Cartography M.Sc.

Predicting, understanding, and visualizing fire dynamics with neural networks

by **Larissa Saad**

Supervisor: JProf. Dr. Matthias Forkel (TU Dresden)
Reviewer: Tichaona Tavare Mukunga MSc. (TU Wien)

Conducted at the Institute of Cartography and the Institute of Photogrammetry and Remote Sensing, Department of Geosciences

TECHNISCHE UNIVERSITÄT DRESDEN

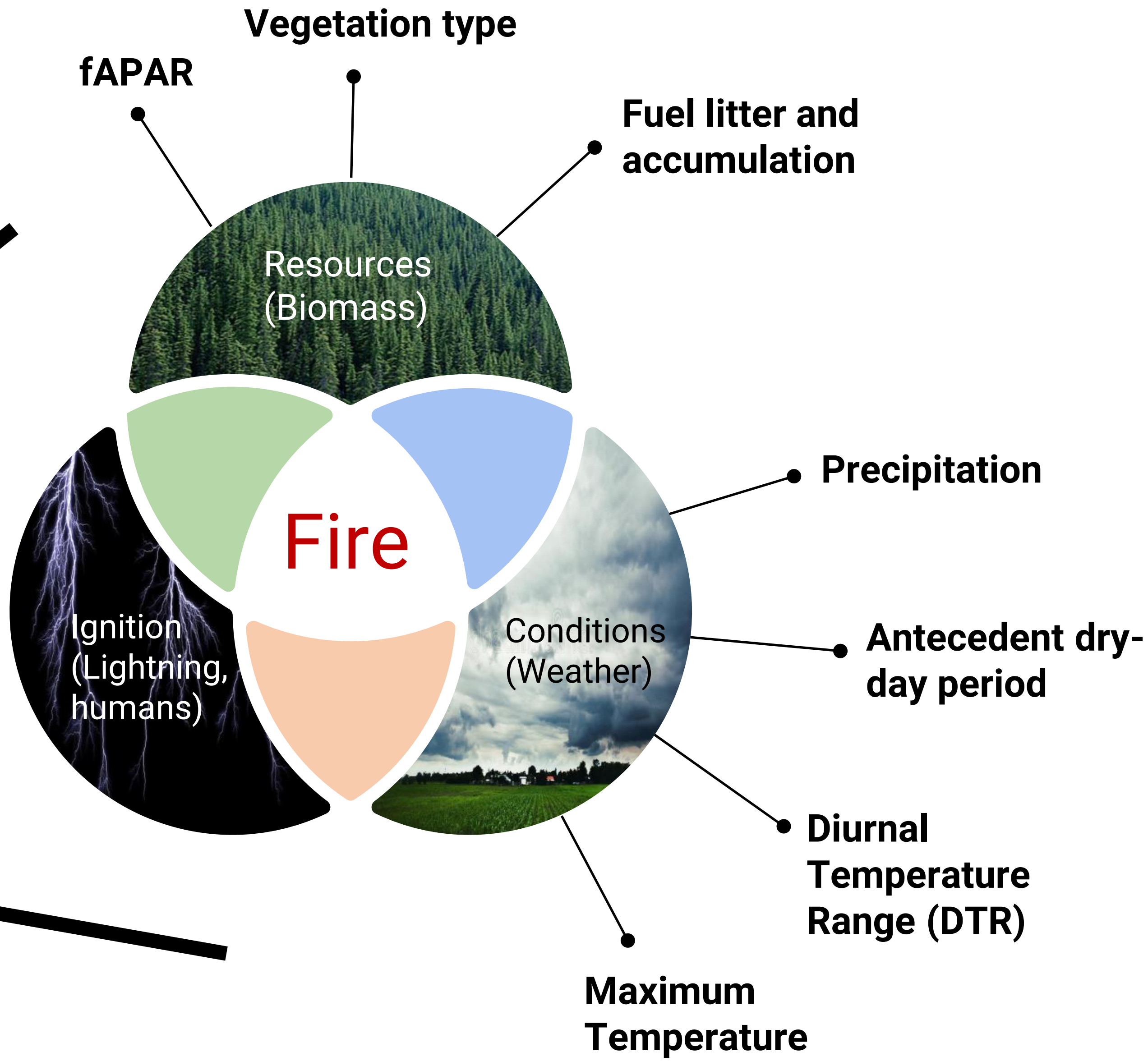
Outline

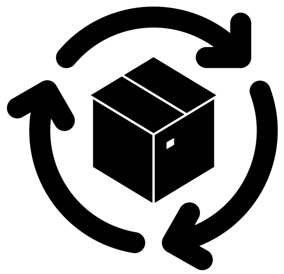
- Introduction and related work
- Research questions
- Methodology
- Results and discussion
- Conclusions
- References

Fire in the Earth system

The most important variables to predict fire

(Aldersley et al., 2011; Archibald et al., 2009; Bistinas et al., 2014; Forkel et al., 2019a; Kuhn-Régnier et al., 2020)





Process-based models– ML models

Predicting Fire

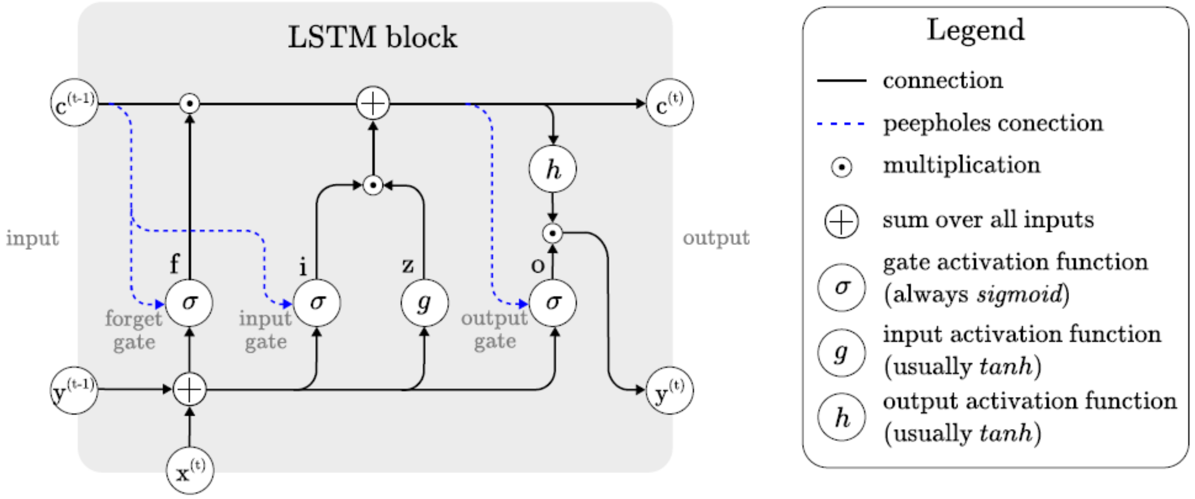
With Process-oriented models

- Fire models coupled with DGVMs / TEMs (Hantson et al., 2016)
- Different results in prediction of future trends (burned areas) (Andela et al., 2017; Forkel et al., 2019b)
- Fire-predictor relationships are not presented correctly (Forkel et al., 2019a)

With Machine Learning models

- Random Forests (Archibald et al., 2009; Aldersley et al., 2011; Forkel et al., 2019a)
- Support Vector Machines (SVMs) (Cortez et al., 2007)
- Neural Networks
  - Feed-forward neural networks (Maeda et al., 2009; Özbayoğlu et al., 2012; Satir et al., 2016; Joshi et al., 2021)
  - High accuracy
  - Challenges with time series with extreme or sudden events

Long-short term memory (LSTM) neural networks



LSTM cell structure (Van Houdt et al., 2020)

LSTM for fire prediction:

- Wildfire scale classification (Liang et al., 2019)
- Wildfire duration and direction (Perumal et al., 2020)
- Burned area map generation based on fire indexes (Kong et al., 2018)

RO1

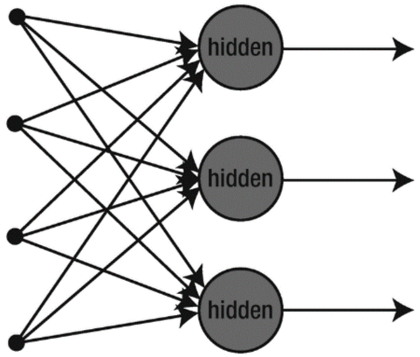
- Explore LSTM predictive ability of fire occurrence

- Detect LSTM ability to correctly capture the relationships between fire and driving factors

Visualization of machine learning models

Interpretation and visualization of machine learning models means determining the most important variables and depicting the predictor-response relationship

Blackbox models



Neural Networks

Interpretation

Model-agnostic methods

RO2

Investigate current interpretation techniques and their ability to characterize

- Global feature importance
- Feature importance spatial distribution
- Feature-output relationships
- Features interactions

Outline

- Introduction and related work
- **Research questions**
- Methodology
- Results and discussion
- Conclusions
- References

Q1: What are the opportunities and limitations of using LSTM neural networks to predict fire occurrence?

Q2: What is the ability of LSTM to record the relationships of fire drivers?

Q3: What is the best available method to interpret and visualize LSTM neural networks in an efficient and understandable way?

Outline

- Introduction and related work
- Research questions
- **Methodology**
  - Fire predictors and datasets
  - Study area
  - Data pre-processing
  - LSTM architecture
  - Visualization techniques
- Results and discussion
- Conclusions
- References

