Firecast Zigzag Convolutional Network for Wildfire Prediction

Yuzhou Chen[1], Joel Chacon Castillo[2], Huikyo Lee[3], Yulia R. Gel[4,5]

[1]Department of Statistics, University of California, Riverside
[2]Department of Computer Science, Center for Research in Mathematics
[3]Jet Propulsion Laboratory, California Institute of Technology
[4]Department of Statistics, Virginia Tech
[5]National Science Foundation

2024 IEEE International Conference on Big Data (BigData) | 979-8-3503-6248-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/BigData62323.2024.10825303

***Abstract*—Each year wildfires result in billions of dollars in property damage. Being one of the major natural hazards, wildfires nowadays are also a global affair whose negative impact is particularly devastating in developing countries. As wildfires are expected to become more frequent and severe, more accurate models to predict wildfires are vital to mitigating risks and developing more informed decision-making. Artificial intelligence (AI) has a potential to enhance wildfire risk analytics on multiple fronts. For example, deep learning (DL) has been successfully used to classify active fires, burned scars, smoke plumes and to track the spread of active wildfires. Since wildfire spread tends to exhibit highly complex spatio-temporal dependencies which often cannot be accurately described with conventional Euclidean-based approaches, we postulate that the tools of topological and geometric deep learning, specifically designed for non-Euclidean objects such as manifolds and graphs, may offer a more competitive solution. We validate the proposed methodology to predict wildfire occurrences in Greece and several regions of Africa. Our results indicate that the Firecast Zigzag Convolutional Network (F-ZCN) outperforms the current baseline methods for wildfire prediction and opens a path for more accurate wildfire risk analytics, even in scenarios of limited and noisy data records.**

***Index Terms*—Time-Series Analysis, Topological Data Analysis, Geometric Deep Learning, Wildfire Risk Analytics**

I. Introduction

Modeling wildfire occurrences is a spatio-temporal complex process which includes non-linear interrelated fire drivers factors (e.g., weather, vegetation, and soil moisture) along with human drivers, further exacerbated by current climate change [1]. Although significant advances in wildfire modeling have been made, there still exist major limitations, such as computational power, quality, and availability of the data [2]. Furthermore, despite recent breakthroughs in machine learning (ML) in Earth science, the simultaneous and complementary use of spatial structures and temporal evolution from space and suborbital observations is still in its infancy. One of the key obstructing challenges here is the inability of the current wildfire models to efficiently characterize nonstationary spatio-temporal structures and to detect anomalies arising in the time-evolving environment of the modeling steps. Furthermore, one of the main challenges in providing decision-makers with comprehensive, quantitative, and consumable information on wildfire forecasts, especially in developing countries, is the limited spatial coverage of ground-based stations to measure the key predictors. As such, the key to reliable forecasting of wildfires is novel DL models that can fully take advantage of satellite observations and the blend of observations and numerical simulations, such as the European Centre for Medium-Range Weather Forecasts Reanalysis v5 (ERA5; [3]).

We propose to address these challenges by harnessing the power of geometric deep learning (GDL) enhanced by the tools of topological data analysis (TDA), particularly, zigzag persistence. The benefits of this approach are multi-fold. First, by using a graph representation of wildfire spread, it allows us to capture the complex spatio-temporal dependencies that the traditional Euclidean-based methods cannot. Second, by integrating zigzag persistence, we better describe latent dynamic higher interactions among various wildfire driving factors that are the key behind the wildfire spread mechanism. Third, topological tools that are inherently more robust to noisy data, enable us to better address the problem of limited and irregular data records. As a result, we develop a more systematic time-aware learning of wildfire spread.

The key novelties of this paper are summarized as follows:

- We introduce the concepts of geometric deep learning to wildfire analytics, particularly, spatio-temporal prediction of wildfire occurrence.
- We propose a time-aware topological layer which learns patterns of persistent features across multiple scales of the spatio-temporal fire drivers, using the notion of zigzag persistence.
- We illustrate the proposed Firecast Zigzag Convolutional Network (F-ZCN), in application to wildfire prediction in Greece and Africa. Our experiments indicate that F-ZCN leads to a more accurate wildfire prediction and shows promise in successfully detecting hidden spatiotemporal relationships among multi-source fire drivers.

