Optimising Forest Fire Predictions for Sustainable Ecosystem Management: A Comparative Analysis of Traditional and Modern Machine Learning Algorithms

Harikesh[1] , Li-minn Ang (Kenneth)[1] , Sanjeev Srivastava[1]

Introduction

- Forest fires pose significant threats to human life, biodiversity, and properties, with increasing frequency and intensity due to climate change and human activities. Accurate prediction models are essential for timely interventions, resource allocation, and strategic planning to mitigate these impacts.
- Traditional models like linear regression and decision trees offer interpretability but struggle with complex interactions in forest fire data. Modern techniques such as Gradient Boosting Machines (GBM), Neural Networks (NN), XGBoost, and LightGBM handle large, complex datasets more effectively, providing higher predictive accuracy and robustness.
- There is limited research comparing the performance of traditional and modern machine learning models specifically for forest fire prediction. This study aims to fill this gap by evaluating these models using a standardized dataset to determine which offers the best predictive accuracy and computational efficiency.
- This study provides a comprehensive comparison of multiple machine learning algorithms for forest fire prediction, employing a standardised dataset with various environmental and geographical factors. By enhancing predictive capabilities, this research supports robust disaster preparedness and response efforts, aligning with sustainable ecosystem management.

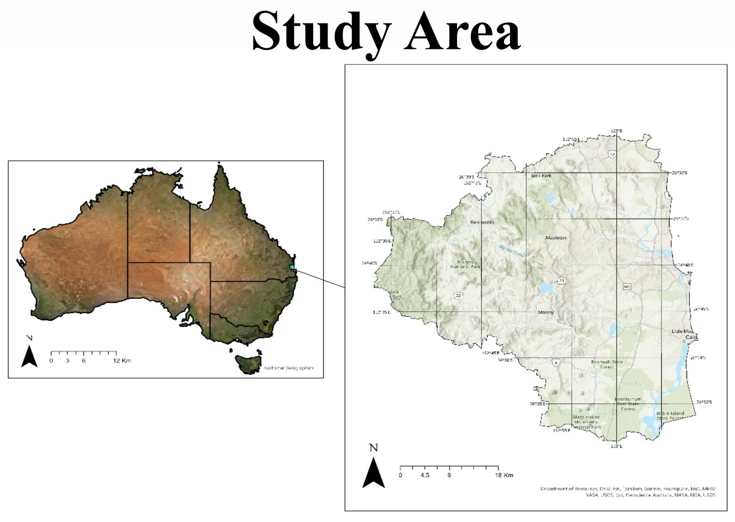
Objective

- Optimise forest fire predictions through comparative analysis of traditional and modern machine learning algorithms.
- Enhance accuracy and reliability of predictive models.