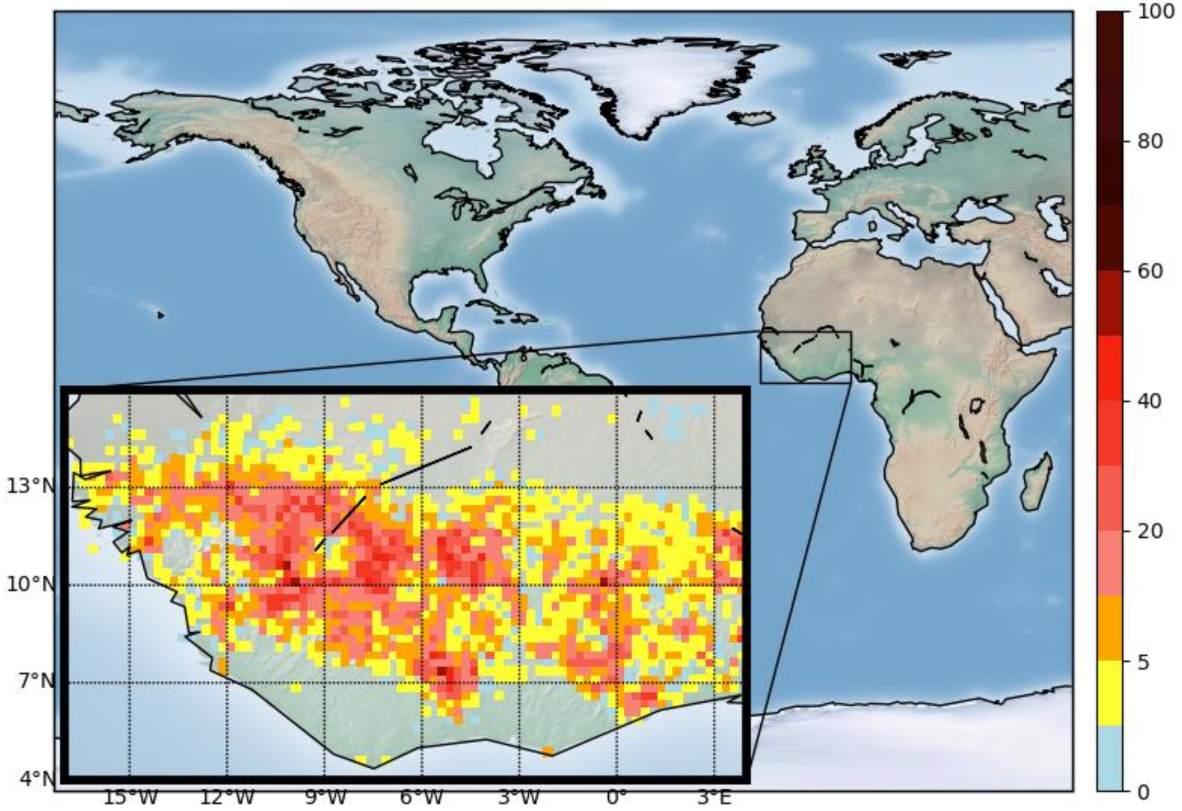
Fire predictors and datasets

Variable	Description	Data source	Measuring unit	Spatial Resolution	Temporal Resolution
Predictors					
Tmax	Mean of monthly maximum temperature	CRU JRA v2.0 (Harris, 2019)	Kelvin	0.5x0.5	Monthly
Tmin	Mean of monthly minimum temperature	CRU JRA v2.0 (Harris, 2019)	Kelvin	0.5x0.5	Monthly
DTR	Diurnal temperature range	CRU JRA v2.0 (Harris, 2019)	Kelvin	0.5x0.5	Monthly
Pre	Total precipitation	CRU JRA v2.0 (Harris, 2019)	kg/m2	0.5x0.5	Monthly
NI	Nesterov Index	CRU JRA v2.0 (Harris, 2019)	-	0.5x0.5	Monthly
Wind	Wind Speed	CRU JRA v2.0 (Harris, 2019)	m/s	0.5x0.5	Monthly
fAPAR	Fraction of Absorbed Photosynthetically Active Radiation	MODIS: MOD15A2H (Myneni et al., 2015)	-	0.25x0.25	Monthly
Target variables					
Fire ignitions	Count of ignition points per cell	Fire Atlas (Andela et al., 2019)	-	0.25x0.25	Monthly

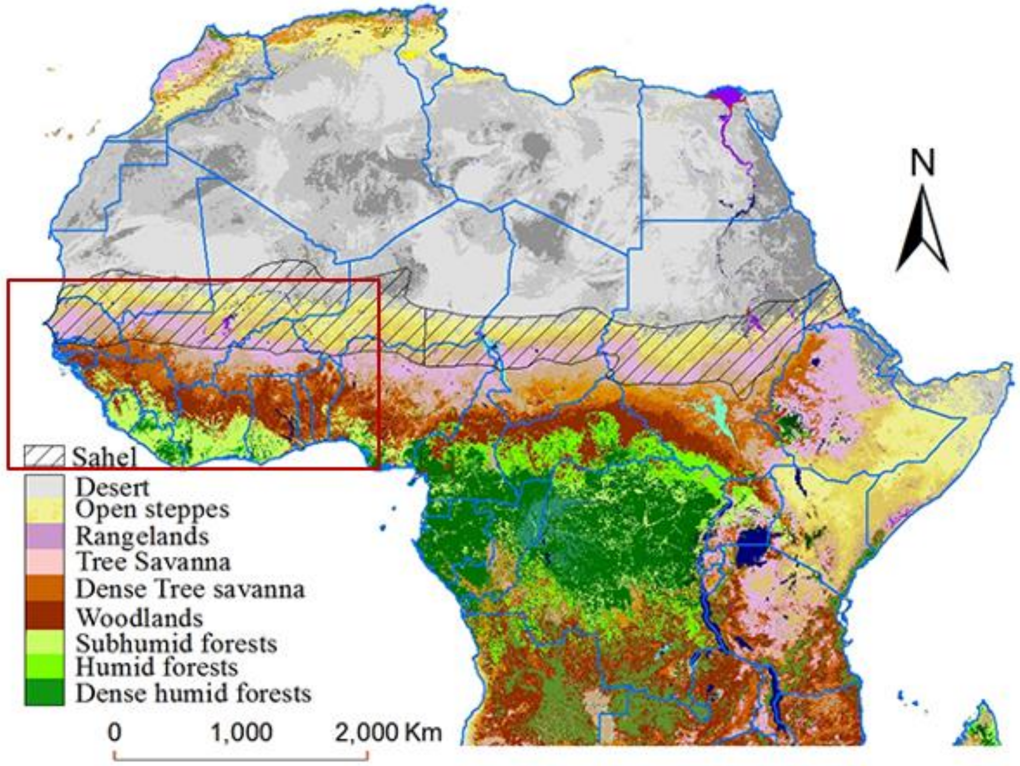
The Nesterov Index: $NI = \sum_{i=1}^{w} T_i(T_i - D_i)$

W = number of days since last rainfall > 3 mm
T = midday temperature (°C)
D = dew point temperature (°C)

Study Area



Study Area with a sample of ignition count for one month



Land cover map in the Sahel zone. The red rectangle was added to indicate the location of the study area in this thesis. Source: From GLC 2000, EU-JRC data (Mbow, 2017)

Data pre-processing

Data aggregation and resampling

0.25 deg latitude x 0.25 deg longitude grid

Merge

One NetCDF file with 7 variables

Missing values

Variable	Count of Missing Values
DTR	0
fAPAR	524669
Ignitions	2025479
NI	0
Pre	0
Tmax	0
Tmin	0
Wind	0

Multicollinearity test

$$VIF = \frac{1}{1 - R_i^2} \quad for\ i = 1,2,\ldots,k$$ (Shrestha, 2020)

Variable	VIF	VIF after dropping tmin
Wind	1.96	1.96
DTR	538577.07	1.66
fPAR	2.91	2.91
NI	1.31	1.31
Pre	2.09	2.09
tmax	2257348.76	1.49
tmin	2229555.77	

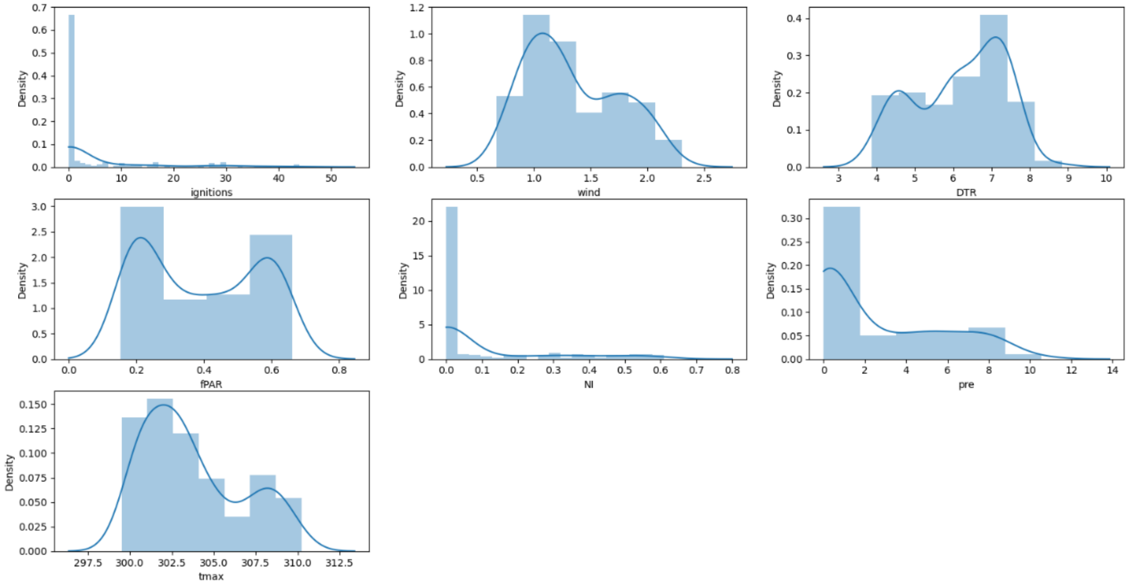
Data pre-processing

Data aggregation and resampling

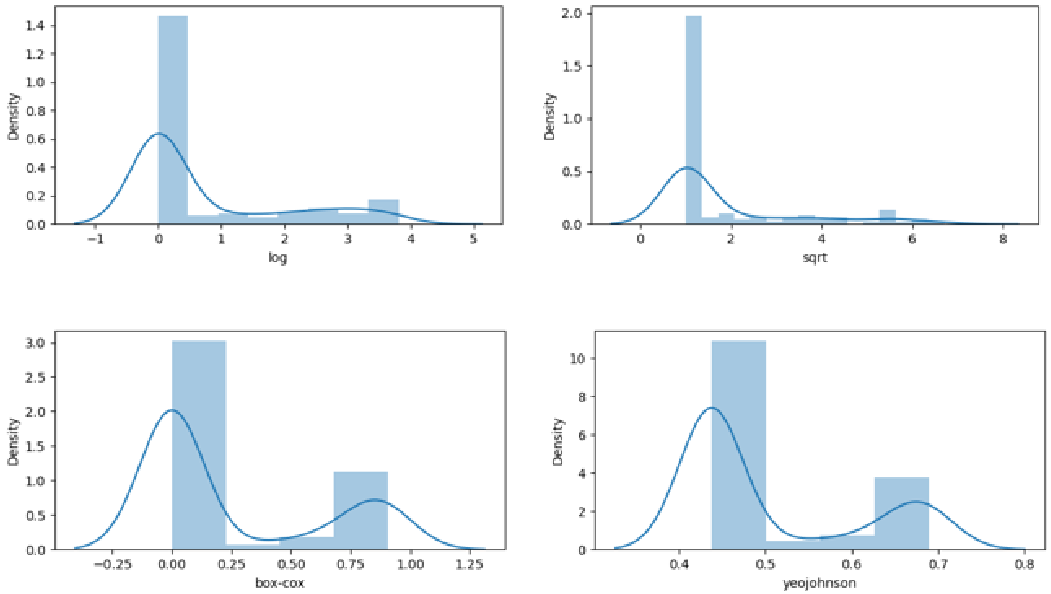
Missing values

Multicollinearity test

Data transformation



Histogram plots with kernel density for all variables of a random pixel in the dataset



The effect of applying multiple transformation technique on the highly skewed output variable (ignitions count)

