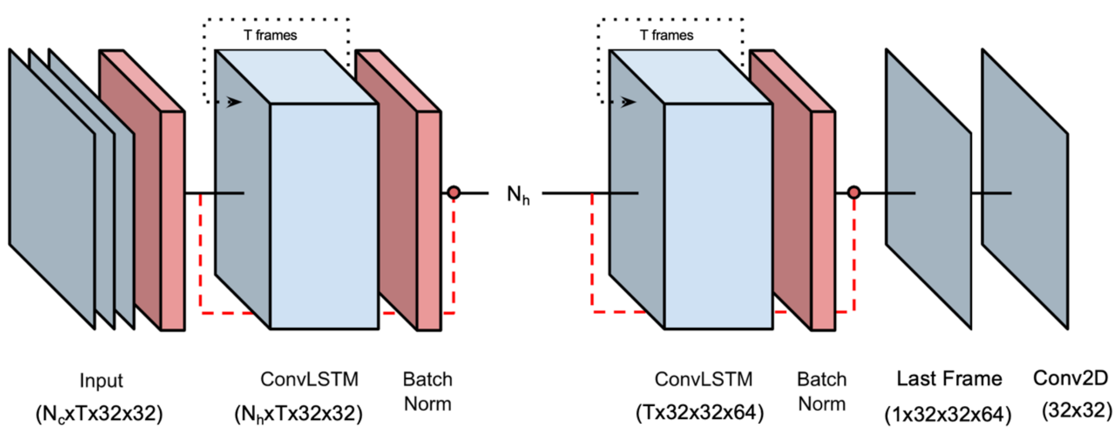


Figure 4. Model Architecture (illustration schematics adapted from (Burge et al., 2020))

[1] https://sladewinter.medium.com/video-frame-prediction-using-convlstm-network-in-pytorch-b5210a6ce582

4 Results

We completed the training 53 of 54 networks using the parameter combinations discussed in the previous section by the time of completing this paper. The model with the minimum validation loss was trained with 5 ConvLSTM layers, 32 hidden kernels of size 7 within each layer, and at a learning rate of 0.0001. In general, increasing the kernel size resulted in higher validation loss, while the increasing the number of ConvLSTM layers result in a decrease in the validation loss. The validation loss also appeared to be less with a smaller number of kernels and a slightly faster learning rate (Figure 5). The rise in validation loss along with the increase in kernel size may be related to the sporadic pattern of fire spread, as increasing the kernel size may result in higher variabilities in the fire pattern that makes training more intractable. However, using a deeper network with more ConvLSTM layers may allow us to extract more interactions between the fire dynamics and other input variables. The poorer performance on validation dataset at higher number of kernels may be related to the difficulty of tuning a large number of parameters with the limited amount of input data.