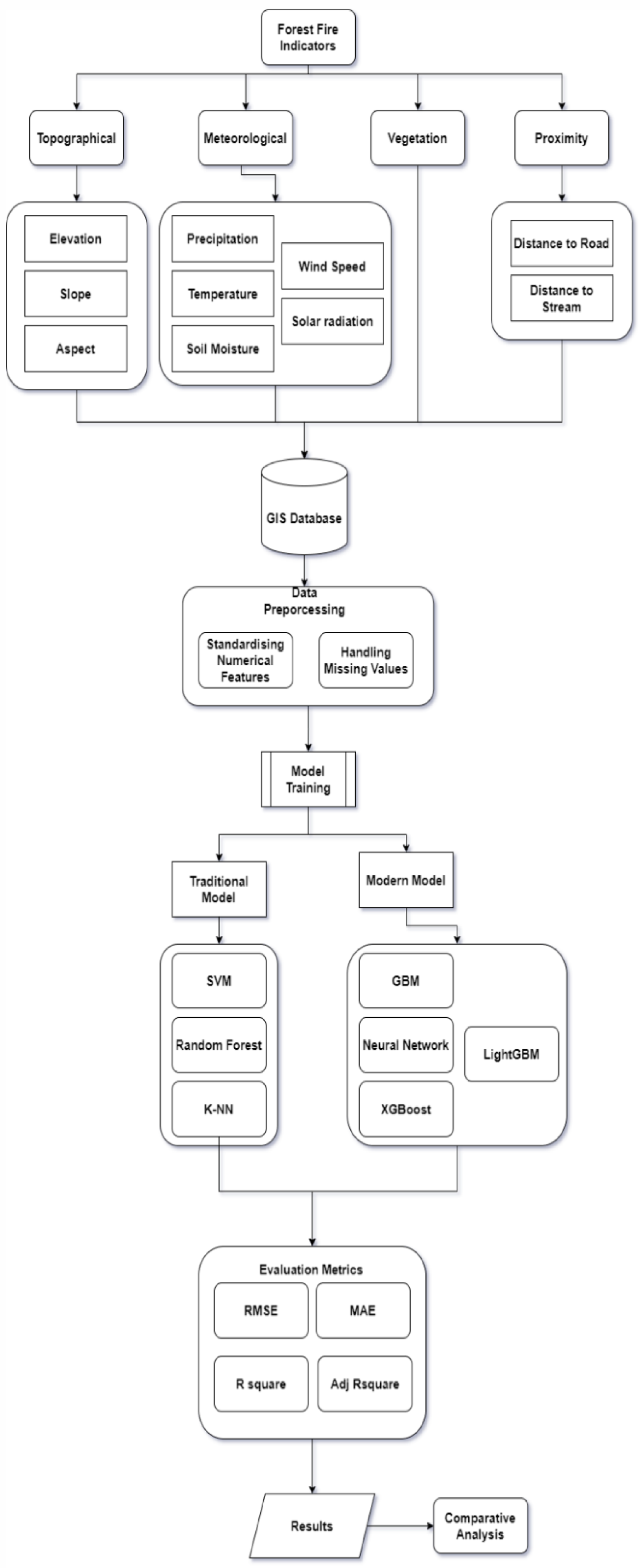
Data Inventory

Data	Variable	Description	Initial data format	Data source and map scale
Digital elevation model (DEM)	Elevation	Height above sea level	Raster	Geoscience Australia, LiDAR-derived 5m DEM
Digital elevation model (DEM)	Slope	Steepness or incline of the terrain	Raster	Derived from 5m DEM
Digital elevation model (DEM)	Aspect	Compass direction that a slope faces	Raster	Derived from 5m DEM
CHIRPS Daily	Precipitation	Amount of rainfall	Raster	4.8 km
Broad vegetation group (BVG)	Vegetation Type	Type of vegetation	Vector	Department of Environment and Science, 1:100-K
TerraClimate - Monthly	Soil Moisture	Moisture content of the soil	Raster	4km
Terraclimate	Temperature	Average temperature	Raster	4km
TerraClimate - Monthly	Wind Speed	Average wind speed	Raster	4km
TerraClimate- Monthly	Solar Radiation	Amount of solar radiation received	Raster	4km
Distance to Road	Distance to Road	Proximity to the nearest road	Vector	Derived from Roads and Tracks data available with the Department of Resources (map-scale of 1:200K)
Distance to stream	Distance to stream	Proximity to the nearest stream	Vector	Derived from drainage network data available with the Department of Resources (map scale 1: 100 000)
Forest fire history data		Historical forest fire occurrences	Vector	Queensland Parks and Wildlife Service (2m accuracy) (The data was utilised as point locations and incorporated into training and testing datasets.)