Data pre-processing

Data aggregation and resampling

Missing values

Multicollinearity test

Data transformation

Feature scaling

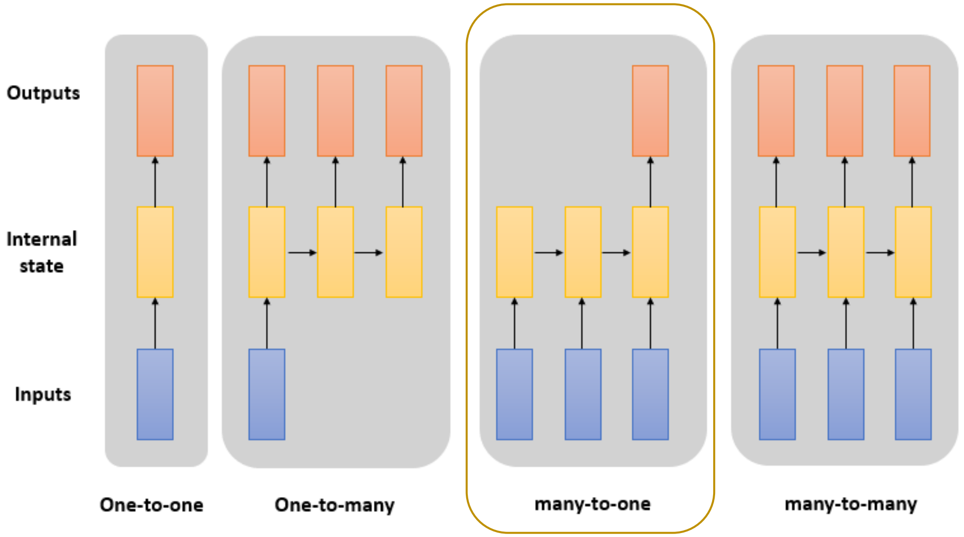
Min-Max Scaler within range [0, 1]

$$x_{new} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

LSTM architecture

Chosen methodology:
Pixel-based LSTM

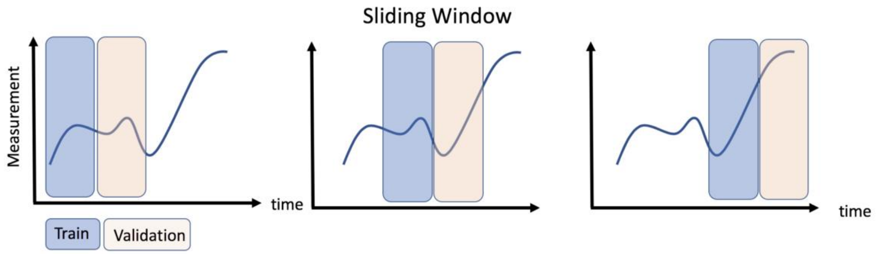
Model type



Data transformation:
Log transformation

Train-test split:
Training set: 01/2003 – 09-2015
Test set: 10/2015 – 09/2016

3D data structure using sliding window approach:
[samples, timesteps, features]



LSTM architecture

Experiment	RMSE	MAE
Different Number of hidden layers		
One LSTM layer	**5.915882**	2.801811936
Two hidden layers	6.239518	2.867196477
Three hidden layers	6.499306	3.01
Different time lags		
One layer 6 months	5.915882	2.801811936
One layer 12 months	**5.760793**	2.71326
One layer 18 months	5.815041	2.759578
One layer 24 months	5.766851	2.734103
Different loss Functions		
RMSE	5.760793	2.71326
MAE	**5.715536**	2.6674773
Different types of LSTM		
Vanilla LSTM	**5.715536**	2.6674773
Bidirectional LSTM	5.807794393	2.707526329

Setup tests results for different parameters to determine LSTM structure

hyperparameter	Value
Learning rate	0.001
Batch size	12
Window size	12
Loss function	Mean Absolute Error (MAE)
Activation function	Relu
Optimizer	Adam
Hidden layers	1
Input data size	(12,12,7)
Drop out	True (0.2)
Feature scaling	True [0,1]

LSTM neural network hyperparameters used in this thesis

Visualization techniques of LSTM-based model

1. **Permutation feature importance**

2. **Variance-based Feature Importance** (de Sá, 2019)

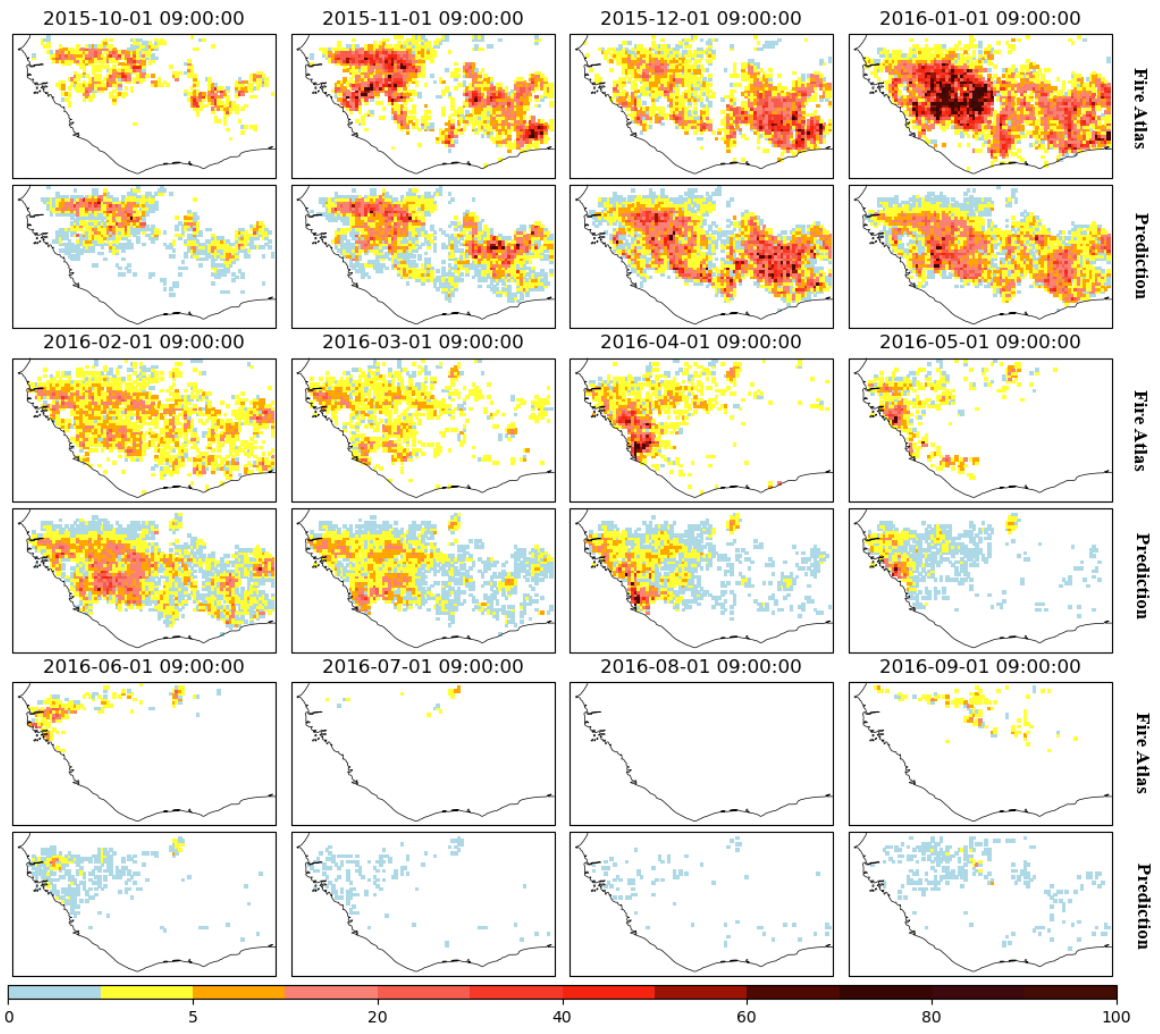
3. **Explaining model predictions through explanation method**

    - **SHAP (DeepSHAP Explainer)** (Lundberg et al., 2017)

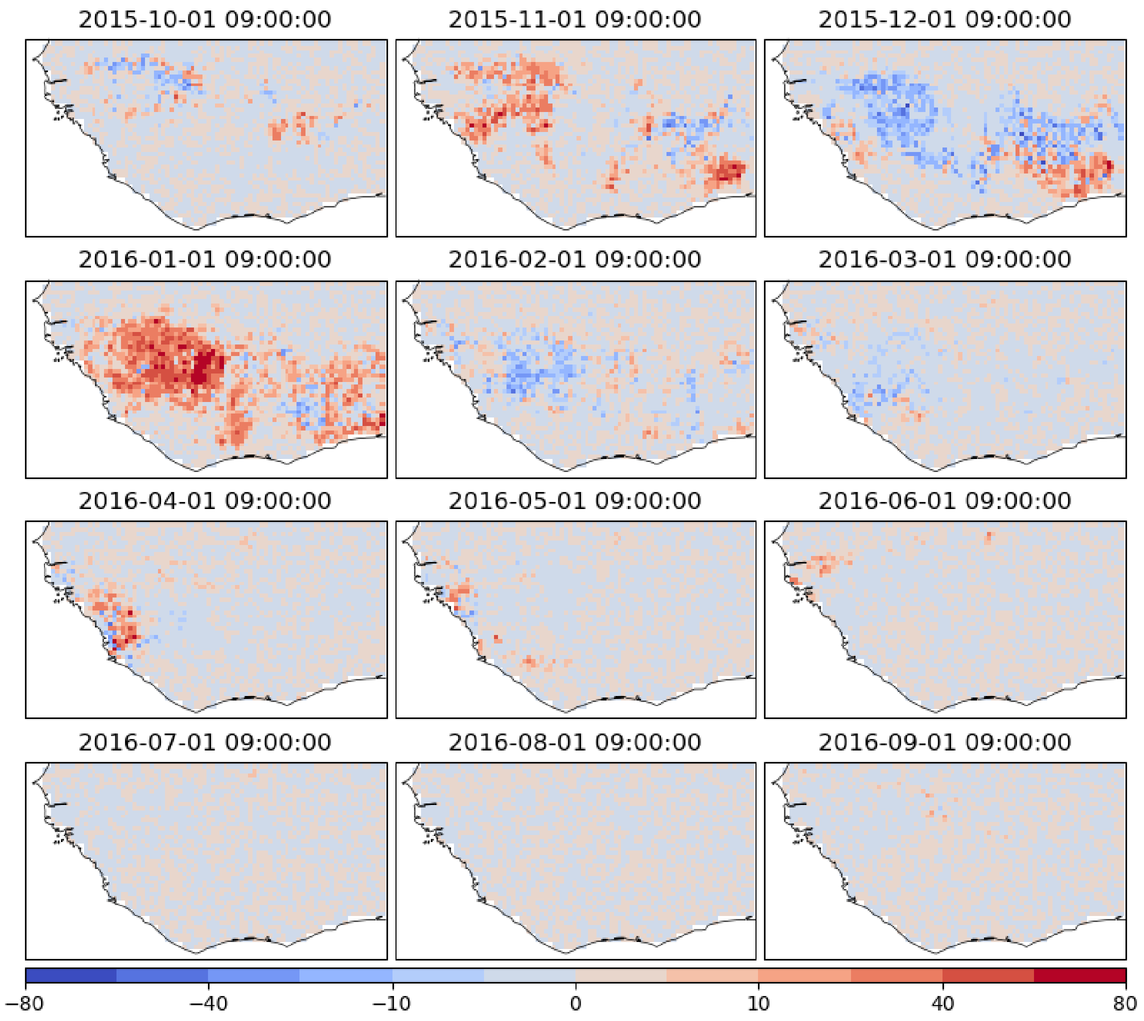
Outline

- Introduction and related work
- Research questions
- Methodology
- **Results and discussion**
- Conclusions
- References

Model evaluation



Comparison between the results of LSTM neural network and Fire Atlas data for one year prediction