II. Related Work

Fire Weather Index. The most widespread tool to identify periods of higher wildfire risk is the Fire Weather Index (FWI), which relies on atmospheric factors, fuel moisture, and physics-based modelling to provide consistent ratings of fire susceptibility [4]. FWI is successful as an indicator of fire

danger since this index correlates with burned areas across non-arid global ecoregions and wetter climates. Despite FWI success, there are regions, e.g., the central African rainforest, in which anthropogenic factors, such as intentional ignitions and landscape fragmentation, produce weak climate-fire relationships making FWI less effective as wildfire predictor [5], [6]. Furthermore, there is societal and economic pressure to provide as more reliable information for fire management and future fire suppression decisions, in both, developed and developing countries [7]. As prediction of widespread events is a complex-multifaceted problem, we aim to further enrich the pool of tools by showing what value modern persistent homology approaches can bring to wildfire prediction in developed and developing countries.

Wildfire Prediction via DL. Application of DL within wildfire science is a fairly underexplored area, with newer models emerging during the last few years to tackle different aspects of wildfires [8]–[10]. In the context of wildfire prediction, some of the first efforts include the use of spatial DL for forest fire risk prediction in Australia [11], and applications of Convolutional Neural Networks (CNNs) to predict which areas surrounding a burning wildfire have a higher risk of near-future wildfire spread in Colorado-US [12]. In [13], a variety of DL models are introduced to capture the spatio-temporal features, from which the Convolutional Long-Short Term Memory (ConvLSTM) achieved the best metrics against traditional ML methods. In turn, the utility of Graph Neural Networks (GNNs) and other GDL models for unstructured wildfire data still remains noticeably less unexplored [10], [14]. Finally, WildfireDB [15] contains over 17 million data that capture how fires spread in the continental USA in the last decade and shows benchmark results of ML models for wildfire spread prediction. To the best of our knowledge, there is no prior studies using GNNs enhanced by TDA for wildfire prediction.

Topological Data Analysis within Wildfire Research. Topological Data Analysis (TDA) is a modern approach to characterize the shapes and patterns of data based on its topology, mainly by means of persistent homology (PH) [16]. Although TDA is being actively used and developed across multiple research areas, there are almost virtually no applications in the context of wildfire research. [17] uses the Mapper algorithm and PH to detect and classify non-linear patterns in weather data during active and inactive fire seasons. In turn, [18] considers PH to associate wildfire sizes, i.e., burned areas, with environmental conditions, thus, showing new ways to classify wildfire events. These results further highlight the idea that ML models containing TDA-based signatures are likely to improve situational awareness of wildfire events. To the best of our knowledge there are not any previous studies that extract persistent spatio-temporal features to perform geographical wildfire prediction.

III. Preliminaries

Zigzag Persistence. Zigzag persistence (ZP) generalizes the theory of traditional PH [19] to a case of inclusions going into multiple directions [20]–[22]. In this paper ZP is used to track relevant topological features of the data over time, by considering bidirectional maps (i.e. forward and backward maps) over a time-indexed data.

Similarly to PH, information extracted from ZP can be represented in the form of a persistence image [23]; this representation is also known as the Zigzag Persistence Image (ZPI) [24]. Alternatively, the ZP summary can be represented in a form of zigzag curve [25].

Graphs and Image Representations. Suppose we have a d-dimensional spatio-temporal dataset $\mathbf{Y} = \{y^1, ..., y^d\} \in \mathbb{R}^d$. We can construct a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$ from $\mathbf{Y}$, by viewing each point as a node and computing the edge weights through a certain similarity function. Here $\mathbf{V}$ denotes a set of nodes, and $\mathbf{E}$ denotes a set of edges, weighted by a similarity function. If we observe the data $\mathbf{Y}$ over T steps, it will then result in a sequence of the corresponding time-evolving graphs $\{\mathbf{G}_t\}_1^T$. In addition, we associate each graph $\mathbf{G}_t$ with a node-features matrix $\mathbf{X}_t \in \mathbb{R}^{N\times d}$, where N is the dimension of node attributes. In this context the node-feature matrix can represent various spatio-temporal information about locations (i.e., nodes). Furthermore, we can associate $\mathbf{X}_t$ with an image defined as $\mathbf{I}_t \in \mathbb{R}^{Cin\times W\times H}$, where Cin is the number of input channels, W is the width, and H height of the image. In this image representation the number of channels corresponds to the feature dimension ($Cin = d$) and the number of pixels in the image corresponds to the number of nodes ($W \times H = N$). The type of connectivity given an image can be modeled in different ways, for example, using connections of each pixel (i.e. node) with its eight surrounded neighbors.

