Proceedings of Machine Learning Research 1–14

# High-Resolution Live Fuel Moisture Content (LFMC) Maps for Wildfire Risk from Multimodal Earth Observation Data

**Patrick Alan Johnson** PATRICKJ@ALLENAI.ORG
*Allen Institute for AI (Ai2), United States of America*

**Gabriel Tseng** GABRIELT@ALLENAI.ORG
*Allen Institute for AI (Ai2), United States of America*

**Yawen Zhang** YAWENZ@ALLENAI.ORG
*Allen Institute for AI (Ai2), United States of America*

**Heather Heward** HHEWARD@UIDAHO.EDU
*University of Idaho, United States of America*

**Virginia Sjahli** VIRGINIAS@ALLENAI.ORG
*Allen Institute for AI (Ai2), United States of America*

**Favyen Bastani** FAVYENB@ALLENAI.ORG
*Allen Institute for AI (Ai2), United States of America*

**Joseph Redmon** JOER@ALLENAI.ORG
*Allen Institute for AI (Ai2), United States of America*

**Patrick Beukema** PATRICKB@ALLENAI.ORG
*Allen Institute for AI (Ai2), United States of America*

## Abstract

Wildfires are increasing in intensity and severity at an alarming rate. Recent advances in AI and publicly available satellite data enable monitoring critical wildfire risk factors globally, at high resolution and low latency. Live Fuel Moisture Content (LFMC) is a critical wildfire risk factor and is valuable for both wildfire research and operational response. However, ground-based LFMC samples are both labor intensive and costly to acquire resulting in sparse and infrequent updates. In this work, we explore the use of a pretrained, highly-multimodal earth-observation model for generating large-scale spatially complete (wall-to-wall) LFMC maps. Our approach achieves significant improvements over previous methods using randomly initialized models (> 20% reduction in RMSE). We provide an automated pipeline that enables rapid generation of these LFMC maps across the United States, and demonstrate its effectiveness in two regions recently impacted by wildfire (Eaton and Palisades).

## 1. Introduction

Live Fuel Moisture Content (LFMC) is a measurement of the amount of water in live vegetation. It is a critical parameter for both wildfire research and operations. LFMC has been shown to have a strong influence on wildfire ignition (Chuvieco et al., 2004), fuel availability (Kelley et al., 2019), and wildfire spread (Rossa, 2017). Live Fuel Moisture Content is calculated using the following formula:

$$\text{LFMC [\%]} = \frac{W_f - W_d}{W_d} \times 100 \quad (1)$$

© P.A. Johnson, G. Tseng, Y. Zhang, H. Heward, V. Sjahli, F. Bastani, J. Redmon & P. Beukema.

Where $W_f$ is the weight of fresh plant material and $W_d$ is the weight of the material after drying (often in an oven). Lower LFMC values indicate higher wildfire risk Rao et al. (2023).

LFMC sampling in the field is a time-consuming and labor-intensive process. Variations in land cover, plant species, and topography can all significantly impact LFMC. Effective and accurate sampling requires expert judgment. Plant material is collected, weighed, dried for one to two days in an oven, and weighed again to calculate LFMC as the percentage change in weight. Depending on site access, equipment, and fuel type, sampling just a single site takes 12 hours – 4 days. Due to these constraints, only a few samples can be gathered at a time, limiting spatial coverage.

Accurate, *spatially complete*, assessments of LFMC have broad applications, including historical analysis of wildfires, planning prescribed burns, and supporting decision-making during active wildfires. Recent work has investigated a combination of machine learning and remote sensing-based techniques to provide wall-to-wall LFMC maps Miller et al. (2023); Rao et al. (2020); Marino et al. (2020). In tandem, the machine learning community has been investigating self-supervised models, with the stated intent of improving the spatial and temporal generalization of machine learning models applied to remote sensing Astruc et al. (2024); Szwarcman et al. (2024); Tseng et al. (2023). These models can be critical tools for LFMC predictions, which often suffer from imbalanced and heterogeneous labels.

In this work, we develop a pipeline which leverages Galileo Tseng et al. (2025) — a pretrained remote sensing model — to generate on-demand LFMC maps. We validate the effectiveness of using a pretrained model, and demonstrate the effectiveness of our pipeline in two study areas.

The contributions of this paper are threefold:

- We develop a pipeline to generate LFMC maps for spatiotemporal areas of interest, allowing users to generate LFMC maps for future treatment planning and in response to historical wildfire events.
- We demonstrate the utility of a highly multimodal *pretrained* geospatial model to make accurate LFMC predictions.
- We apply our pipeline to two real-world case studies without available labels, showing that the model's predictions of LFMC are consistent with expert field observations.

The code is available on GitHub [1].

## 2. Related Works

Previous work on estimating LFMC with machine learning has primarily focused on leveraging multimodal remote sensing data and *fully supervised* approaches to generate coarse resolution maps. For example, Rao et al. (2020) developed a physics-assisted, recurrent neural network (RNN) model to map the LFMC for the western United States at a 250-meter spatial resolution, mainly leveraging Sentinel-1 backscatter and Landsat-8 optical imagery. Miller et al. (2023) proposed a temporal convolutional neural network (tempCNN) model to forecast LFMC at a 500-meter resolution with a three-month lead time, mainly using ground observations, MODIS images, meteorological variables, climate zone classifications, and elevation data. In this study, we investigate a model *pretrained* on multimodal remote sensing data, and produce maps at a far higher spatial resolution (10-meter).