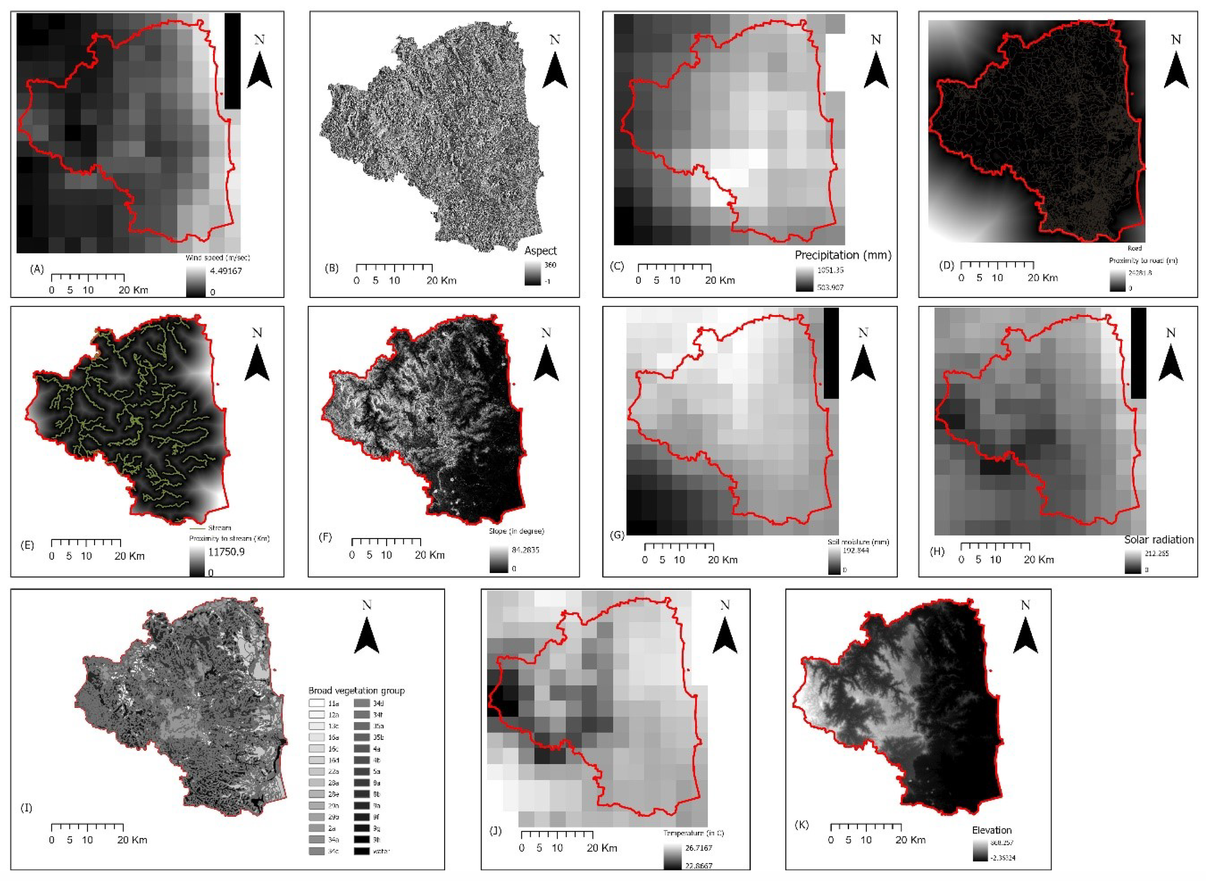
Study Area



Methods:



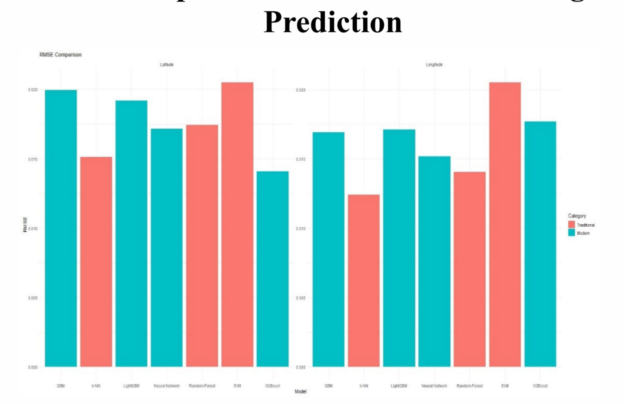
Forest fire indicators



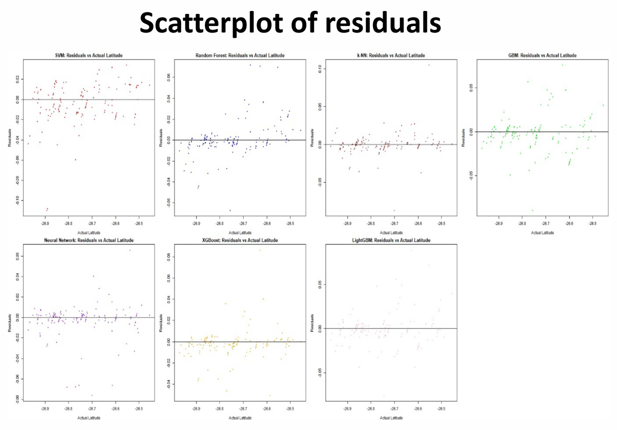
Performance Metrics for Latitude Prediction