Graph Convolutions. A graph convolution is an operation to extract feature information from each node given its structure [26]. It can be interpreted as smoothing the information of each node by aggregating and transforming the information of its neighborhood. A graph convolution layer can be defined as follows [27]:

$$\mathbf{Z} = \sum_{k=0}^{K} \mathbf{P}_1^k \mathbf{X}\mathbf{W}_1 + \hat{\mathbf{A}}_{apt}^k \mathbf{X}\mathbf{W}_2, \tag{1}$$

where $\mathbf{P}^k$ represents the power series of the transition matrix, $\mathbf{W} \in \mathbb{R}^{Cin\times M}$ denotes a trainable paramater matrix (where M denotes the hidden dimension), $\hat{\mathbf{A}} \in \mathbb{R}^{N\times Cin}$ the normalized adjacency matrix with self-loops, $\hat{\mathbf{A}}_{apt}$ the self-adaptive adjacency matrix, and $\mathbf{Z} \in \mathbb{R}^{N\times M}$ the output signal. Note that the self-adaptive adjacency matrix is designed to discover hidden spatial dependencies by itself, which is computed as follows:

$$\hat{\mathbf{A}}_{apt} = Softmax(\sigma(\boldsymbol{E}_1 \boldsymbol{E}_1^T)), \tag{2}$$

where $\boldsymbol{E}_1 \in \mathbf{R}^{N\times c}$ is the node embedding dictionary with learnable parameters, c is the dimension of the embedding, $0 \le k \le K$ denotes the power series. The power series of the transition matrix $\mathbf{P^k}$ can be created in several ways, in the particular case of an undirected graph can be computed as $\mathbf{P} = \mathbf{A}/\sum_{i=1}^{N} a_{ji}$. Note that in some works the self-adaptive adjacency matrix is known as Laplacianlink [24], [28].

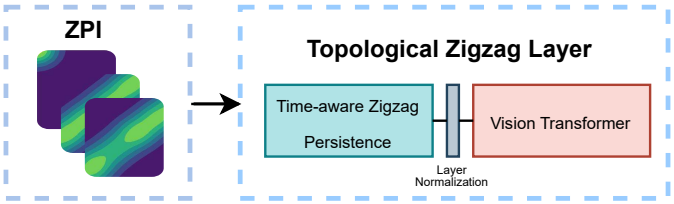

Fig. 1: Topological Zigzag Layer.

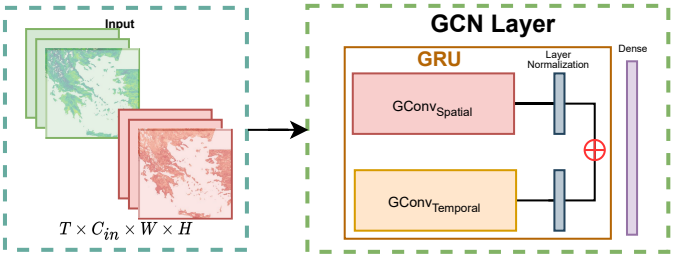

Fig. 2: Graph Convolutional Network Layer.

IV. Firecast Zigzag Graph Convolutional Network

We formulate the task of wildfire risk prediction as a binary classification problem. That is, given spatio-temporal information $\mathbf{X}_t$, we assess whether a pixel-of-interest (POI) will be on fire at the next time step $t+1$. This problem is widely known as pixel-wise classification.

To tackle this task, we propose a model, namely, Firecast Zigzag Convolutional Network (F-ZCN) that combines information of a topological zigzag layer with a topological Graph Convolutional Layer by means of a Gate Recurrent Unit. This process is detailed in Figure 3, where the spatio-temporal information is used as input for each layer.

First, it applies a topological zigzag layer (Figure 1) which uses ZPI information. Formaly, it is defined as follows:

$$\mathbf{ZO} = ViT(\mathbf{PI}) \tag{3}$$

where $\mathbf{PI} \in \mathbb{R}^{p \times p}$ and $\mathbf{ZO} \in \mathbb{R}^{2 \times 1}$. Note that depending on the configuration, a zigzag filtration might generate different signatures. For simplicity, in our case we concatenate p-dimensional topological features (0-dimensional and 1-dimensional holes) as different input channels. Then, the zigzag information is passed through a layer normalization, which enables smoother gradients, faster training times, and achieves better generalization [29]. After that, a ViT is used to learn the topological patterns that can be found in the ZPI. Note that this ViT version uses multi-head mechanisms attentions, and position embedding that allows us to extract of ZPI information efficiently [30]. Second, we include a Graph Convolutional Layer (Figure 2). This layer includes spatial and temporal graph convolution, respectively. While the former is to capture spatial correlation between nodes, the latter is to capture temporal correlation between features in different time slices. We adopt Equation (III) in our architecture to apply spatial convolutions, where $\mathbf{W}_1$ and $\mathbf{W}_2$ are trainable weights of the network, $\mathbf{P}_1$ is computed with a power series of the adjacency matrix $\mathbf{A}$. The output of this spatial convolution is multiplied by a hidden matrix $\mathbf{W}_3$:

$$\mathbf{S} = \mathbf{Z}\mathbf{W}_3 \tag{4}$$

where $\mathbf{W}_3 \in \mathbb{R}^{M \times H/2}$ and $\mathbf{S} \in \mathbb{R}^{N \times H/2}$.

In addition to the spatial domain, we apply a convolution to all the time slices of the graph as follows:

$$\mathbf{L} = \mathbf{W}_4 \mathbf{X} \mathbf{W}_5 \tag{5}$$

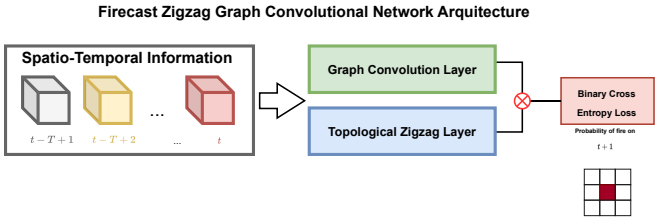

Fig. 3: Architecture of F-ZCN. Spatio-temporal information is given as input to GCN layer and Topological Zigzag Layer, then the output of each layer is combined and processed by the binary cross-entropy loss, which represents the probability of fire on a pixel-of-interest.

where $\mathbf{W}_4 \in \mathbb{R}^{T \times 1}$ is a vector of learnable weights, $\mathbf{W}_5 \in \mathbb{R}^{C_{in} \times H/2}$ is a hidden matrix, and $\mathbf{L} \in \mathbb{R}^{N \times H/2}$ is the output of the temporal convolution. On the last step of the graph convolutional layer is concatenated the outputs of the spatial convolution ($\mathbf{Z}$) and temporal convolution ($\mathbf{L}$) as $\mathbf{Q} = Dense(COMBINE(\mathbf{Z}, \mathbf{L}))$, where $\mathbf{Q} \in \mathbb{R}^{2 \times 1}$. The output of combining both layers is downsized by a dense layer. Note that in several parts of the network is included a layer normalization layer which enables smoother gradients, faster training times, and achieves better generalization [29]. Finally, information of both layers $\mathbf{ZO}$ and $\mathbf{Q}$ are added and passed to the binary cross entropy loss function.

V. Experiments

We illustrate the proposed F-ZCN model in application to forecasting wildfire occurrences in Eastern Mediterranean and the African continent at various resolutions. We also conduct an ablation study to assess individual contributions of the F-ZCN components.

A. Datasets

We consider two regions: Eastern Mediterranean and surrounding areas (defined in a lower resolution) and the African continent (defined in a higher resolution).

Eastern Mediterranean. We consider the dataset proposed in [13], which consists of 1 km $\times$ 1 km $\times$ 1 day resolution. The region of interest of this dataset is centered around Greece covering a total area of $1,253$ km $\times$ 983 km. Note that, we consider the following 25 variables as predictors and can be grouped as follows: (i) *Daily wheater data* from ERA-5 [3]: Land of maximum 2m temperature, maximum wind speed, minimum relative humidity, total precipitation,

TABLE I: Wildfire risk prediction comparison between different DL methods on test set 2020 and test set 2021. Note that the year 2020 represents a typical fire season and the year 2021 represents an extreme one.

Year 2020	Precision	Recall	F1-Score	Accuracy	AUC	AUCPR
ConvLSTM	0.842 ± 0.032	0.660 ± 0.030	0.739 ± 0.027	84.52% ± 0.015	0.911 ± 0.023	0.857 ± 0.030
LSTM	0.834 ± 0.026	0.696 ± 0.046	0.757 ± 0.023	85.18% ± 0.010	0.923 ± 0.005	0.864 ± 0.008
GCN	0.877 ± 0.014	0.683 ± 0.024	0.768 ± 0.015	86.24% ± 0.007	0.936 ± 0.004	0.893 ± 0.007
F-ZCN	**0.907** ± 0.016	**0.706** ± 0.025	**0.793** ± 0.012	**87.76%** ± 0.005	**0.951** ± 0.003	**0.919** ± 0.006
Year 2021	**Precision**	**Recall**	**F1-Score**	**Accuracy**	**AUC**	**AUCPR**
ConvLSTM	0.901 ± 0.019	0.905 ± 0.027	0.903 ± 0.015	93.49% ± 0.009	0.978 ± 0.006	0.958 ± 0.012
LSTM	0.891 ± 0.015	0.859 ± 0.039	0.874 ± 0.019	91.77% ± 0.010	0.942 ± 0.003	0.942 ± 0.007
GCN	0.923 ± 0.012	0.905 ± 0.012	0.914 ± 0.009	94.31% ± 0.006	0.981 ± 0.002	0.964 ± 0.005
F-ZCN	**0.946** ± 0.008	**0.908** ± 0.005	**0.927** ± 0.003	**95.21%** ± 0.002	**0.985** ± 0.000	**0.971** ± 0.002

