A Comparative Evaluation of Grid-Invariant Deep Learning Surrogate Models for Wildfire Simulation

Matheu Boucher[1], Jidan Zhang[1], Christopher Pain[1], Yueyan Li[1], Aniket Joshi[2], Ben Moseley[1], and Philip Cunningham[3]
[1]Department of Earth Science and Engineering, Imperial College London, London, UK
[2]Department of Civil and Environmental Engineering, Imperial College London, London, UK
[3]Gallagher Re, Natick, MA, USA

EGU26 GENERAL ASSEMBLY

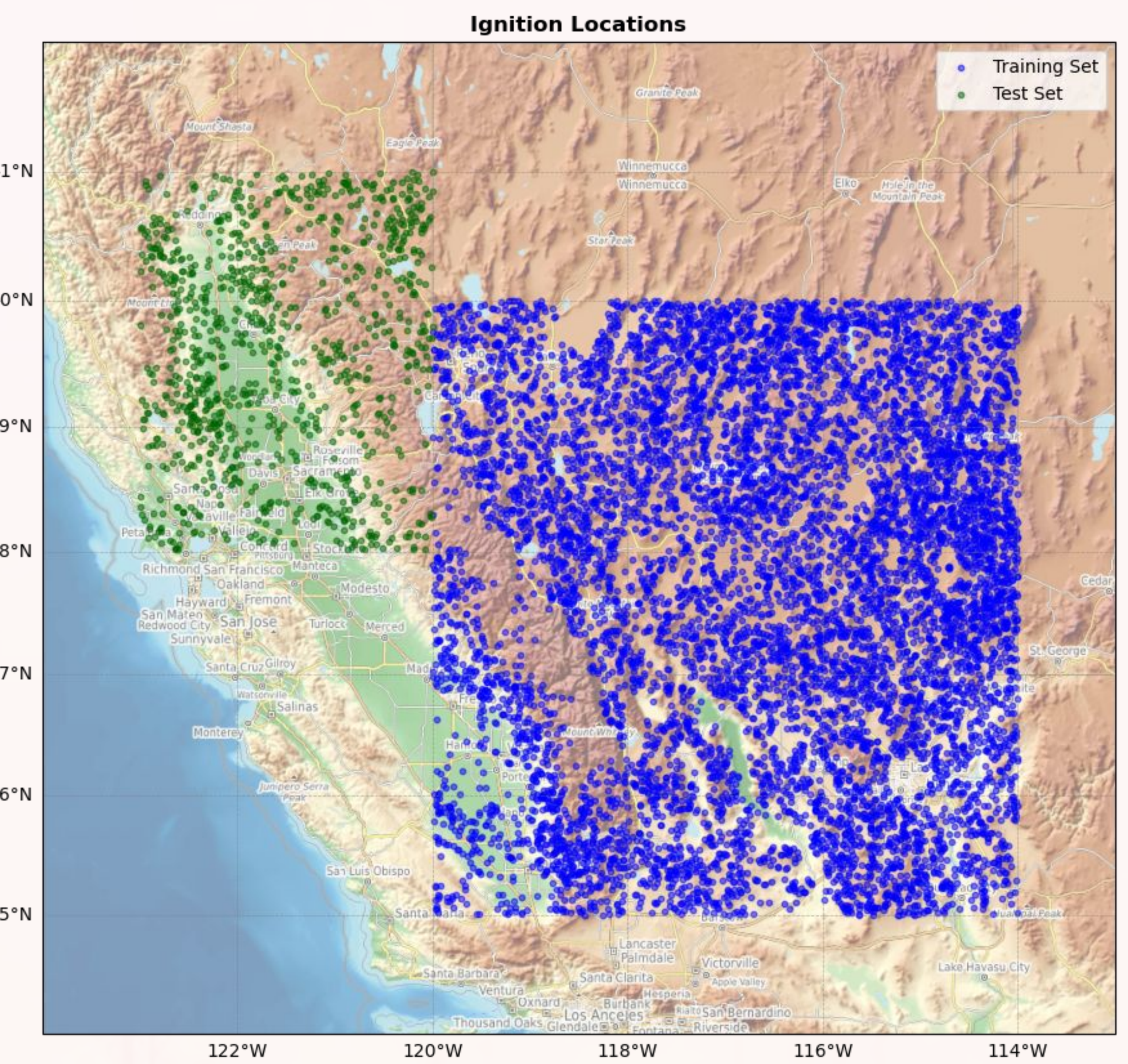
Introduction

- As the climate continues to change, wildfires grow both in size and frequency, impacting urban as well as rural environments with often devastating environmental and human costs.
- Wildfire modelling presents a unique challenge given the reciprocal effects it has on its environment – the fire itself influences the environment in which it propagates. Traditional techniques utilizing numerical, empirical, and semi-empirical methods are often computationally expensive.
- Wildfire modeling thus presents a unique opportunity for the application of deep learning methods that can improve inference time and thus potentially contribute to more robust risk assessments.