As part of efforts to modernize the U.S. National Fire Danger Rating System, Jolly

1. https://github.com/allenai/lfmc

et al. (2024) introduced a physiologically based LFMC model grounded in plant response mechanisms. Their approach uses a Growing Season Index (GSI) model primarily driven by daily surface weather variables, including minimum temperature, vapor pressure deficit, photoperiod, and rainfall derived from both point-source and gridded weather observations and forecasts. In contrast, we investigate an empirical model that relies on observation data, without explicitly modeling plant physiological processes.

## 3. Materials and methods

### 3.1. The Globe-LFMC 2.0 dataset

In 2024, the second iteration of the global LFMC dataset was published, Globe-LFMC 2.0 (Yebra et al., 2024). This dataset comprises over 280,000 LFMC values derived from samples gathered at more than 2,000 locations across 15 countries collected between 1977 and 2023. It includes data from more than 500 different species or combinations of species.

Filtering and aggregating the Globe-LFMC 2.0 dataset resulted in 41,214 samples across 1,031 sites (described in Section 3.3.1). The median LFMC value was 102.0%, the 1% LFMC value was 51%, and the 99.9% LFMC value was 301.67%. To prevent outliers from having an outsized impact on predictions, all samples were capped and normalized with an LFMC value of 302%, rounded up from the 99.9 percentile.

As shown in Figure 1, when grouped by meteorological season, most samples were taken during summer and the fewest were taken during winter. Spring and autumn had a similar number of samples. This aligns with the typical annual wildfire risk during a calendar year in the United States. Additionally, the spring and autumn months are often used for forest treatments such as prescribed burning due to more favorable weather conditions (Hosten and Stringham, 2020).

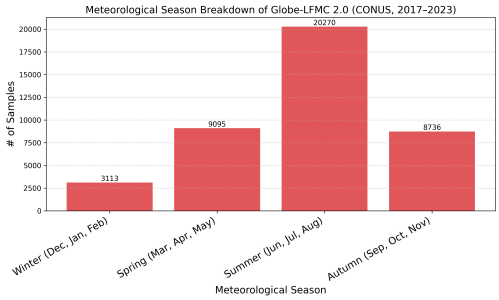



Figure 1: Meteorological season breakdown of the Globe-LFMC 2.0 dataset for CONUS samples 2017-2023.

Also noticeably, as shown in Figure 2, most winter samples were taken in warmer climates such as Southern California, Arizona, and Texas.

### 3.2. The Galileo Pretrained Remote Sensing Model

Recent advancements have created new opportunities for applications that combine remote sensing with deep learning (Vatsavai, 2024). Notably, by pretraining on a large volume of modalities across time and space, geospatial foundation models enable fine-tuned models to be produced efficiently and with lower data labeling requirements than traditional machine learning techniques (Jakubik et al., 2023). A well trained foundation model can be adapted to a diverse set of remote sensing applications, from crop mapping to land cover classification to flood detection (Dionelis et al., 2024).

LFMC estimation is a highly multimodal task; Rao et al. (2020) report the beneficial impact of including a diversity of remote sensing modalities, ranging from directly sensed products such as optical and SAR data to derived products including topography and land cover. We therefore require a *highly multimodal* foundation model, which can process

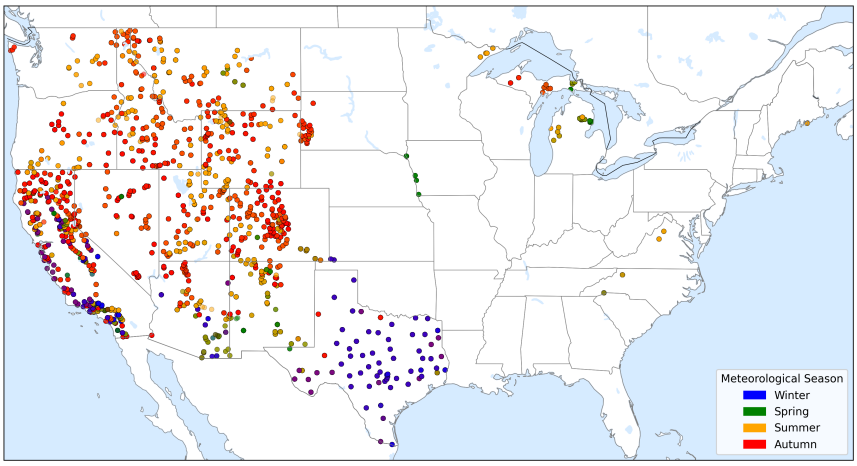



Figure 2: Location breakdown of Globe-LFMC 2.0 dataset for CONUS samples 2017-2023 by meteorological season.

this diversity of inputs. We select the Galileo model Tseng et al. (2025), a model which can ingest 10 directly-sensed and derived remote sensing products across both the spatial and temporal dimensions. We document the inputs we used in more detail below.

Galileo is a vision transformer based model Dosovitskiy et al. (2020) — we use a Galileo-Tiny model, which contains 5.3M parameters. The Galileo-Tiny model is chosen for its balance between a lightweight architecture and strong performance.

#### 3.2.1. Galileo's remote sensing inputs

We use a subset of Galileo's remote sensing inputs when producing our LFMC maps. All remote sensing data was collected from the Google Earth Engine Data Catalog (Gorelick et al., 2017).