TABLE II: Wildfire risk prediction comparison between spatio-temporal DL methods on test sets 2019 (left side) and 2020 (right side) on the African continent.

Year 2019	Precision	Recall	F1-Score	Accuracy	AUC	AUCPR
ConvLSTM	0.949 ± 0.005	0.984 ± 0.007	0.966 ± 0.004	0.966 ± 0.004	0.997 ± 0.001	0.997 ± 0.001
GCN	0.965 ± 0.011	**0.994** ± 0.001	0.979 ± 0.005	0.979 ± 0.006	0.999 ± 0.000	0.999 ± 0.000
F-ZCN	**0.979** ± 0.006	0.992 ± 0.002	**0.986** ± 0.002	**0.985** ± 0.002	**0.999** ± 0.000	**0.999** ± 0.000
Year 2020	**Precision**	**Recall**	**F1-Score**	**Accuracy**	**AUC**	**AUCPR**
ConvLSTM	0.966 ± 0.005	0.967 ± 0.006	0.966 ± 0.004	0.966 ± 0.004	0.996 ± 0.001	0.996 ± 0.001
GCN	0.972 ± 0.010	0.986 ± 0.003	0.979 ± 0.004	0.979 ± 0.005	0.998 ± 0.000	0.999 ± 0.000
F-ZCN	**0.978** ± 0.006	**0.987** ± 0.002	**0.982** ± 0.003	**0.982** ± 0.003	**0.999** ± 0.000	**0.999** ± 0.000

TABLE III: Ablation study of the F-ZCN.

Year 2020	Precision	Recall	F1-Score	Accuracy	AUC	AUCPR
F-ZCN	0.907 ± 0.016	**0.706** ± 0.025	**0.793** ± 0.012	**0.878** ± 0.005	**0.951** ± 0.003	**0.919** ± 0.006
W/o Zigzag learning	0.877 ± 0.014	0.683 ± 0.024	0.768 ± 0.015	0.862 ± 0.007	0.936 ± 0.004	0.893 ± 0.007
W/o GCNSpatial	**0.918** ± 0.014	0.660 ± 0.017	0.768 ± 0.009	0.867 ± 0.004	0.946 ± 0.003	0.909 ± 0.005
W/o Adjacency	0.893 ± 0.007	0.686 ± 0.012	0.776 ± 0.006	0.868 ± 0.003	0.940 ± 0.003	0.900 ± 0.004
W/o Self-Adaptive Adjacency	0.880 ± 0.033	0.701 ± 0.054	0.778 ± 0.020	0.867 ± 0.006	0.942 ± 0.003	0.902 ± 0.006
Year 2021	**Precision**	**Recall**	**F1-Score**	**Accuracy**	**AUC**	**AUCPR**
F-ZCN	**0.946** ± 0.008	0.908 ± 0.005	**0.927** ± 0.003	**0.952** ± 0.002	**0.985** ± 0.000	**0.971** ± 0.002
W/o Zigzag learning	0.923 ± 0.012	0.905 ± 0.012	0.914 ± 0.009	0.943 ± 0.006	0.981 ± 0.002	0.964 ± 0.005
W/o GCNSpatial	0.944 ± 0.009	0.907 ± 0.005	0.925 ± 0.003	0.951 ± 0.002	0.984 ± 0.001	0.970 ± 0.004
W/o Adjacency	0.933 ± 0.008	**0.911** ± 0.009	0.922 ± 0.003	0.949 ± 0.002	0.983 ± 0.001	0.966 ± 0.005
W/o Self-Adaptive Adjacency	0.944 ± 0.009	0.907 ± 0.005	0.925 ± 0.003	0.951 ± 0.002	0.984 ± 0.001	0.970 ± 0.004

maximum 2m dew point temperature, and maximum surface pressure; (ii) *Saltellite variables* from MODIS: Normalized Difference Vegetation Index (NDVI), Day and Night Land Surface Temperature; (iii) *Soil* from the European Drought Observatory [31]: *Soil moisture index*; (iv) *Geographic and Demographic* from Worldpop [32] and Copernicus EU-DEM: Roads distance, waterway distance, yearly population density, elevation and slope. Data collection on land from Copernicus Corine Land Cover [33].