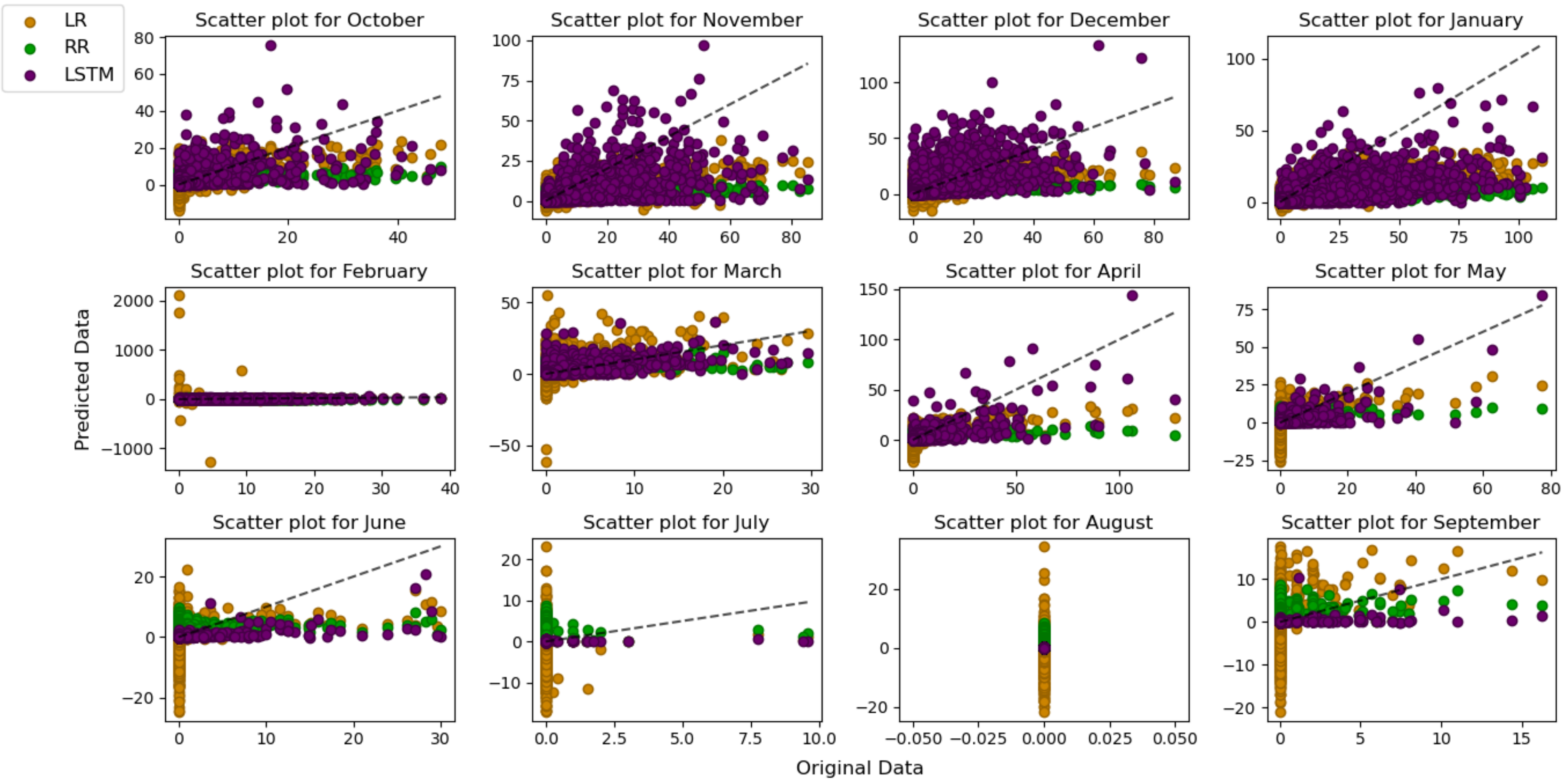


The difference between original and predicted fire ignitions for one year prediction

Model evaluation

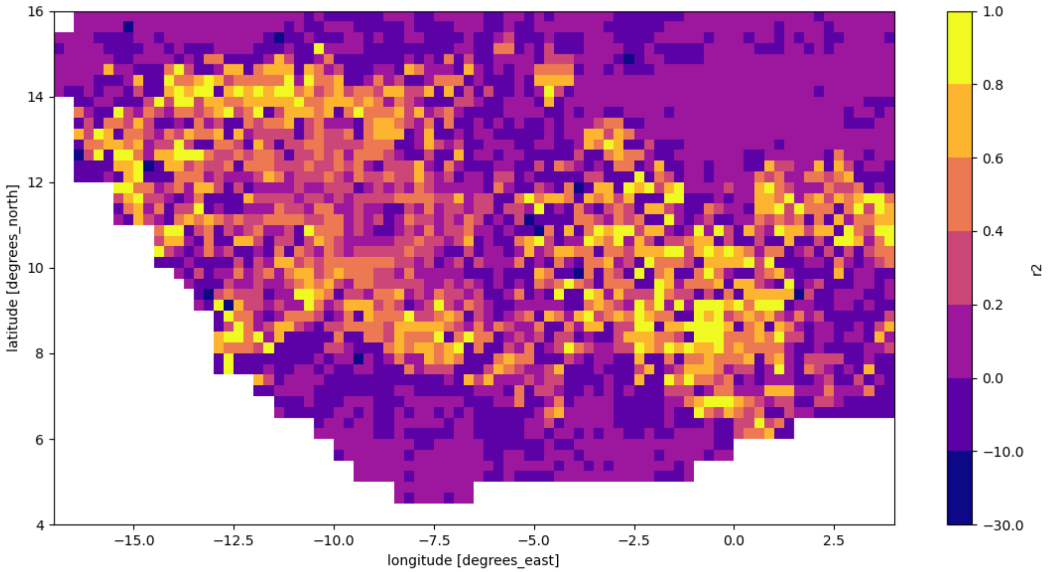
Prediction for one year	RMSE	MAE
LSTM	**3.333**	**1.509**
Linear Regression	4.48353	2.8453
Ridge Regression	4.0209	2.5585

Comparison of RMSE and MAE for one year prediction for the entire study area

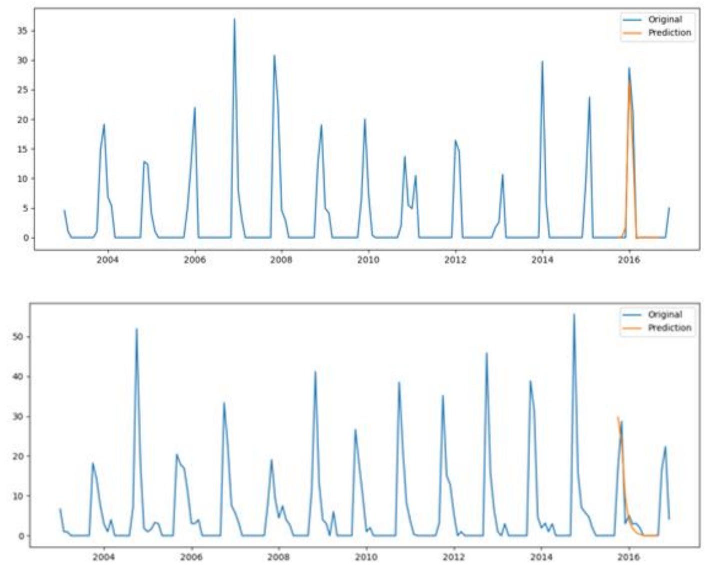


Comparative scatter plots of LSTM predictions (purple) with two baseline models, Linear Regression (LR) (orange) and Ridge Regression (RR) (green)

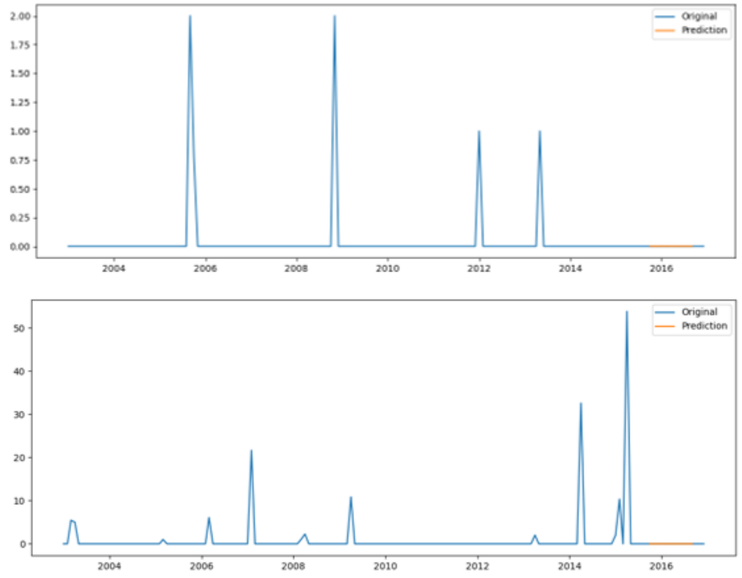
Model evaluation



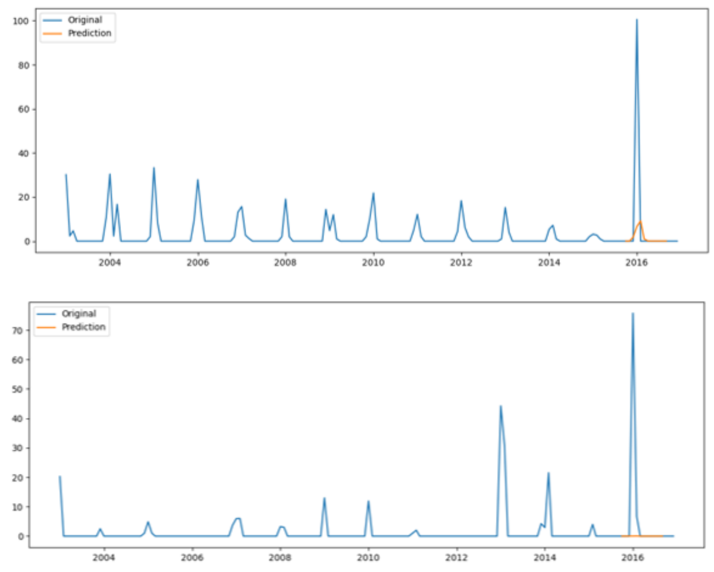
Pixel-based map of coefficient of determination (R2) values for one year prediction



Two samples where R2 values are high and fire ignitions are more frequent and annual