- **Multispectral Optical Data**: We use optical imagery from the Sentinel-2 satellite, which include the visible, near-infrared and shortwave infrared bands. In addition, we compute NDVI Tucker (1979) from the near-infrared and red bands. Following Galileo's pretraining, we use the L1C Top of Atmosphere (TOA) product.
- **Synthetic Aperture Radar Data**: We include the dual VV and VH bands from the Sentinel-1 satellite, which can be useful for distinguishing between water, land and vegetation.
- **Night Lights**: We use the night lights from the Visible Infrared Imaging Radiometer Suite (VIIRS) missions, including the average (Day/Night Band) DNB radiance values.
- **Weather Data**: We include precipitation and temperature data from the ERA-5 Land reanalysis product Muñoz Sabater (2019).
- **Climate and Water Balance Data**: We include climate water deficit, soil moisture, and actual evapotranspiration from the TerraClimate dataset Abatzoglou et al. (2018).
- **Topography Data**: We include elevation and slope computed from the SRTM Digital Elevation Model Farr and Kobrick (2000).

- **Location Data** We provide location awareness to the model via latitude and longitude coordinates.

The input datasets used in this analysis span a wide range of spatial and temporal resolutions—for example, spatial resolution ranges from approximately 10 meters per pixel for certain Sentinel-2 bands to tens of kilometers per pixel for ERA5 weather data. Temporal resolution also varies considerably, from a 5-day revisit period for Sentinel-2 to monthly intervals for datasets such as ERA5, TerraClimate, and VIIRS. Some inputs, such as SRTM, are static and do not vary over time. To accommodate these differences, Galileo categorizes inputs based on whether they vary spatially, temporally, or both. Inputs with coarse spatial resolution, such as weather data, are treated as spatially static within the context of a single prediction instance. All temporally varying inputs are aggregated or resampled to a consistent monthly time scale to ensure comparability across datasets.

### 3.3. Full coverage LFMC mapping pipeline

We combine the Globe-LFMC dataset and the pretrained Galileo model to generate a fine-tuned LFMC model. We then apply this fine-tuned model to generate wall-to-wall LFMC maps. We outline this pipeline in Figure 5.

#### 3.3.1. Generating a training dataset from Globe LFMC-2.0

The Globe-LFMC dataset includes samples from around the globe, though the vast majority originate from the continental United States (CONUS). Accordingly, this study focuses solely on samples within CONUS. To ensure satellite data availability for all data sources, only samples from 2017 to 2023 were included.

In the Globe-LFMC 2.0 dataset, it is common to find multiple samples collected at the same location on the same day. Following the guidance of the Globe-LFMC 2.0 paper, such samples were averaged into a single observation.

For each sample in the Globe-LFMC 2.0 dataset, a 1 km × 1 km bounding box was created around its latitude–longitude coordinates. Data within this bounding box was exported from Google Earth Engine for the following products: Sentinel-1, Sentinel-2, VIIRS, SRTM, ERA5-Land, and TerraClimate, and elevation data from the SRTM product.

In total, the dataset included 41,214 ground-based LFMC samples, randomly split into training (≈ 70%), validation (≈ 15%), and test (≈ 15%) sets. Sample elevations ranged from 15 to 3,187 meters and were grouped into 500-meter ranges for analysis, as detailed in Figure 3. The breakdown of all land cover classifications (defined by the ESA WorldCover 2021 product) is shown in Figure 4.

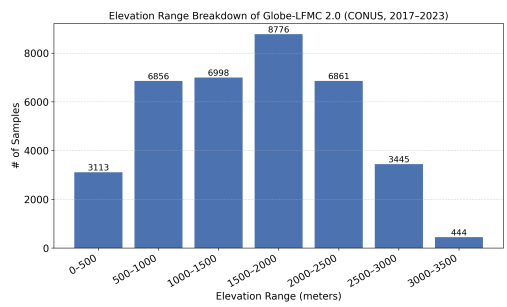



Figure 3: Elevation range breakdown of the Globe-LFMC 2.0 dataset for CONUS samples (2017–2023) in 500 m increments.

#### 3.3.2. Fine-tuning an LFMC Model

We fine-tuned the Galileo-Tiny pretrained on the Globe LFMC-2.0 dataset (combined with the exported remote sensing data, described above in Section 3.3.1) under mean squared error (MSE) loss.

During fine-tuning, we used the ≈70% training set and used the ≈15% validation

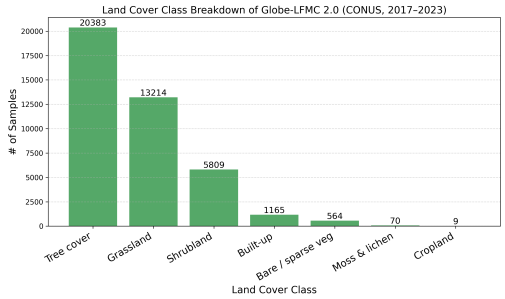



Figure 4: Coverage of the Globe-LFMC 2.0 dataset for CONUS samples 2017-2023 in terms of land cover classes covered by the labels. We use WorldCover Zanaga et al. (2021) to obtain the land cover labels.

set as an early stopping mechanism. Training was conducted for up to 100 epochs, with early stopping applied if no improvement was observed against the validation set for 5 consecutive epochs. This yielded a fine-tuned Galileo$_{LFMC}$ model, which could then be deployed to generate LFMC maps.

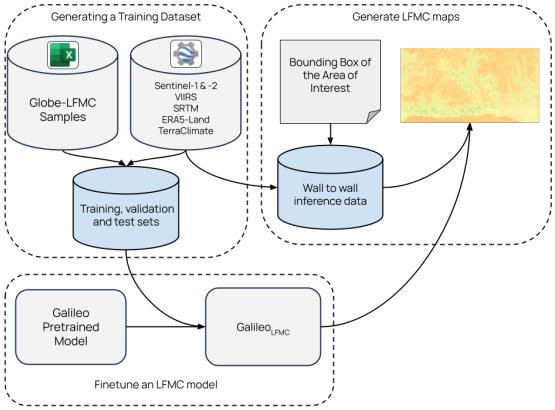



Figure 5: A pipeline to produce LFMC map for a given spatiotemporal window. By leveraging a pretrained remote sensing model and the Globe-LFMC 2.0 dataset, we can efficiently generate maps for new areas and timesteps.