Note that the first ten variables are dynamic and the remaining ones are static. For this dataset the years from 2009 to 2018 are for training set and 2019 for validation set. Regarding testing sets, we take into account two sets, one for each year; years 2020 and 2021. All four datasets consists of $40,554$ training ($27,036$ nonfire, $13,518$ fire), $3,900$ validation ($2,600$ nonfire, $1,300$ fire), $3,684$ testing ($2,456$ nonfire, $1,228$ fire) samples for 2020, and $13,221$ testing ($8,814$ nonfire, $4,407$ fire) samples for 2021. To overcome the problem of highly unbalanced classes caused by the scarcely amount of fires, for each run, the negatives are randomly sampled two times compared to positives [34]. Since our goal is to predict the center pixel and to decrease the risk of sampling negatives that in fact represent fire dangers different from the center pixel, this sampling process selects negatives from days when no fire occurred on the entire patch or region of interest. In this dataset, we use two modalities of samples: temporal and spatio-temporal. The temporal dataset consists of the time series of days $\{t-1, t-2, \ldots, t-10\}$ of the dynamic input observations, which exploits the temporal context. Furthermore, the spatial-temporal dataset consists of 25 km $\times$ 25 km $\times$ 10 days blocks of the dynamic input observations centered spatially around the given cell. Note that some features are static (e.g., yearly population density) therefore are repeated in time.

African Continent. We further study the benefits of our proposed method by analyzing its performance across developing

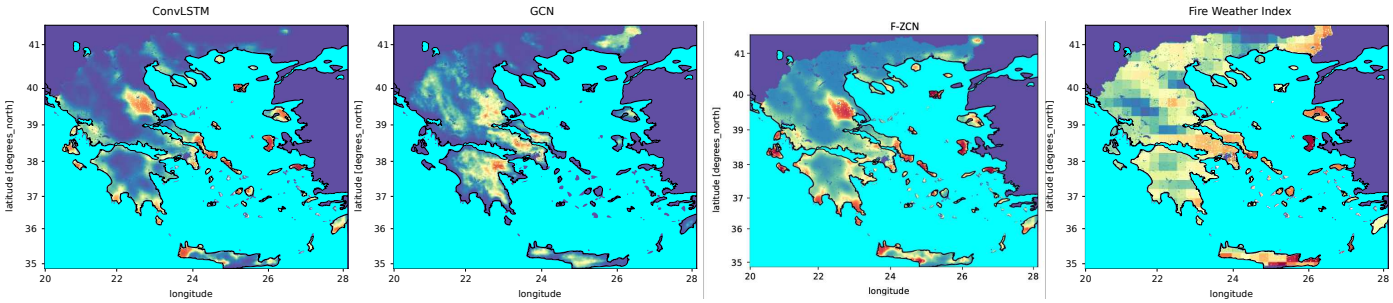

Fig. 4: Fire danger for Greece on days 19/07/2020. DL methods offer a better resolution than the empirical Fire Weather Index.

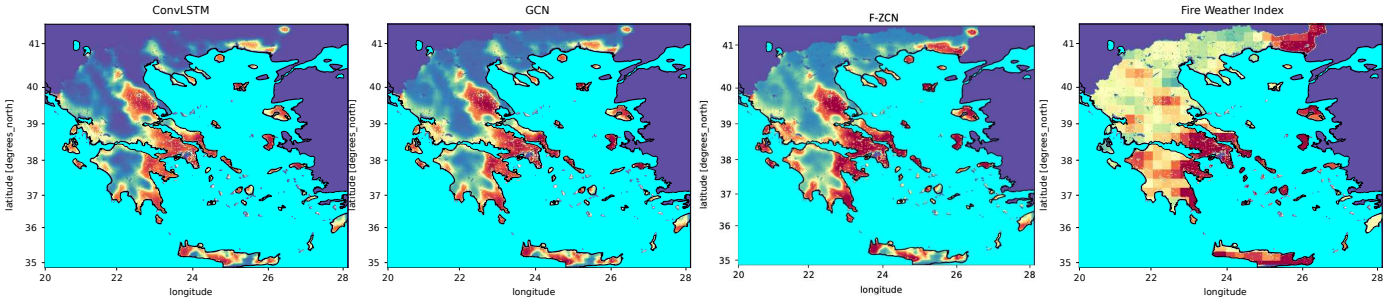

Fig. 5: Fire danger for Greece on days 21/07/2020. DL methods offers a better resolution than the empirical Fire Weather Index.