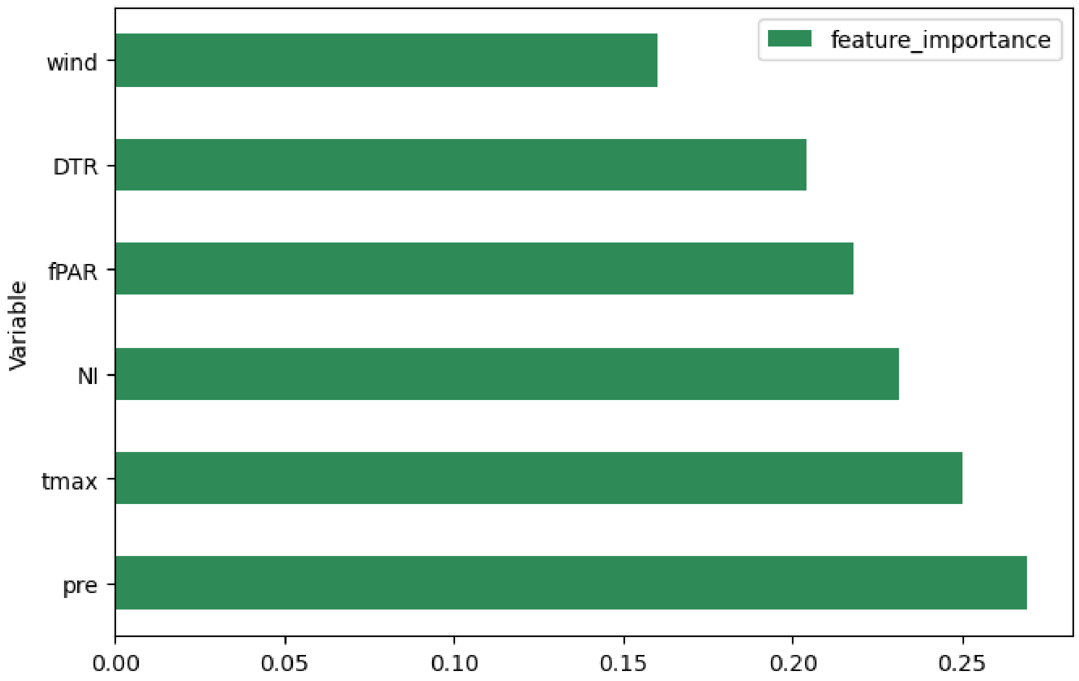


Two samples where R2 values are high and fire ignitions are rare

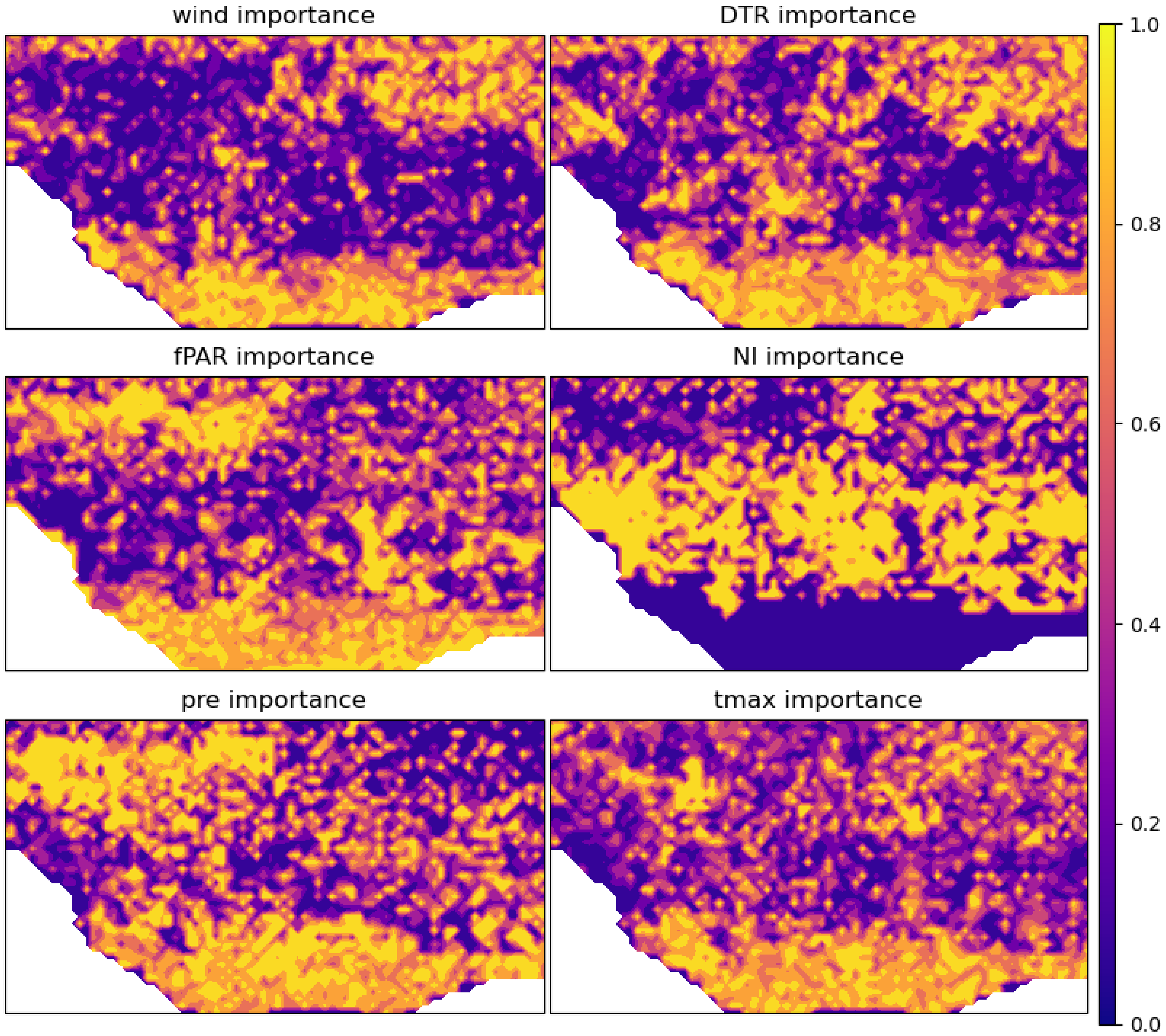


Two samples where R2 values are low. Fire occurrence has different frequencies but with extreme values in the last fire season

Comparison of visualization techniques for explainable LSTM-based fire modelling



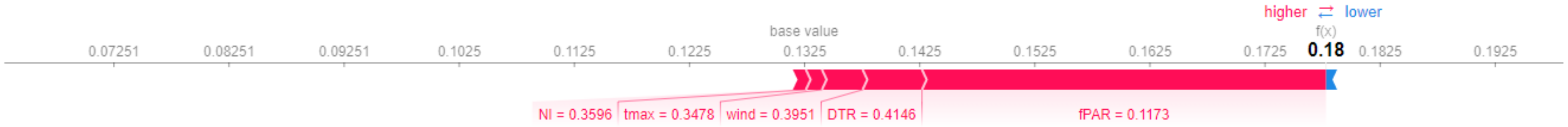
Permutation feature importance for the entire study area. The error increase is represented as percentage of the original RMSE of the model



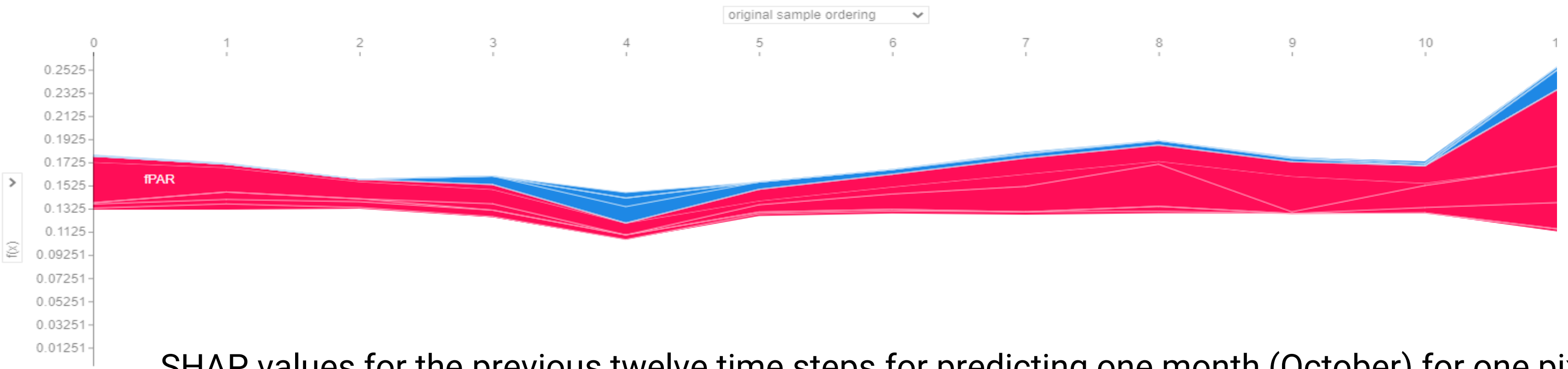
Variance-based feature importance for the entire study area. The relative importance of each feature in each pixel is represented as a percentage

Comparison of visualization techniques for explainable LSTM-based fire modelling

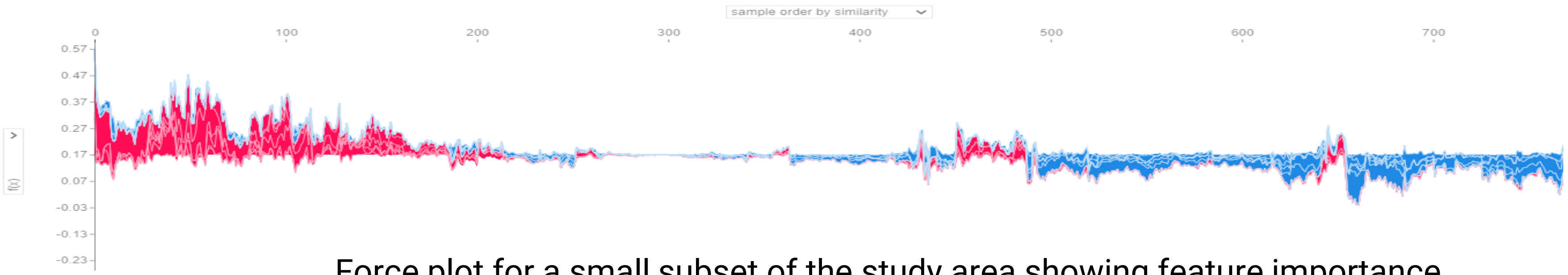
SHAP Force plots



SHAP values for the first time steps for predicting one month (October) for one pixel



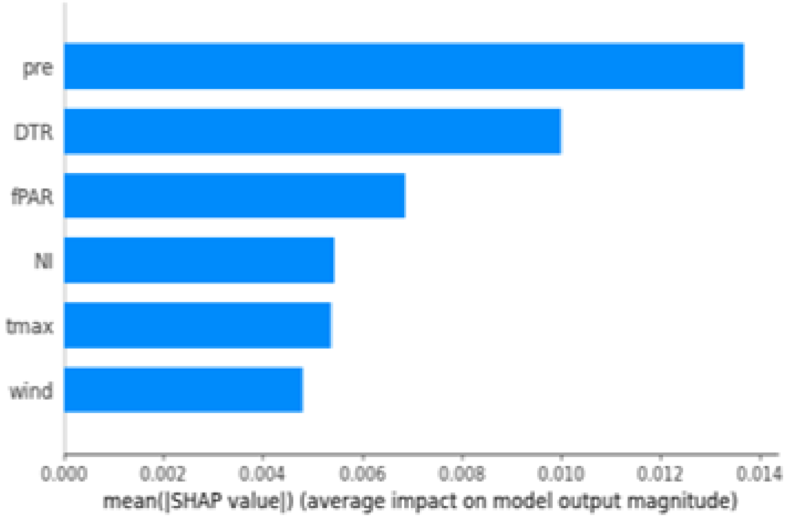
SHAP values for the previous twelve time steps for predicting one month (October) for one pixel