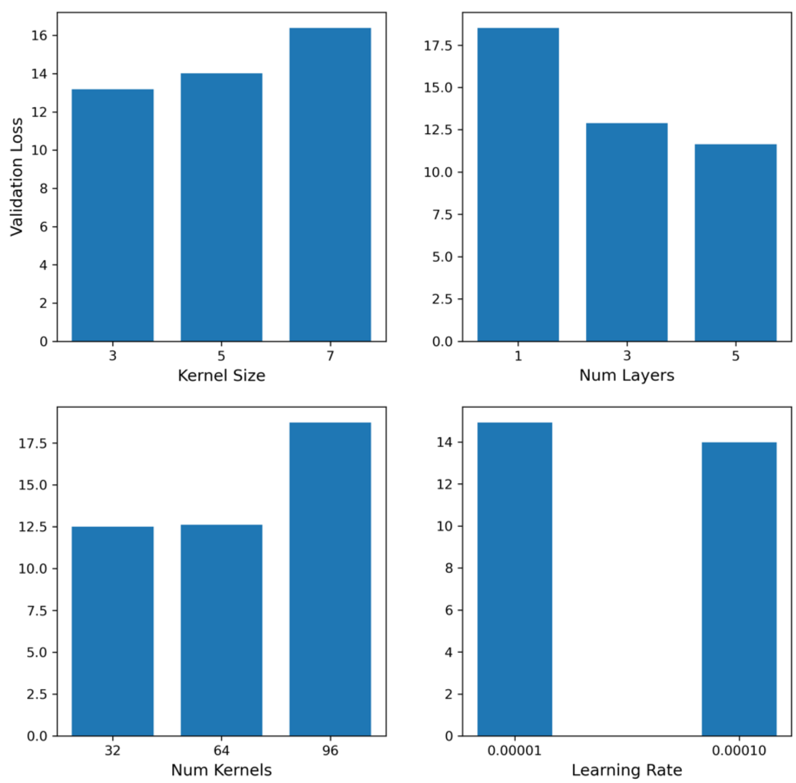


Figure 5. Effect of kernel size, number of layers, number of kernels, and learning rate on the validation loss of the model

However, when evaluated against the test dataset, we found significant deficiencies in the forecast outcomes (Figure 6). Even when evaluated against the model with the minimal validation loss, we found the model only learns the mean distribution of fire locations in the input dataset but does not predict the fire dynamics given the fire location. The best performing model only produces a near normal distribution with higher fire probability in the center of the frame. The elliptical shape of the high probability region is due to the sinusoidal projection of the fire dataset, and should appear near circular if projected onto a conformal coordinate system. Other models output similar fire patterns and none of the models were able to effectively capture the fire dynamics.

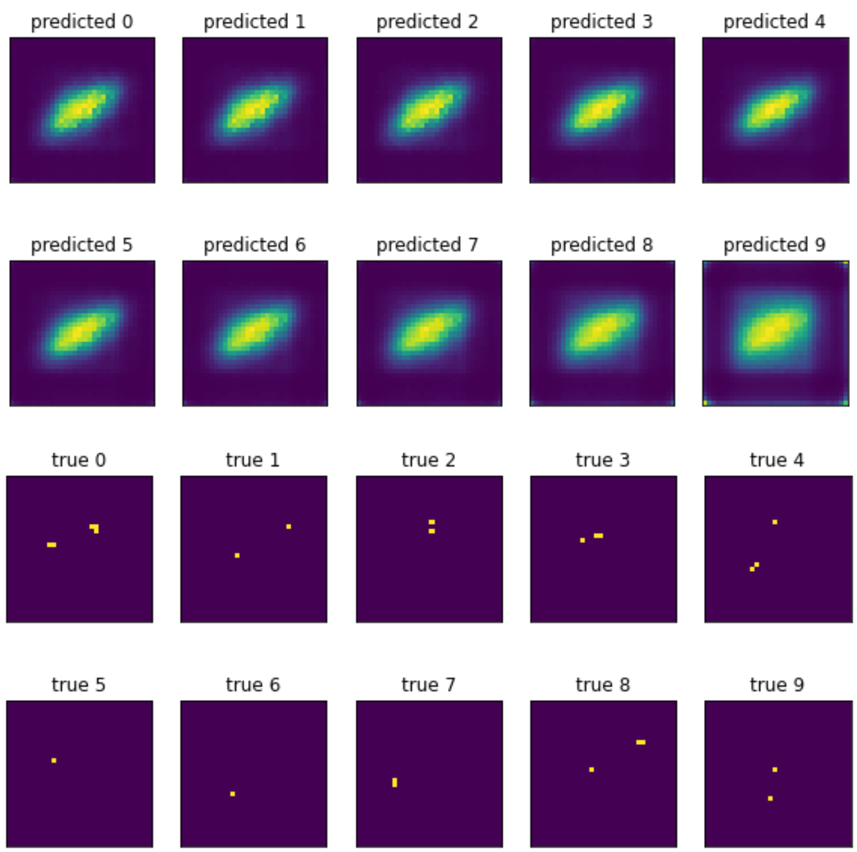
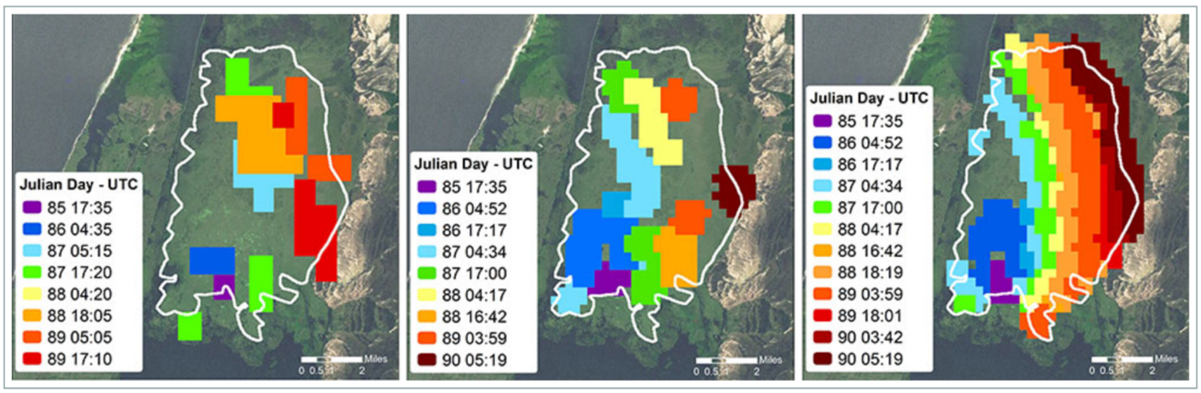