Model	Category	RMSE	MAE	R2	Adj_R2
SVM	Traditional	0.020498	0.014181	0.975924	0.975582
Random Forest	Traditional	0.017437	0.010048	0.982578	0.982331
k-NN	Traditional	0.015127	0.00763	0.986887	0.986701
GBM	Modern	0.019928	0.012906	0.977243	0.97692
Neural Network	Modern	0.017144	0.008801	0.983157	0.982918
XGBoost	Modern	0.014075	0.007616	0.988648	0.988487
LightGBM	Modern	0.019187	0.012051	0.978906	0.978607

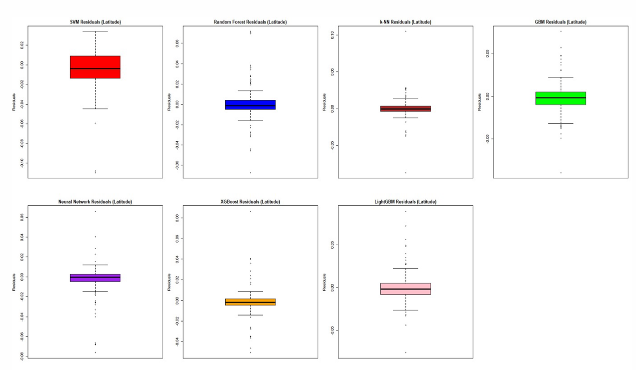
RMSE Comparison for Latitude and Longitude Prediction