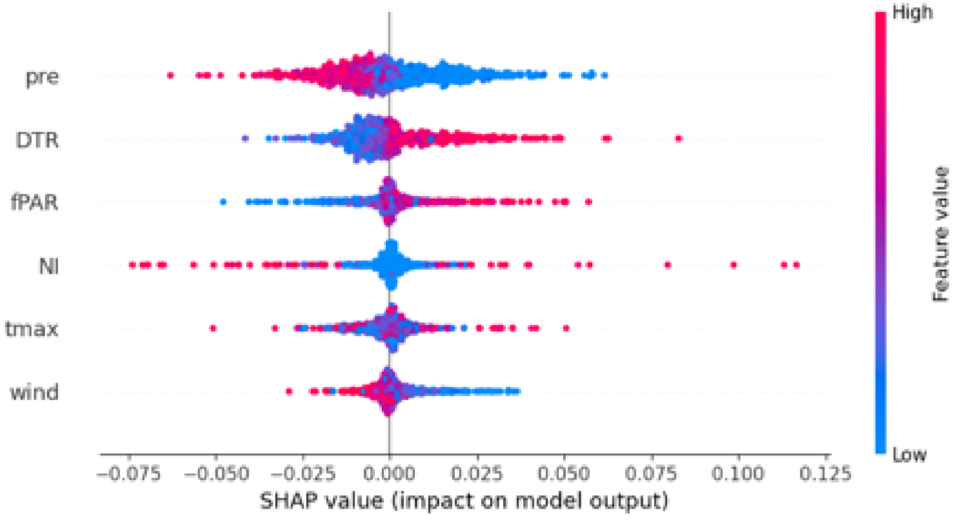


Force plot for a small subset of the study area showing feature importance

Comparison of visualization techniques for explainable LSTM-based fire modelling

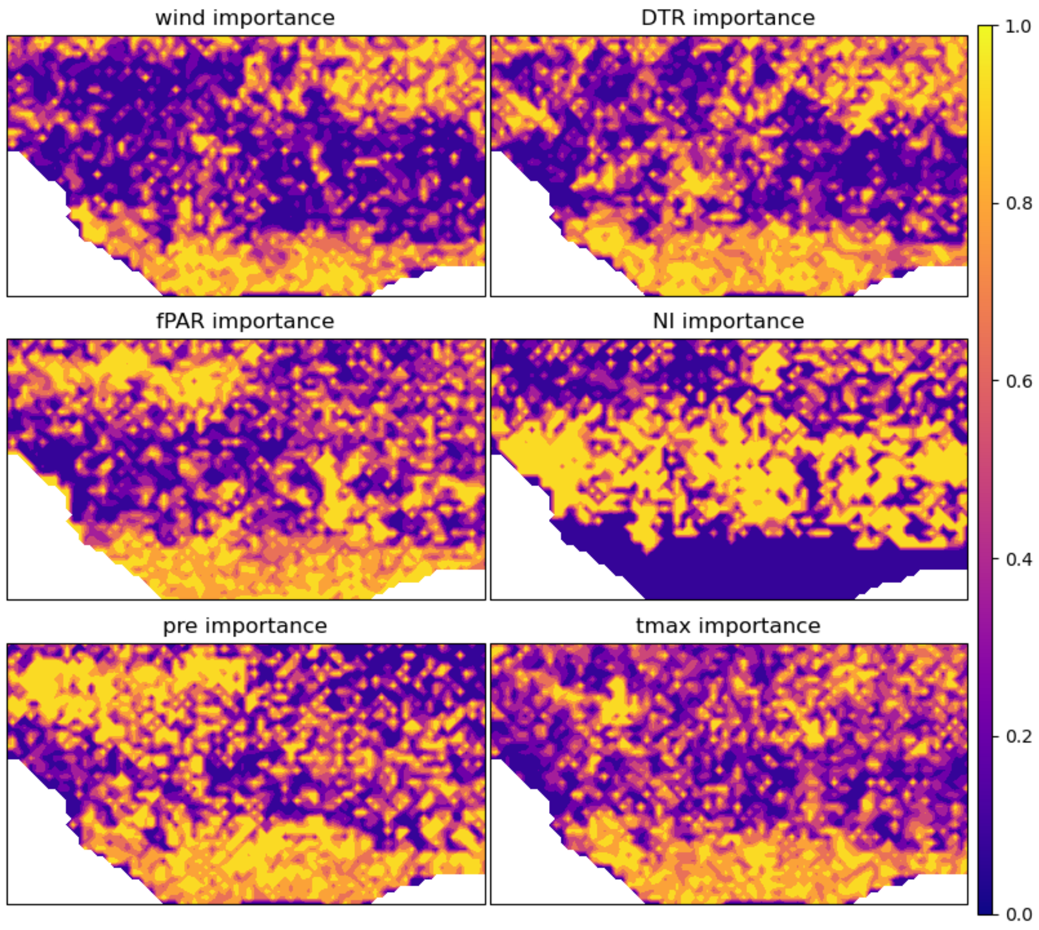
SHAP summary plots



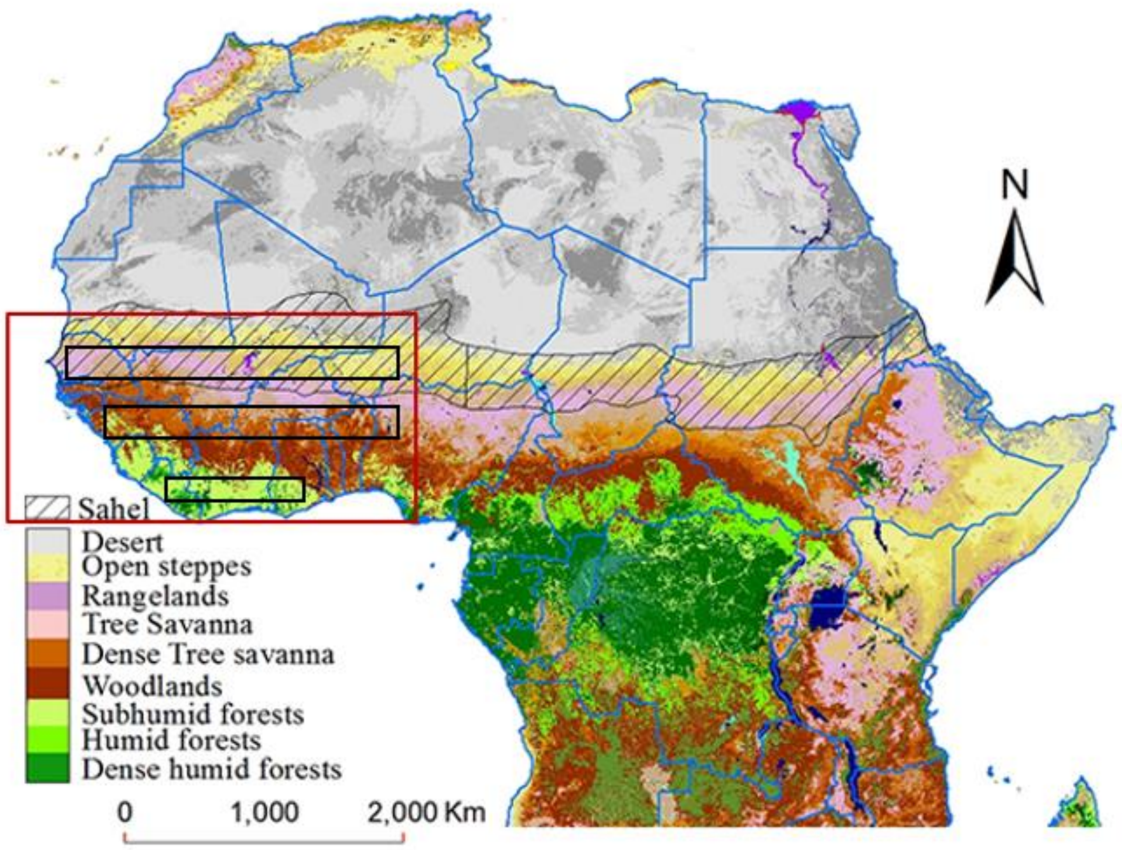


SHAP feature importance plot (left) and SHAP summary plot (right) for the precedent month

Importance of predictor variables

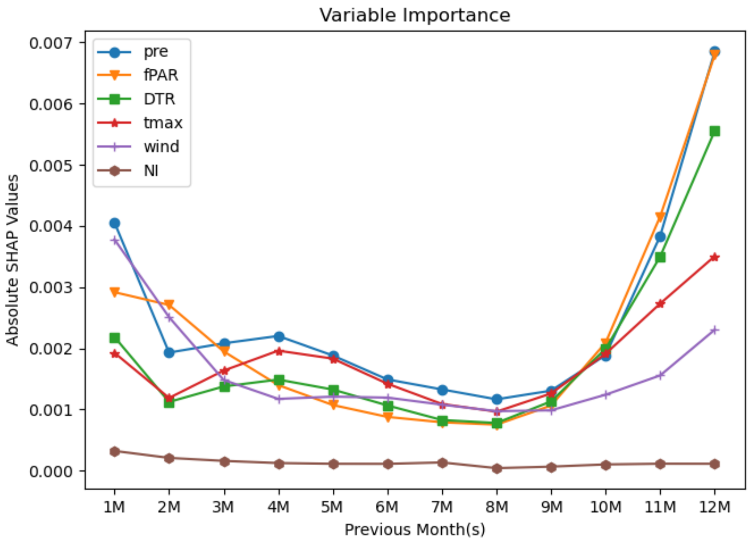


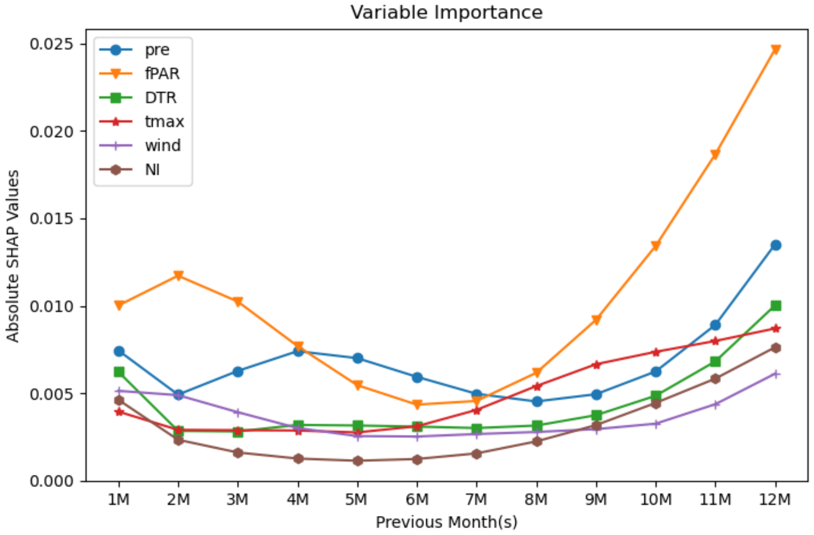
Spatial distribution of feature importance using variance-based feature importance method

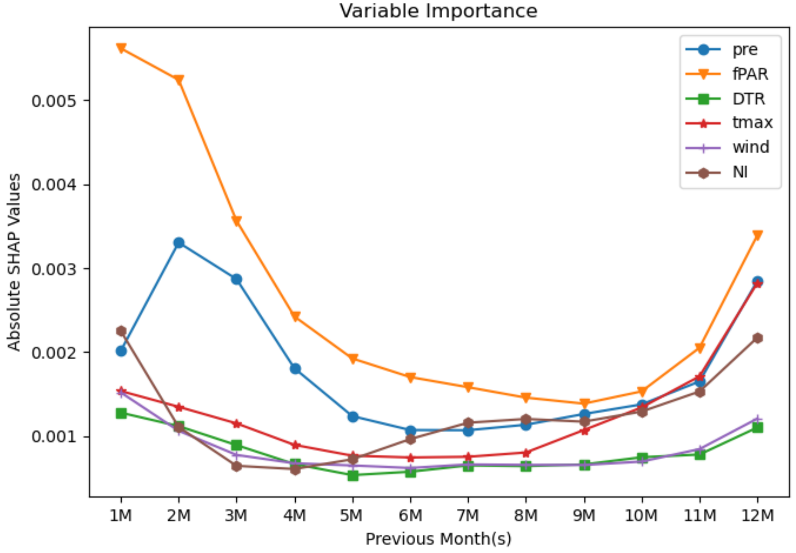


Sub-regions for SHAP feature importance analysis (black rectangles), taken from land cover map in the Sahel zone (Mbow, 2017). The red rectangle represents the entire study area.