#### 3.3.3. Generating LFMC maps

Generating LFMC maps requires a fine-tuned Galileo$_{LFMC}$ model, and a bounding box describing the desired mapping area. We export the relevant remote sensing products using Google Earth Engine, and run inference over the exported area using Galileo$_{LFMC}$.

#### 3.3.4. Infrastructure

The model was fine-tuned on a single NVIDIA H100 GPU. Fine-tuning took approximately 30-60 minutes depending on the number of epochs it took for the model to reach a steady validation loss.

## 4. Experiments and Results

We report results on the (unseen) $\approx 15\%$ test set, using the random split described in Section 3.3.1. We report root mean squared error (RMSE), mean absolute error (MAE) and the $R^2$ score (coefficient of determination) of our predictions against the test set.

We compare the performance of Galileo$_{LFMC}$ against two baselines: (1) a fine-tuned model initialized with random weights and (2) a model that predicts values using the monthly average of all training samples corresponding to the sample's month. Previous works applying deep learning to LFMC estimation have adopted a fully supervised approach Rao et al. (2020); Miller et al. (2023); Zhu et al. (2021) - our "random weights" baseline reflects this previous approach.

### 4.1. Results

#### 4.1.1. Pretrained weights achieve the best results

Galileo$_{LFMC}$ achieved an overall MAE of 12.58 and an RMSE of 18.91, with an $R^2$ score of 0.72 (Table 1). This is a significant improvement over the randomly initialized

Table 1: Using a pretrained model significantly improves performance compared to a randomly initialized model. In this table, we compare a pretrained Galileo architecture to two baselines: (1) a randomly initialized model and (2) a monthly average, and find a significant reduction in RMSE and an increase in $R^2$ score.

| Category | RMSE | MAE | $R^2$ |
|---|---|---|---|
| Pretrained | **18.91** | **12.58** | **0.72** |
| Random initialized | 23.61 | 16.33 | 0.57 |
| Monthly predictions | 33.66 | 25.38 | 0.11 |

Table 2: Fine-tuned model evaluation across meteorological seasons. Galileo$_{LFMC}$ performs well across all seasons - including the winter season, where it sees fewer training data points.

| Season | RMSE | MAE | $R^2$ |
|---|---|---|---|
| Overall | 18.91 | 12.58 | 0.72 |
| Winter season | 15.31 | 10.74 | 0.77 |
| Spring season | 22.85 | 15.35 | 0.69 |
| Summer season | 19.70 | 13.05 | 0.67 |
| Autumn season | 12.70 | 9.27 | 0.75 |

model ($\approx$ 20% reduction in RMSE), highlighting the effectiveness of pretrained models for LFMC predictions. Our results are comparable to, and in some cases exceed, the performance of existing LFMC models developed without the use of pretrained weights Miller et al. (2023); Rao et al. (2020); Marino et al. (2020).

In addition to the results in Table 1, we plot model performance spatially in Figure 6. To assess the spatial autocorrelation of the residuals, we compute a Moran's I Moran (1950) using K-Nearest Neighbors based spatial weights matrix. The results yielded a Moran's I of 0.057 with a p-value of 0.001, suggesting a statistically significant weak positive spatial autocorrelation in the model errors.

Table 3: Fine-tuned model evaluation across land cover classes. Galileo$_{LFMC}$ performs well across all land cover types.

| Land Cover Class | RMSE | MAE | $R^2$ |
|---|---|---|---|
| Overall | 18.91 | 12.58 | 0.72 |
| Trees | 18.00 | 11.97 | 0.68 |
| Grass | 20.09 | 13.62 | 0.73 |
| Shrub | 19.53 | 12.28 | 0.74 |
| Built-up | 16.79 | 11.78 | 0.77 |
| Bare / Sparse | 20.52 | 15.67 | 0.79 |

Table 4: Fine-tuned model evaluation across elevation bands. $R^2$ values drop at higher elevations, likely due to a sparsity of training labels above 3,000m.

| Category | RMSE | MAE | $R^2$ |
|---|---|---|---|
| Overall | 18.91 | 12.58 | 0.72 |
| Elevation: 0-500m | 18.34 | 11.59 | 0.73 |
| Elevation: 500-1000m | 17.93 | 11.98 | 0.77 |
| Elevation: 1000-1500m | 21.65 | 14.54 | 0.73 |
| Elevation: 1500-2000m | 18.91 | 13.56 | 0.75 |
| Elevation: 2000-2500m | 19.35 | 12.44 | 0.61 |
| Elevation: 2500-3000m | 15.25 | 10.00 | 0.54 |
| Elevation: 3000-3500m | 14.54 | 10.41 | 0.32 |

A weak but statistically significant spatial autocorrelation suggests potential information leakage from the random split, as geographically proximate samples may appear across training, validation, and test sets. Future work could adopt spatial partitioning methods to better assess model generalization.

When analyzed across meteorological seasons (Table 2), the model's lowest RMSE was for autumn (12.70) and its highest RMSE was for spring (22.85). It did not demonstrate lower performance nor explainability ($R^2$ score) in the winter, despite lower representation in the dataset. This suggests that Galileo's pretraining—which exposes it to data from many different timesteps—is helpful in ensuring Galileo$_{LFMC}$ can generalize temporally.