regions in the African continent. In particular, we build a dataset for the Africa continent by taking as a source the scientific datacube for seasonal fire forecasting around the globe [35]. This dataset consists of a 0.25 degrees spatial resolution, and 8 days of temporal resolution. Next 7 variables serve as input features in our experiments: total precipitation, sea surface temperature, average temperature at 2 Meters, drought code average, Normalized Difference Vegetation Index (NDVI), population density, and Downwards Surface Solar Radiation (DSSR). Similarly, this dataset consists of spatio-temporal blocks of 5 km $\times$ 5 km $\times$ 10 time dimensions, where each time slice is the average along 8 days. Here, the training set is taken from years 2009-2017, thus the validation set comes from 2018; whilst for testing we use years 2019 and 2020. All our datasets consist of $9,000$ samples for training (1000 samples per year), and $16,000$ samples for each remaining set.

B. Experimental Settings

This experimental validation takes into account two baseline methods widely used for wildfire prediction: LSTM, from [36], and ConvLSTM, from [37]. Note that both methods are considered as standard for image processing tasks. Our experimentation was carried out on an NVIDIA Tesla T4 GPU card with 32GB of memory. LSTM and ConvLSTM were configured as suggested by their authors [13]. GCN and F-ZCN use 1 hidden GCN layer with size 18, node embedding of size 64, and spatial convolution size $K = 2$. Each method was run 10 times on their validation set using different seeds, stopping criterion of 30 epochs, batch size is 256, and L2-regularization weight of 0.001. Source codes are available at https://github.com/yuzhouguangc/F-ZCN.git.

VI. Ablation Study

To have compelling insights on the importance of each component within F-ZCN, we conduct ablation studies on the Eastern Mediterranean dataset, focused on Greece, and present the results in Table III. In particular, we test F-ZCN w/o the zigzag learning module, spatial graph convolution, fixed adjacency matrix, and self-adaptive adjacency matrix. Note that the last two elements aim to verify the contribution using a static adjacency matrix, and learning an adjacency matrix, respectively. Results confirm that each component is beneficial for F-ZCN. Specifically, we can found that, for AUC, the relative gains of F-ZCN over F-ZCN w/o Zigzag learning are 1.603% and 0.408% on datasets in Year 2020 and Year 2021; for AUCPR, the relative gains of F-ZCN over F-ZCN w/o Zigzag learning are 2.911% and 0.721% on datasets in Year 2020 and Year 2021. Although removing the spatial graph convolutions produces an improvement in terms of precision, this action deteriorates the F1-Score metric and Accuracy. Similarly, removing a fixed adjacency matrix improves the Recall but deteriorates the remaining metrics. These results reveal that the inclusion of the zigzag topological layer is needed in order to improve the wildfire prediction and overall performance of the proposed F-ZCN. Finally, it is relevant to remark that each variant in this ablation study performs better than all baseline methods in terms of F1-Score and Accuracy, which supports the effectiveness of our TDA-based DL model.

VII. Lessons Learned and Path to Deployment

Table I shows the six scores that represent the performance of the two baseline methods and GCNs. Overall, regardless of the score, F-ZCN shows the best performance followed by GCN. For instance, for accuracy, the relative gains of F-ZCN