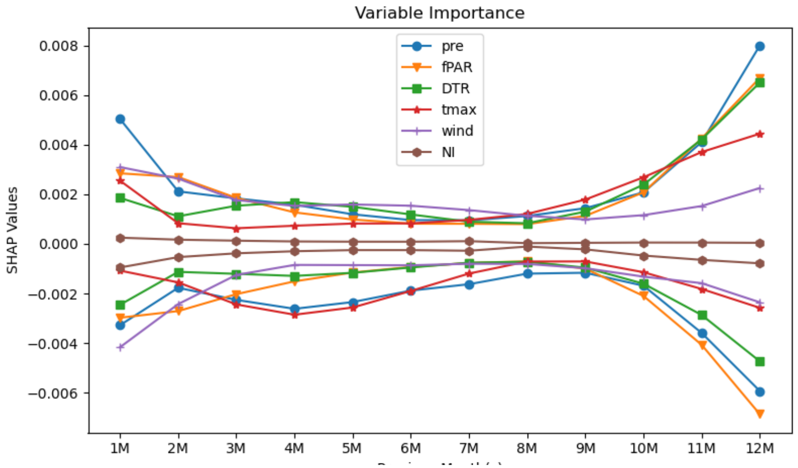
Southern region | Middle region | Northern region

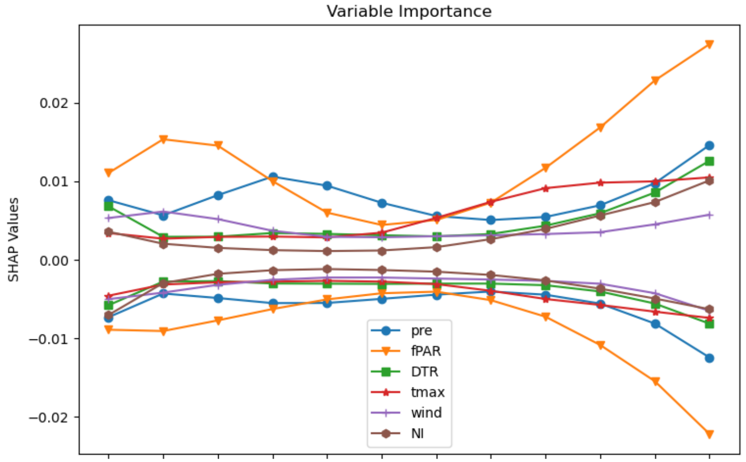


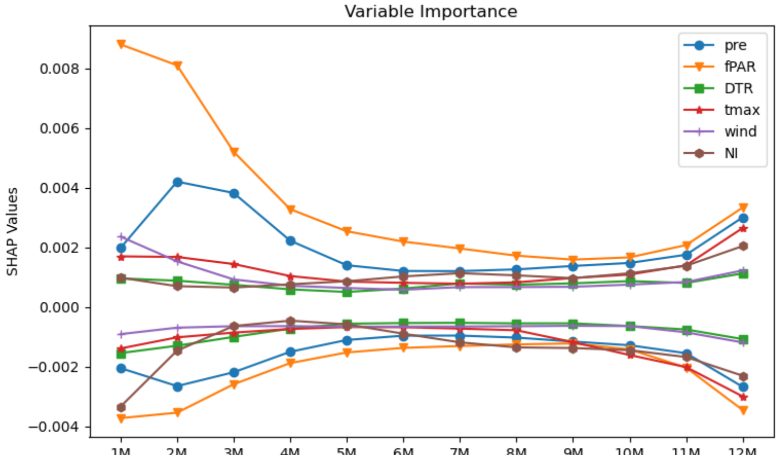




Feature importance for the previous 12 months using SHAP absolute values

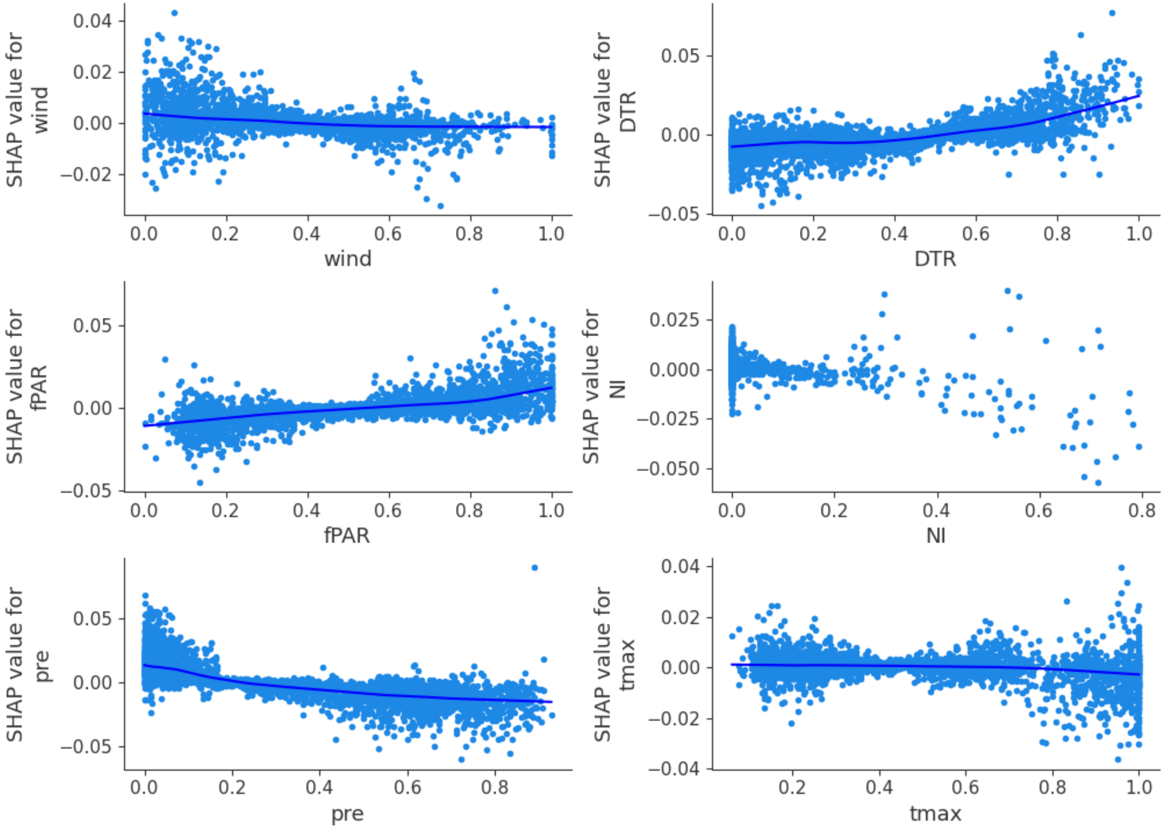




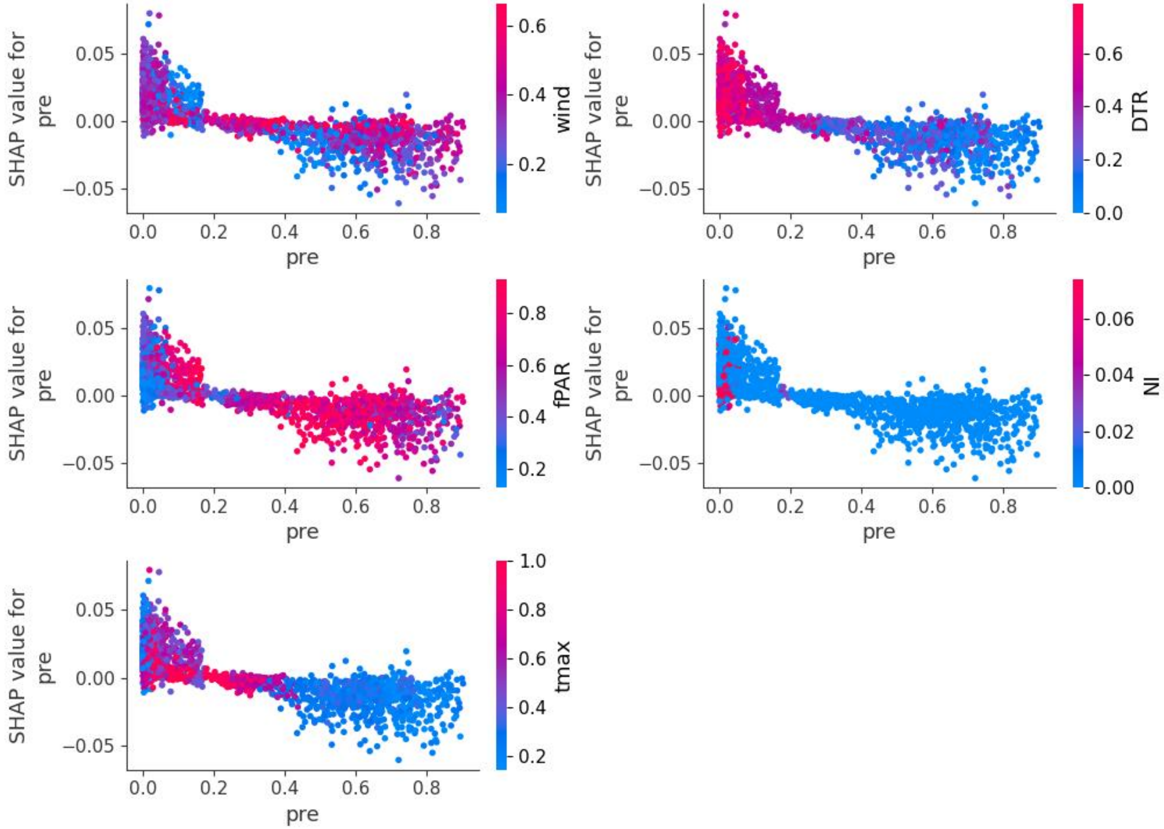


The negative and positive effect of each feature for the previous 12 months

Predictor-response relationships and interactions

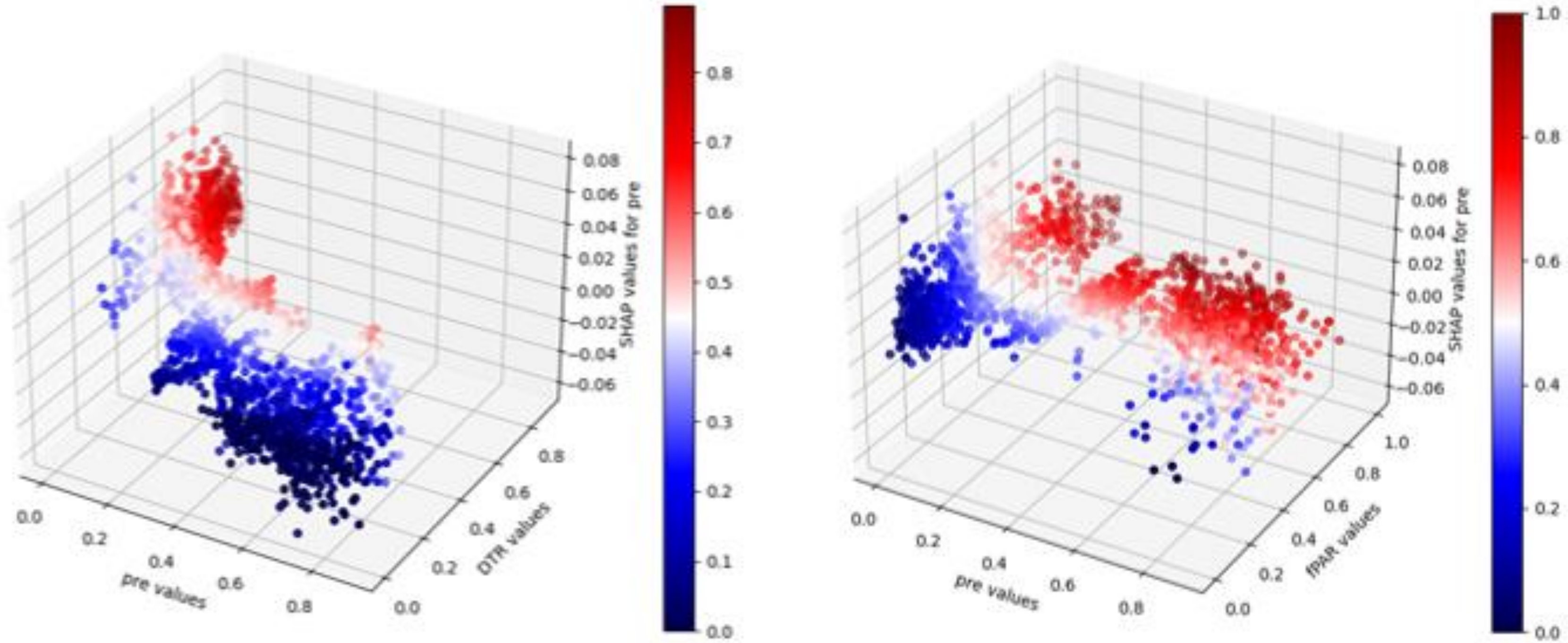


SHAP dependence plots



SHAP interaction plots for precipitation with the other variables

Predictor-response relationships and interactions



3D interaction plots. DTR-pre interaction plot (left) and fAPAR-pre interaction plot (right).

Outline

- Introduction and related work
- Research questions
- Methodology
- Results and discussion
- **Conclusions**
- References

Q1: What are the opportunities and limitations of using LSTM neural networks to predict fire occurrence?

- The pixel-based LSTM captured the seasonal and spatial varieties with RMSE value computed at 3.333 for the entire study area
- LSTM underestimated the high values of ignitions during the peak of fire season.
- LSTM was not able to capture the extreme values and performed better during the months of lower fire occurrence

Limitation:
- The inclusion of all important fire drivers as recommended in the literature was not possible
- Limited length of time series

Q2: What is the ability of LSTM to record the relationships of fire drivers?

- LSTM was able to model the fire-predictor relationship correctly only for precipitation, DTR and fAPAR
- Maximum temperature and wind the relationship were vague
- The Nesterov index did not play a major role for LSTM and no clear relationship was concluded from the model

The most important features to predict fire ignitions were mainly fAPAR, precipitation and maximum temperature. The order of importance for other variables differs based on location and precedent month.

Q3: What is the best available method to interpret and visualize LSTM neural networks in an efficient and understandable way?

- No general approach was able to visualize local and general feature importance

  - Global feature importance → Permutation feature importance
  - Feature importance spatial distribution → Variance-based feature importance
  - Precedent conditions → SHAP summary plots
  - Feature-output relationships → SHAP dependence plots
  - Features interactions → 3D SHAP dependence plots

- Visualization techniques have contributed to better understanding of the machine learning model and presented useful insights for further developments.

Future research

- More advanced types of LSTM such as Attention LSTM
- Inclusion of spatial information with convolutional neural networks (Multivariate spatio-temporal convolutional LSTM)
- One comprehensive model facilitates implementing different visualization techniques

Outline

- Introduction and related work
- Research questions
- Methodology
- Results and discussion
- Conclusions
- **References**

Aldersley, A., Murray, S. J., & Cornell, S. E. (2011). Global and regional analysis of climate and human drivers of wildfire. Science of the Total Environment, 409(18), 3472–3481. https://doi.org/10.1016/j.scitotenv.2011.05.032

Andela, N., Morton, D. C., Giglio, L., Chen, Y., Van Der Werf, G. R., Kasibhatla, P. S., DeFries, R. S., Collatz, G. J., Hantson, S., Kloster, S., Bachelet, D., Forrest, M., Lasslop, G., Li, F., Mangeon, S., Melton, J. R., Yue, C., & Randerson, J. T. (2017). A human-driven decline in global burned area. Science, 356(6345), 1356–1362. https://doi.org/10.1126/science.aal4108

Andela, N., Morton, D. C., Giglio, L., Paugam, R., Chen, Y., Hantson, S., Van Der Werf, G. R., & Anderson, J. T. (2019). The Global Fire Atlas of individual fire size, duration, speed and direction. Earth System Science Data, 11(2), 529–552. https://doi.org/10.5194/essd-11-529-2019

Archibald, S., Roy, D. P., van Wilgen, B. W., & Scholes, R. J. (2009). What limits fire? An examination of drivers of burnt area in Southern Africa. Global Change Biology, 15(3), 613–630. https://doi.org/10.1111/j.1365-2486.2008.01754.x

Bistinas, I., Harrison, S. P., Prentice, I. C., & Pereira, J. M. C. (2014). Causal relationships vs. emergent patterns in the global controls of fire frequency. Biogeosciences Discussions, 11(3), 3865–3892. https://doi.org/10.5194/bgd-11-3865-2014

Cortez, P., & Morais, A. (2007). A Data Mining Approach to Predict Forest Fires using Meteorological Data. Proceedings of 13th Portugese Conference on Artificial Intelligence, 512–523. http://www.dsi.uminho.pt/~pcortez/fires.pdf