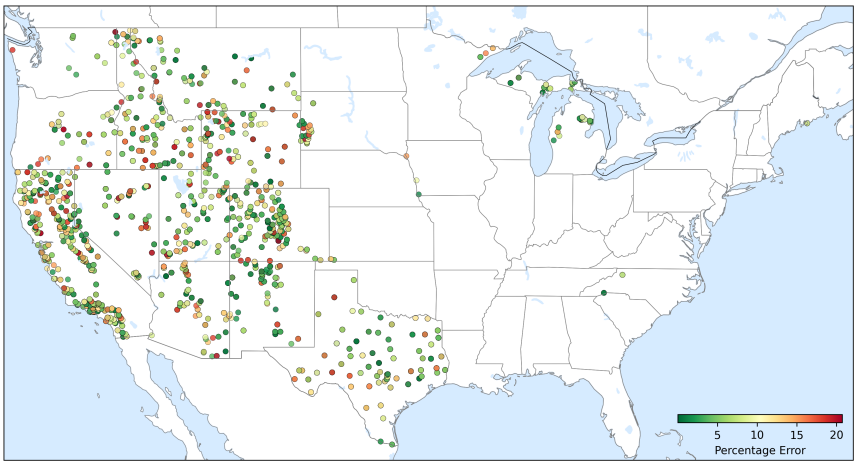



Figure 6: Percentage error for the test dataset. Galileo$_{LFMC}$ performs well across the entire spatial distribution of the test set. Moran's I is 0.057 with a p-value of 0.001, indicating a statistically significant weak positive autocorrelation in model errors.

The model's performance across the various land cover classes was consistent with its overall accuracy (Table 3). RMSE values ranged from 16.79 for built-up areas to 20.52 for sparse vegetation, with a median value of 18.00. Notably, despite their lower representation, built-up and sparse areas did not exhibit significantly different errors compared to more prevalent classes such as tree-covered areas, grasslands, and shrublands. Land cover types with very limited representation in the filtered Globe-LFMC 2.0 dataset—specifically moss/lichen and cropland—were not covered by the test set.

The model maintained an $R^2$ score above 0.7 for elevation ranges below 2,000 meters (Table 4). However, performance declined substantially at higher elevations, with $R^2$ values dropping to 0.54 for the 2,500–3,000 meter range and 0.32 for the 3,000–3,500 meter range. This decrease is likely due to the limited representation of high-elevation areas in the Globe-LFMC 2.0 dataset—in particular, there were only 444 samples above 3,000 meters. RMSE values across elevation bands ranged from 14.54 (3000–3500 meters) to 21.65 (1000–1500 meters). Among the bands with $R^2 > 0.7$—those under 2,000 meters—the lowest RMSE was 17.93. The overall median RMSE was 18.34.

#### 4.1.2. Model performance is not sensitive to input shape

In addition to being highly multimodal, Galileo can ingest a variety of input shapes (i.e. varying spatial areas and temporal ranges). We investigate the effect of passing different input shapes as inputs (compared to our default spatial area of $[32 \times 32]$ pixels and 12 timesteps) in Table 5. While more spatial and temporal context helps, we find that the model remains relatively stable as the spatial area and temporal range are reduced.

#### 4.1.3. Pretraining makes the model more robust to missing inputs

We use highly multimodal remote sensing inputs when fine-tuning the Galileo model and making LFMC maps (described in Section 3.2.1). To understand the contribution of each of these modalities, we fine-tune the pre-

Table 5: We fine-tune the pretrained Galileo model with a variety of input shapes (i.e. varying [height× width] and number of input timesteps). While more spatial and temporal context is beneficial, the fine-tuned model's performance stays relatively stable as these parameters are changed.

| H, W | T | RMSE | MAE | $R^2$ |
|---|---|---|---|---|
| 32 | 12 | 18.91 | 12.58 | 0.72 |
| 32 | 6 | 18.93 | 12.68 | 0.72 |
| 32 | 3 | 19.45 | 13.10 | 0.70 |
| 16 | 12 | 19.36 | 12.67 | 0.71 |
| 8 | 12 | 19.71 | 13.20 | 0.70 |
| 1 | 12 | 20.25 | 13.46 | 0.68 |

trained and randomly initialized models with one input removed in Table 6. We find that for the randomly initialized model, removing TerraClimate yields a significant degradation in performance (8% reduction in RMSE). On the other hand, the pretrained model's performance remains relatively stable even as inputs are removed, which suggests Galileo's self-supervised pretraining increases its robustness to missing inputs at fine-tuning time. This allows the pretrained model to be leveraged even if certain remote sensing products are missing (e.g. due to clouds).

### 4.2. Case study: 2025 Palisades and Eaton Wildfires

In January 2025, a series of devastating wildfires burned across the Los Angeles metropolitan area. These fires were exacerbated by prolonged drought, low humidity, unusually strong northeasterly (Santa Ana) winds, and a buildup of vegetation following two consecutive springs of above-average rainfall (Giuseppe et al., 2025).