Methodology

- The goal of this study was to adapt two grid-invariant architectures that modelled fluid flow in urban environments, RAPIDS and SCALED – developed by Aniket Joshi and Yueyan Li respectively at Imperial College London – to the challenge of wildfire spread prediction.
- 10,000 wildfire simulations were generated using ELMFIRE, a semi-empirical model.
- The models predict time of arrival (s), flame length (ft), and the resulting burn scar at 15-minute intervals.

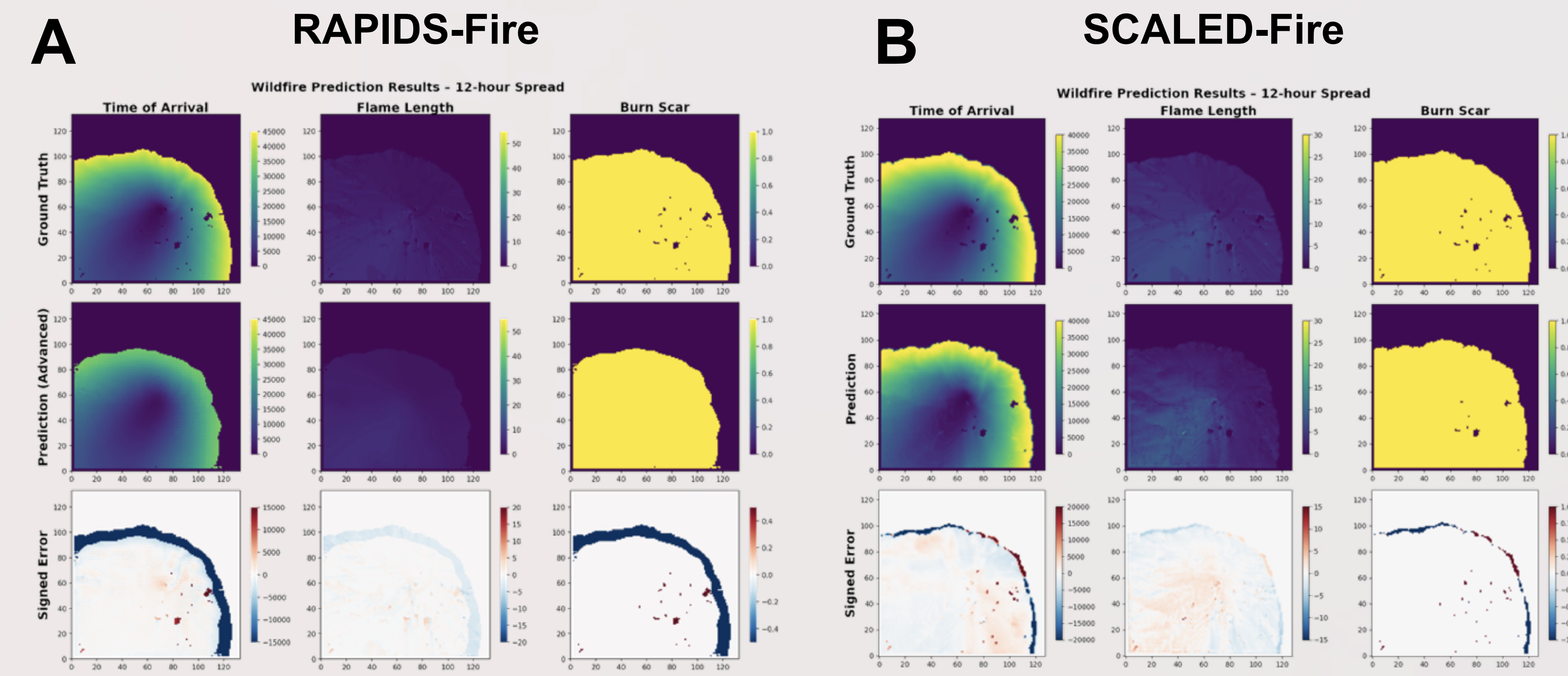
Data and Study Area



- The 10,000 simulation dataset was split into 8,000 training cases, 1,000 validation cases, and 1,000 reserved for testing.
- The test set was kept geographically distinct to assess the models' ability to generalize to unseen fuel models.
- Ten input variables were varied within physically realistic ranges to generate the data.