Figure 6. True and predicted next frame fire distribution for 10 fire events in the test dataset.

5 Discussion

To our knowledge, our experiment was the first study to apply convLSTM for fire spread prediction on actual fire spread datasets derived from satellite imagery, while previous studies were conducted on simulated datasets. However, the results suggested there were several challenges in reliably predicting fire using the MODIS dataset. First, the existence of observation gaps in the MODIS dataset complicated the prediction and introduced inconsistent intervals between adjacent timesteps. One potential way to mitigate issue is to replace the missing input timestep with the predicted timestep. This strategy has been adopted for the scheduled sampling of RNN for sequence prediction using a less guided scheme to mitigate the discrepancy between training with known samples and autoregressive prediction with unknown samples(Bengio et al., 2015).



Daily fire spread mapped by 1km Aqua/MODIS (left), 750 m VIIRS (center) and 375 m VIIRS (right) data at the Taim Ecological Reserve in southern Brazil. The data cover the period 26 -31 March 2013. The white outline shows represents the burned area mapped using 30m Landsat-7 on 31 March. This figure is reproduced here courtesy of Wilfrid Schroeder, University of Maryland.

Figure 7. Comparison of VIIRS active fire data with MODIS 1km Burn area product. Notice that while the mapped fire locations appear to jump from one day to the next in the MODIS dataset, the fire transitions appear to be more continuous in the 375m VIIRS dataset[2].

However, as Figure 2 shows, even in absence of observation gaps (e. g. event day 14-18) the movement of fire from one day to the next appear to be highly sporadic. Instead of a gradual translation movement from one region to the next, the fire appeared to jump from one location to the other location, whereas the area and shape of the fire perimeter showed highly

[2] Figure source: https://earthdata.nasa.gov/earth-observation-data/near-real-time/firms/viirs-i-band-active-fire-data

irregular changes in adjacent days. This “jumpy” behavior of the fire makes it difficult for the convolutional operator to capture useful features across the timeseries. Preliminary comparison with other sensors with higher spatial resolution seems to suggest the inconsistencies in the fire data are not reflecting the actual fire dynamics but are instead related to the sensor limitations. For example, the VIIRS Band-I 375m Active Fire Product provides a more accurate mapping of smaller fire perimeters due to higher spatial resolution and may provide a more dataset more suitable for training with convLSTMs (Figure 7). Using input datasets with higher spatial and temporal resolution will also allow us to include fire with shorter durations and smaller burned areas, which are currently excluded from the dataset.

Given the limited amount of available data, we used the entire sequence of each fire event for training despite the different durations for different fire events. Therefore, a simplifying assumption is that the fires included in the training dataset are governed by the same dynamics regardless of the duration, and the same transition dynamics will apply from the beginning to the end of the fire event. However, such assumptions may break down due to the large heterogeneities in the spatial and temporal variations of different fire events. Future efforts to improve the training result should focus on either first segmenting the dataset and training the model on more homogenous fire types or developing model structures that can more effectively accommodate varying length fire events. One possible avenue to approach this problem is to train models on a large number of simulated fires and then fine-tune the models on actual fire events (Burge et al., 2020).This transfer learning approach can help us mitigate the limitations in access to high-quality fire data while improving the fidelity of model simulations of wildfire spread. Although our initial attempt of predicant fire spread using data derived from MODIS burned area product and ConvLSTMs yielded limited success, the lessons we learned from this exercise can help us develop more effective ML predictions in the future.