The two largest and most destructive fires—by both area burned and structures lost—were the Palisades and Eaton wildfires, both located in Los Angeles County. The Palisades

Table 6: We fine-tune the pretrained and randomly initialized Galileo models while removing one of the at fine-tuning time. Pretraining makes the model more robust to missing inputs; the randomly initialized model experiences significantly larger changes in performance depending on the removed inputs (e.g. a 14 % degradation in $R^2$ score when the TerraClimate input is removed).

| | Pretrained | | | Random | | |
|---|---|---|---|---|---|---|
| W/O IN | RMSE | MAE | $R^2$ | RMSE | MAE | $R^2$ |
| None | **18.91** | 12.58 | 0.72 | 23.61 | 16.33 | 0.57 |
| S2 | 19.51 | 13.10 | 0.70 | 23.46 | 16.45 | 0.57 |
| S1 | **18.82** | 13.10 | 0.72 | 23.84 | 16.57 | 0.56 |
| ERA5 | 19.27 | 13.09 | 0.71 | **22.42** | 15.90 | 0.61 |
| TC | 19.51 | 13.02 | 0.70 | 25.57 | 17.59 | 0.49 |
| SRTM | 19.61 | 13.34 | 0.70 | **22.02** | 15.34 | 0.62 |
| loc. | 20.08 | 13.91 | 0.69 | 23.80 | 16.54 | 0.56 |

Fire originated in the Santa Monica Mountains and severely impacted Pacific Palisades, Topanga, and Malibu. The Eaton Fire started in the San Gabriel Mountains and destroyed foothill communities, particularly Altadena. Both fires caused unprecedented levels of damage.

Perimeter data for both fires was retrieved on January 13, 2025, from the WFIGS Current Interagency Fire Perimeters database (National Interagency Fire Center, 2025). At the time, both fires were active and largely uncontained. Bounding boxes were constructed for each fire, and data was exported from Google Earth Engine covering the period from January 1, 2020, to December 31, 2024.

The fine-tuned model was used to generate monthly LFMC predictions for the affected areas, producing outputs at a 10-meter spatial resolution. We show a seasonal example of these predictions for the Eaton Fire in Figure 7.

The model revealed a consistent seasonal pattern: average LFMC levels rose through winter and peaked in late spring before de-

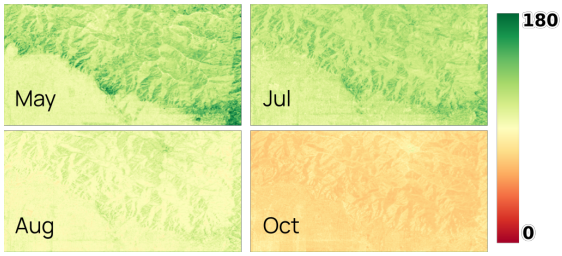
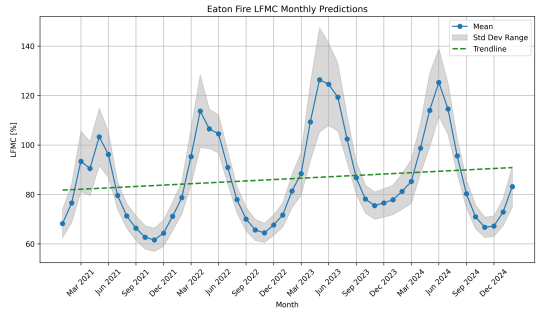
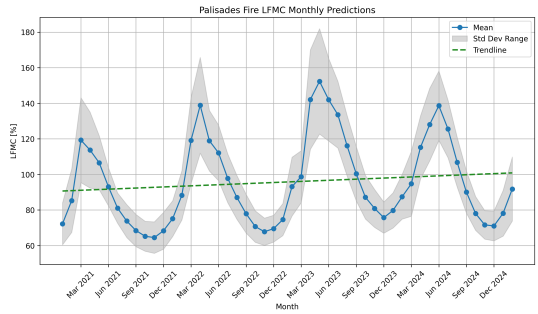



Figure 7: LFMC predictions for the San Gabriel Mountains, site of the Eaton Fire in 2025. These maps show 4 months in 2024 and demonstrate our pipeline's ability to make spatially complete maps that capture seasonal patterns and geographic variability in LFMC values. LFMC values range from 0% (red) to 180% (green).




**(a)** Eaton Fire area




**(b)** Palisades Fire area

Figure 8: Monthly LFMC average predictions for two wildfire areas from 2021–2024: (a) Eaton and (b) Palisades. These averages are computed over wall-to-wall maps like the ones shown in Figure 7 and Figure 9.

clining into late autumn. This is shown in the average LFMC prediction for all pixels in the bounding boxes (see Figure 8). Additionally, the model predicted higher LFMC values in 2023 and 2024 compared to 2021 and 2022 (Figure 9). While the link between LFMC and fuel accumulation is not fully understood, this pattern may reflect increased vegetation growth following wetter years. Such growth has been linked to greater fuel loads, a contributing factor in wildfire severity (Giuseppe et al., 2025).

## 5. Discussion

Live Fuel Moisture Content (LFMC) is a key environmental variable for assessing wildfire risk, as it directly affects both the likelihood of ignition and the potential for fire spread. However, current live fuel moisture sampling is restricted by field and processing constraints, as samples must be collected from easily accessible areas and quickly transported to laboratory facilities for drying. These limitations reduce the spatial and temporal resolution of traditional measurements.

An accurate and reliable LFMC prediction model may offer significant value to land managers, wildland firefighters, and the communities potentially threatened by wildfires. Improved ability to analyze past wildfire events enhances our understanding of the conditions that contribute to large, destructive fires. This enables more effective anticipation and mitigation strategies. Furthermore, such insights support the evaluation of conditions under which prescribed fire has been successfully implemented, thereby promoting its strategic use as a tool for reducing wildfire risk.