Scatterplot of residuals

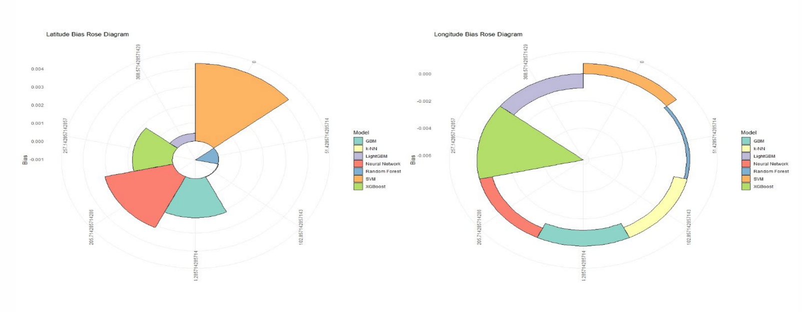


Box plots of residuals

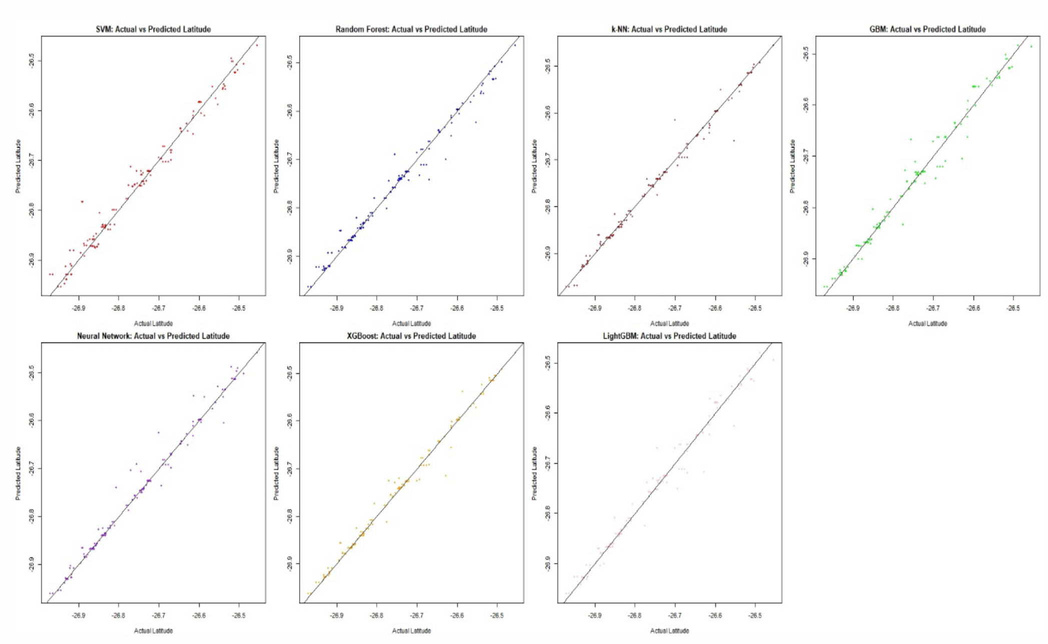


Bias analysis

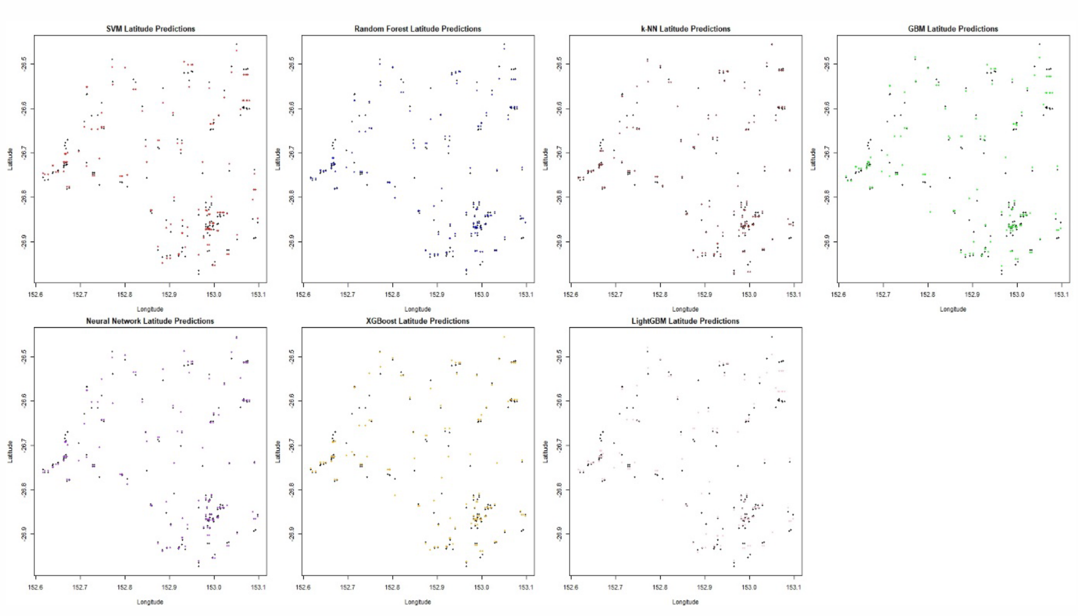
Model	Category	Bias
SVM	Traditional	0.004297
Random Forest	Traditional	-0.00104
k-NN	Traditional	1.11E-05
GBM	Modern	0.002164
Neural Network	Modern	0.003125
XGBoost	Modern	0.001796
LightGBM	Modern	0.000436



Scatter plots of predicted vs actual values



Predicted result of all models



Conclusion

- Modern machine learning models like Neural Networks, XGBoost, and LightGBM generally provide superior predictive accuracy and robustness compared to traditional models.

[1]School of Science Engineering and Technology, University of the Sunshine Coast, Queensland, Australia

Email: Harikesh@research.usc.edu.au

SMARTSAT Cooperative Research Centre
University of the Sunshine Coast Australia
Australian Government Department of Industry, Science, Energy and Resources
AusIndustry Cooperative Research Centres Program