de Sá, C. R. (2019). Variance-Based Feature Importance in Neural Networks. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11828 LNAI(October), 306–315. https://doi.org/10.1007/978-3-030-33778-0_24

Forkel, M., Andela, N., P Harrison, S., Lasslop, G., Van Marle, M., Chuvieco, E., Dorigo, W., Forrest, M., Hantson, S., Heil, A., Li, F., Melton, J., Sitch, S., Yue, C., & Arneth, A. (2019a). Emergent relationships with respect to burned area in global satellite observations and fire-enabled vegetation models. Biogeosciences, 16(1), 57–76. https://doi.org/10.5194/bg-16-57-2019

Forkel, M., Dorigo, W., Lasslop, G., Chuvieco, E., Hantson, S., Heil, A., Teubner, I., Thonicke, K., & Harrison, S. P. (2019b). Recent global and regional trends in burned area and their compensating environmental controls. Environmental Research Communications, 1(5), 051005. https://doi.org/10.1088/2515-7620/ab25d2

Hantson, S., Arneth, A., Harrison, S. P., Kelley, D. I., Prentice, I. C., Rabin, S. S., Archibald, S., Mouillot, F., Arnold, S. R., Artaxo, P., Bachelet, D., Ciais, P., Forrest, M., Friedlingstein, P., Hickler, T., Kaplan, J. O., Kloster, S., Knorr, W., Lasslop, G., … Yue, C. (2016). The status and challenge of global fire modelling. Biogeosciences Discussions, 2016(January), 1–30. https://doi.org/10.5194/bg-2016-17

Harris, I. C. (2019). CRU JRA v2.0: A forcings dataset of gridded land surface blend of Climatic Research Unit (CRU) and Japanese reanalysis (JRA) data. Centre for Environmental Data Analysis. https://catalogue.ceda.ac.uk/uuid/7f785c0e80aa4df2b39d068ce7351bbb

Joshi, J., & Sukumar, R. (2021). Improving prediction and assessment of global fires using multilayer neural networks. Scientific Reports, 11(1), 1–14. https://doi.org/10.1038/s41598-021-81233-4

Kong, Y. L., Huang, Q., Wang, C., Chen, J., Chen, J., & He, D. (2018). Long short-term memory neural networks for online disturbance detection in satellite image time series. Remote Sensing, 10(3). https://doi.org/10.3390/rs10030452

Kuhn-Régnier, A., Voulgarakis, A., Nowack, P., Forkel, M., Prentice, I. C., & Harrison, S. P. (2020). Quantifying the Importance of Antecedent Fuel-Related Vegetation Properties for Burnt Area using Random Forests. Biogeosciences Discussions, November, 1–24.

Liang, H., Zhang, M., & Wang, H. (2019). A Neural Network Model for Wildfire Scale Prediction Using Meteorological Factors. IEEE Access, 7, 176746–176755. https://doi.org/10.1109/ACCESS.2019.2957837

Lundberg, S. M., & Lee, S.-I. (2017). A Unified Approach to Interpreting Model Predictions. 31st International Conference on Neural Information Processing Systems, 4768–4777.

Maeda, E. E., Formaggio, A. R., Shimabukuro, Y. E., Arcoverde, G. F. B., & Hansen, M. C. (2009). Predicting forest fire in the Brazilian Amazon using MODIS imagery and artificial neural networks. International Journal of Applied Earth Observation and Geoinformation, 11(4), 265–272. https://doi.org/10.1016/j.jag.2009.03.003

Mbow, C. (2017). The Great Green Wall in the Sahel. Oxford Research Encyclopedia of Climate Science, August 2017, 1–24. https://doi.org/10.1093/acrefore/9780190228620.013.559

Miller, T. (2019). Explanation in artificial intelligence: Insights from the social sciences. Artificial Intelligence, 267, 1–38. https://doi.org/10.1016/j.artint.2018.07.007

Myneni, R., Knyazikhin, Y., & Park, T. (2015). MOD15A2H MODIS/Terra Leaf Area Index/FPAR 8-Day L4 Global 500m SIN Grid V006 [Data set]. NASA EOSDIS Land Processes DAAC. https://doi.org/https://doi.org/10.5067/MODIS/MOD15A2H.006

Özbayoğlu, A. M., & Bozer, R. (2012). Estimation of the burned area in forest fires using computational intelligence techniques. Procedia Computer Science, 12, 282–287. https://doi.org/10.1016/j.procs.2012.09.070

Perumal, R., & van Zyl, T. L. (2020). Comparison of Recurrent Neural Network Architectures for Wildfire Spread Modelling. ArXiv. https://doi.org/10.1109/SAUPEC/RobMech/PRASA48453.2020.9078028

Satir, O., Berberoglu, S., & Donmez, C. (2016). Mapping regional forest fire probability using artificial neural network model in a Mediterranean forest ecosystem. Geomatics, Natural Hazards and Risk, 7(5), 1645–1658. https://doi.org/10.1080/19475705.2015.1084541

Shrestha, N. (2020). Detecting Multicollinearity in Regression Analysis. American Journal of Applied Mathematics and Statistics, 8(2), 39–42. https://doi.org/10.12691/ajams-8-2-1

Van Houdt, G., Mosquera, C., & Nápoles, G. (2020). A review on the long short-term memory model. Artificial Intelligence Review, 53(8), 5929–5955. https://doi.org/10.1007/s10462-020-09838-1

Cartography M.Sc.

Thank you for your attention