This study demonstrates that a practical LFMC model can be built using existing state-of-the-art geospatial foundation models, achieving reasonably accurate predictions in many cases. The model offers high spa-

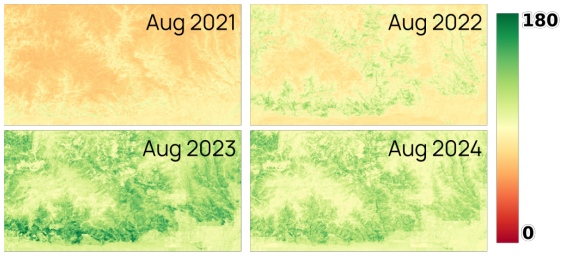



Figure 9: LFMC predictions for the Santa Monica Mountains and nearby coastal communities of Pacific Palisades, Topanga, and Malibu, site of the Palisades fire in 2025. These maps show the month of August for 2021, 2022, 2023, and 2024, demonstrating our pipeline's ability to produce spatially complete maps that capture annual variability in LFMC values driven by differences in seasonal cycles. LFMC values range from 0% (red) to 180% (green).

tial resolution (10 meters), whereas previous studies provided only 500-meter resolution (Miller et al., 2023). A case study of 2025 wildfire events in Los Angeles supported the model's reliability, showing results consistent with observed surface conditions.

Nonetheless, the model has several limitations that warrant further research and development. While the Galileo-Tiny foundation model was trained on global data, the LFMC model was fine-tuned exclusively on data from the continental United States, primarily in the western region. A more extensive global sample could be acquired and used for training. Furthermore, the model has only been applied retrospectively and has not yet been tested for forecasting LFMC. Finally, it currently produces only monthly averages, whereas more frequent weekly and daily predictions—including on-demand generation for current conditions—would be highly valuable, especially for real-time wildfire management.

## 6. Acknowledgments

We thank John J. Battles (UC Berkeley), Rob York (UC Berkeley), Peter Stine (UC Davis; California Wildfire and Forest Resilience Task Force), Jonathan McDuffey (USFS), and Anand Padmanabhan (University of Illinois Urbana-Champaign) for their valuable feedback on this project.

## References

John T. Abatzoglou, Solomon Z. Dobrowski, Sean A. Parks, and Katherine C. Hegewisch. Terraclimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific Data*, 5, 2018.

Guillaume Astruc, Nicolas Gonthier, Clement Mallet, and Loic Landrieu. Anysat: An earth observation model for any resolutions, scales, and modalities. *arXiv preprint arXiv:2412.14123*, 2024.

Emilio Chuvieco, Inmaculada Aguado, and Alexandros Dimitrakopoulos. Conversion of fuel moisture content values to ignition potential for integrated fire danger assessment. *Canadian Journal of Forest Research*, 34:2284–2293, 2004.

Nikolaos Dionelis, Casper Fibaek, Luke Camilleri, Andreas Luyts, Jente Bosmans, and Bertrand Le Saux. Evaluating and benchmarking foundation models for earth observation and geospatial ai. *ArXiv*, abs/2406.18295, 2024.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

Tom G. Farr and Michael Kobrick. Shuttle radar topography mission produces a wealth of data. *Eos, Transactions American Geophysical Union*, 81:583–585, 2000.

Francesca Di Giuseppe, Joe McNorton, Anna Lombardi, and Fredrik Wetterhall. Global data-driven prediction of fire activity. *Nature Communications*, 16, 2025.

Noel Gorelick, Matt Hancher, Michael Dixon, Simon Ilyushchenko, David Thau, and Rebecca Moore. Google earth engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, 202:18–27, 2017.

Peter E. Hosten and Trent K. Stringham. Prescribed fire basics: Fire weather. https://extension.oregonstate.edu/catalog/pub/em-9385-prescribed-fire-basics-fire-weather, 2020. EM 9385.

Johannes Jakubik, Sujit Roy, Christopher Phillips, Paolo Fraccaro, Denys Godwin, Bianca Zadrozny, Daniel Szwarcman, Carlos Gomes, Gabby Nyirjesy, Blair Edwards, Daiki Kimura, Naomi Simumba, Linsong Chu, S. Karthik Mukkavilli, Devyani Lambhate, Kamal Das, Ranjini Bangalore, Dario Oliveira, Michal Muszynski, Kumar Ankur, Muthukumaran Ramasubramanian, Iksha Gurung, Sam Khallaghi, Hanxi Li, Michael Cecil, Maryam Ahmadi, Fatemeh Kordi, Hamed Alemohammad, Manil Maskey, Raghu Kiran Ganti, Kommy Weldemariam, and Rahul Ramachandran. Foundation models for generalist geospatial artificial intelligence. *ArXiv*, abs/2310.18660, 2023.

W. Matt Jolly, Patrick H. Freeborn, Larry S. Bradshaw, Jon Wallace, and Stuart Brittain. Modernizing the us national fire danger rating system (version 4): Simplified fuel models and improved live and dead fuel moisture calculations. *Environ. Model. Softw.*, 181:106181, 2024.

Douglas I. Kelley, Ioannis Bistinas, Rhys Whitley, Chantelle A. Burton, Toby R. Marthews, and Ning Dong. How contemporary bioclimatic and human controls change global fire regimes. *Nature Climate Change*, pages 1–7, 2019.

Eva Marino, Marta Yebra, Mariluz Guillén-Climent, Nur Algeet, José Luis Tomé, Javier Madrigal, Mercedes Guijarro, and Carmen Hernando. Investigating live fuel moisture content estimation in fire-prone shrubland from remote sensing using empirical modelling and rtm simulations. *Remote Sensing*, 12(14):2251, 2020.

Lynn Miller, Liujun Zhu, Marta Yebra, Christoph Rüdiger, and Geoffrey I Webb. Projecting live fuel moisture content via deep learning. *International Journal of Wildland Fire*, 32(5):709–727, 2023.