Code Availability:

Project code is available at: https://github.com/JianingFang/FireML_Submission/tree/jianing

References

Barzegar, R., Aalami, M. T., & Adamowski, J. (2020). Short-term water quality variable prediction using a hybrid CNN–LSTM deep learning model. *Stochastic Environmental Research and Risk Assessment*, *34*(2), 415–433. https://doi.org/10.1007/s00477-020-01776-2

Bengio, S., Vinyals, O., Jaitly, N., & Shazeer, N. (2015). Scheduled Sampling for Sequence Prediction with Recurrent Neural Networks. In *arXiv*. http://arxiv.org/abs/1506.03099

Burge, J., Bonanni, M., Ihme, M., & Hu, R. L. (2020). Convolutional LSTM Neural Networks for Modeling Wildland Fire Dynamics Running Head: ConvLSTMs for Wildland Fire Dynamics. In *arXiv preprint*.

Hochreiter, S., & Urgen Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Computation*, *9*(8), 1735–1780. http://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf

Kim, T. Y., & Cho, S. B. (2019). Predicting residential energy consumption using CNN-LSTM neural networks. *Energy*, *182*, 72–81. https://doi.org/10.1016/j.energy.2019.05.230

Krawchuk, M. A., Moritz, M. A., Parisien, M. A., van Dorn, J., & Hayhoe, K. (2009). Global pyrogeography: The current and future distribution of wildfire. *PLoS ONE*, *4*(4). https://doi.org/10.1371/journal.pone.0005102

Moritz, M. A., Batllori, E., Bradstock, R. A., Gill, A. M., Handmer, J., Hessburg, P. F., Leonard, J., McCaffrey, S., Odion, D. C., Schoennagel, T., & Syphard, A. D. (2014). Learning to coexist with wildfire. In *Nature* (Vol. 515, Issue 7525, pp. 58–66). Nature Publishing Group. https://doi.org/10.1038/nature13946

Parks, S. A., Miller, C., Abatzoglou, J. T., Holsinger, L. M., Parisien, M. A., & Dobrowski, S. Z. (2016). How will climate change affect wildland fire severity in the western US? *Environmental Research Letters*, *11*(3). https://doi.org/10.1088/1748-9326/11/3/035002

Reichstein, M., Camps-Valls, G., Stevens, B., Jung, M., Denzler, J., Carvalhais, N., & Prabhat. (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, *566*(7743), 195–204. https://doi.org/10.1038/s41586-019-0912-1

Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W.-K., & Woo, W.-C. (2015). Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting. *Advances in Neural Information Processing Systems*, 802–810.

Sullivan, A. L. (2009a). Wildland surface fire spread modelling, 1990 - 2007. 1: Physical and quasi-physical models. *International Journal of Wildland Fire*, *18*(4), 349. https://doi.org/10.1071/wf06143

Sullivan, A. L. (2009b). Wildland surface fire spread modelling, 1990 - 2007. 2: Empirical and quasi-empirical models. *International Journal of Wildland Fire*, *18*(4), 369. https://doi.org/10.1071/wf06142

Sullivan, A. L. (2009c). Wildland surface fire spread modelling, 1990 - 2007. 3: Simulation and mathematical analogue models. *International Journal of Wildland Fire*, *18*(4), 387. https://doi.org/10.1071/wf06144

Westerling, A. L., & Bryant, B. P. (2007). Climate change and wildfire in California. *Climatic Change*, *87*(1 SUPPL). https://doi.org/10.1007/s10584-007-9363-z

Yusuf, S. A., Alshdadi, A. A., Alassafi, M. O., AlGhamdi, R., & Samad, A. (2021). Predicting catastrophic temperature changes based on past events via a CNN-LSTM regression mechanism. *Neural Computing and Applications*, *33*(15), 9775–9790. https://doi.org/10.1007/s00521-021-06033-3