Fig 1: Distribution of ignition locations for the datasets used for training and evaluating the models. The region was selected both for high prevalence of burnable fuel and historic fire incidences.

Results and Analysis



- Model performance was evaluated on 12-hr autoregressive rollout (48 steps).
- RAPIDS-Fire achieved a 0.79 Jaccard score over the test dataset.
- SCALED-Fire achieved a 0.68 Jaccard score over the test dataset.
- SCALED-Fire was seen to better capture the interior fire structure, more frequently identifying areas of no-burn (e.g. lakes, roads, etc.).

Fig 3: Model results for a single fire case (case 474 from the test set). A) RAPIDS-Fire produces a lower Jaccard score for this case despite performing better on the test set overall. B) SCALED-Fire generates physically accurate fire spread prediction for this case.

Model Architectures

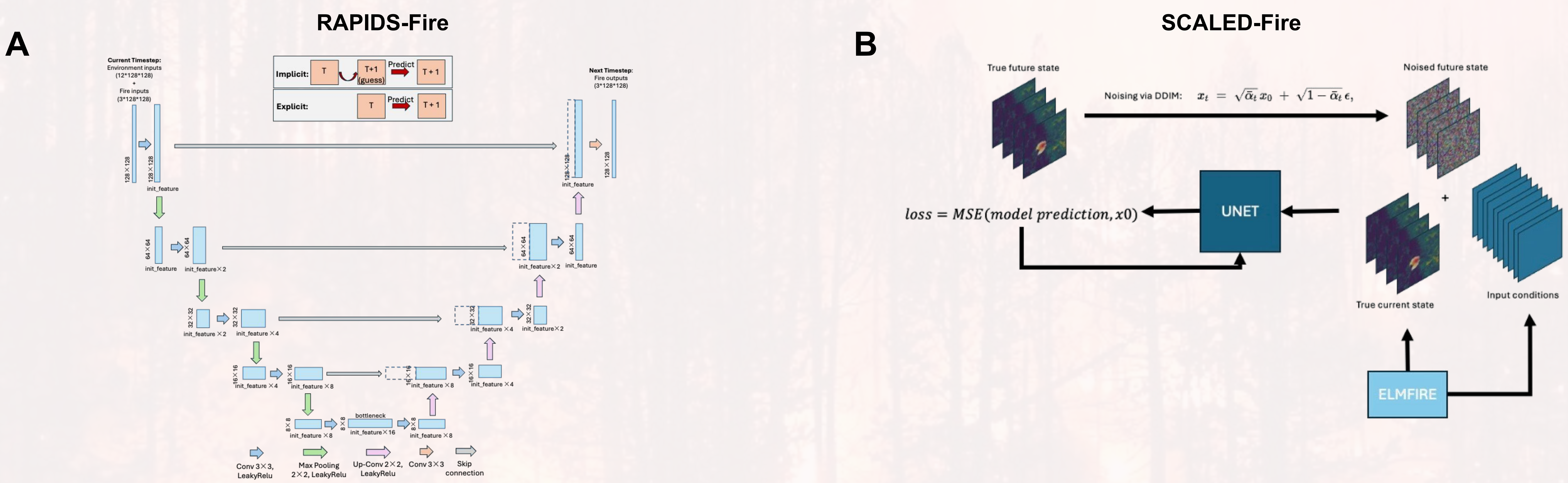


Fig 2: The model architectures used in the study. A) RAPIDS-Fire – A U-Net encoder-decoder with 4 downsampling and 4 upsampling stages. Skip connections at each scale to preserve fine-grained spatial detail. All operations are convolutional; the model produces the next fire state in a single forward pass. B) SCALED-Fire – A diffusion model with a UNet backbone and a DDIM scheduler for inference speedup.

Conclusions

- RAPIDS-Fire outperforms on burn scar extent, however SCALED-Fire performs better at resolving sharp boundaries, identifying unburned areas, and predicting flame length across the entire burned region.
- Both surrogate models are grid invariant owing to the fact that they are fully convolutional models. They are thus able to scale to arbitrary domain sizes.
- Overall, both models stably emulate time of arrival, flame length, and burn scar across diverse, geographically unseen fuel and terrain configurations, using ELMFIRE as a base model.

Future Work

- Future work includes generating larger datasets, and evaluating the models' performance using other base models, including physical solvers.
- To improve model performance, further investigation could be done to integrate physics-informed loss components to the loss functions of the models. While monotonicity was ensured during autoregressive rollout, further physics-informed constraints may improve performance.
- Benchmarking grid-invariant systematically on larger spatial domains.
- Validating across other geographically diverse fire regions.