Percy Moran. Notes on continuous stochastic phenomena. *Biometrika*, 37 1-2:17–23, 1950.

J. Muñoz Sabater. Era5-land monthly averaged data from 1950 to present. https://cds.climate.copernicus.eu/cdsapp#!/dataset/reanalysis-era5-land-monthly-means, 2019. Accessed 2025.

National Interagency Fire Center. Wfigs current interagency fire perimeters. https://data-nifc.opendata.arcgis.com/datasets/nifc::wfigs-current-interagency-fire-perimeters/about, 2025. Accessed 13 Jan 2025.

Krishna Rao, A Park Williams, Jacqueline Fortin Flefil, and Alexandra G Konings. Sar-enhanced mapping of live fuel moisture content. *Remote Sensing of Environment*, 245:111797, 2020.

Krishna Rao, A Park Williams, Noah S Diffenbaugh, Marta Yebra, Colleen Bryant, and Alexandra G Konings. Dry live fuels increase the likelihood of lightning-caused fires. *Geophysical Research Letters*, 50(15): e2022GL100975, 2023.

Carlos G. Rossa. The effect of fuel moisture content on the spread rate of forest fires in the absence of wind or slope. *International Journal of Wildland Fire*, 26:24–31, 2017.

Daniela Szwarcman, Sujit Roy, Paolo Fraccaro, Thorsteinn Elí Gíslason, Benedikt Blumenstiel, Rinki Ghosal, Pedro Henrique de Oliveira, João Lucas de Sousa Almeida, Rocco Sedona, Yanghui Kang, Srija Chakraborty, Sizhe Wang, Ankur Kumar, Myscon Truong, Denys Godwin, Hyunho Lee, Chia-Yu Hsu, Ata Akbari Asanjan, Besart Mujeci, Trevor Keenan, Paulo Arévalo, Wenwen Li, Hamed Alemohammad, Pontus Olofsson, Christopher Hain, Robert Kennedy, Bianca Zadrozny, Gabriele Cavallaro, Campbell Watson, Manil Maskey, Rahul Ramachandran, and Juan Bernabé-Moreno. Prithvi-eo-2.0: A versatile multi-temporal foundation model for earth observation applications. *arXiv preprint arXiv:2412.02732*, 2024.

Gabriel Tseng, Ruben Cartuyvels, Ivan Zvonkov, Mirali Purohit, David Rolnick, and Hannah Kerner. Lightweight, pre-trained transformers for remote sensing timeseries. *arXiv preprint arXiv:2304.14065*, 2023.

Gabriel Tseng, Anthony Fuller, Marlena Reil, Henry Herzog, Patrick Beukema, Favyen Bastani, James R. Green, Evan Shelhamer, Hannah Kerner, and David Rolnick. Galileo: Learning global and local features in pretrained remote sensing models. *ArXiv*, abs/2502.09356, 2025.

Compton J Tucker. Red and photographic infrared linear combinations for monitoring vegetation. *Remote sensing of Environment*, 8(2):127–150, 1979.

Ranga Raju Vatsavai. Geospatial foundation models: Recent advances and applications. *Proceedings of the 12th ACM SIGSPATIAL International Workshop on Analytics for Big Geospatial Data*, 2024.

Marta Yebra, Gianluca Scortechini, Karine R. M. Adeline, Nursema Aktepe, Turkia AbdAlkarem Almoustafa, Avi Bar-Massada, María Eugenia Beget, Matthias M. Boer, Ross Bradstock, Tegan Brown, Francesc Xavier Castro, Rui Chen, Emilio Chuvieco, Mark F. Danson, Cihan Ünal Değirmenci, Ruth Delgado-Dávila, Philip E. Dennison, Carlos M. Di Bella, Oriol Domenech, Jean-Baptiste Féret, Greg Forsyth, Eva Gabriel, Zisis Gagkas, Fatma Gharbi, Elena Granda, Anne Griebel, Binbin He, Matt Jolly, Ivan Kotzur, Tineke Kraaij, Agnes Kristina, Pınar Kütküt, Jean-Marc Limousin, M Pilar Martín, Antonio T Monteiro, Marco Morais, Bruno Moreira, Florent Mouillot, Samukelisiwe T Msweli, Rachael H. Nolan, Grazia Pellizzaro, Yi Qi, Xingwen Quan, Víctor Resco de Dios, Dar A. Roberts, Çağatay Tavşanoğlu, Andy F S Taylor, Jackson Taylor, İrem Tüfekcioğlu, Andrea Ventura, and Nicolas Younes Cardenas. Globe-lfmc 2.0, an enhanced and updated dataset for live fuel moisture content research. *Scientific Data*, 11, 2024.

Daniele Zanaga, Ruben Van De Kerchove, Wanda de Keersmaecker, Niels Souverijns, Carsten Brockmann, Ralf Quast, Jan Wevers, Alexandra Claudia Grosu, Audrey Paccini, Sylvain Vergnaud, Oliver Cartus, Maurizio Santoro, Steffen Fritz, Ivelina Georgieva, Myroslava Lesiv, Sarah Carter, Martin Herold, Linlin Li, Nandin-Erdene

Tsendbazar, Fabrizio Ramoino, and Olivier Arino. Esa worldcover 10 m 2020 v100, 2021.

Liujun Zhu, Geoffrey I Webb, Marta Yebra, Gianluca Scortechini, Lynn Miller, and François Petitjean. Live fuel moisture content estimation from modis: A deep learning approach. *ISPRS Journal of Photogrammetry and Remote Sensing*, 179:81–91, 2021.