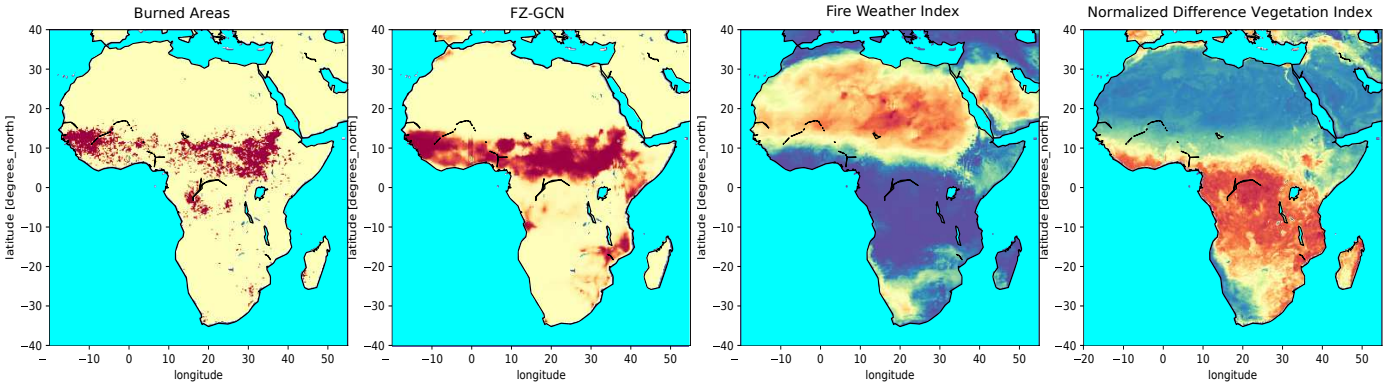

Fig. 6: Fire danger for Africa on day 21/03 /2020. Each date represents the average of 8 days. Reported are Burned Areas, F-ZCN, Fire Weather Index, and Normalized Different Vegetation Index, respectively.

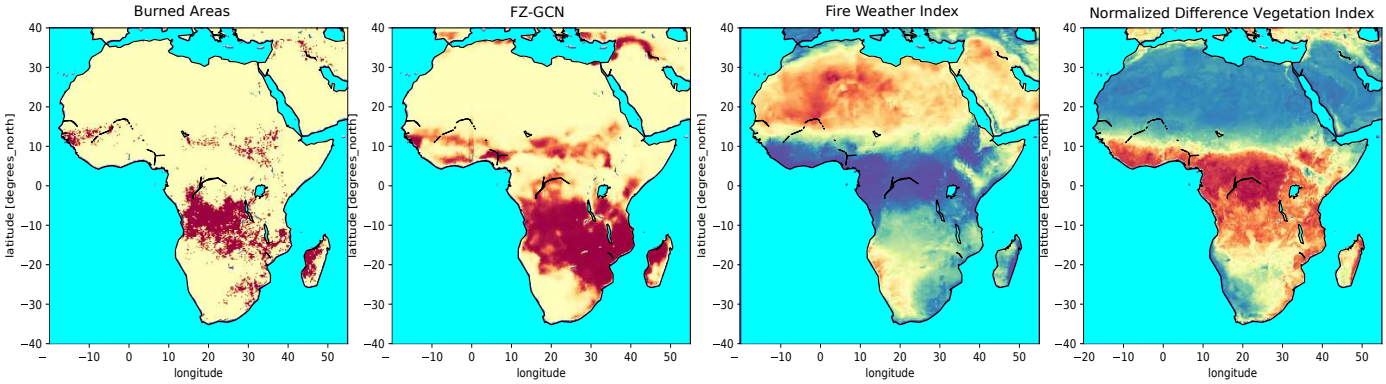

Fig. 7: Fire danger for Africa on day 16/05/2020. Each date represents the average of 8 days. Reported are Burned Areas, F-ZCN, Fire Weather Index, and Normalized Different Vegetation Index, respectively.

over the runner-up, i.e., GCN are 1.763% and 0.954% on datasets in the Year 2020 and Year 2021 respectively. This demonstrates the value added by the time-aware topological layer in modeling the spatio-temporal risk of wildfires. Figures 4-7 show that the fire risk from the DL models outperform FWI. First of all, the DL models highlight the area with actual fire occurrences with greater contrast in the fire danger than FWI. In addition, FWI's spatial resolution is limited to resolution of the meteorological field used to calculate FWI, whereas the DL models can predict fire risk at fine spatial resolution. The high-resolution prediction of fire risk is important, especially for Greece with its complex topography and archipelago of small islands. FWI tends to be relatively high over hot and dry desert areas as shown in Figure 6. However, due to the lack of vegetation, the high FWI over the Sahara is unlikely related to considerable fire danger. The reasonable agreement between the burned area observation and F-ZCN output indicates that it is important to use NDVI in predicting fire danger.

VIII. Conclusion and Future Work

Development of innovative early warning mechanisms for wildfire management is the key towards improving wildfire preparedness, preventing property damages and saving lives. In this project, we have introduced the GNN approach, boosted with time-aware topological signatures for wildfire prediction. The integration of time-conditioned topological descriptors enriches the pool of features from which significant non-linear relationships of fire drivers can be obtained. Our findings in Greece and the African continent have shown that F-ZCN delivers more accurate fire risk prediction regardless of the local climate conditions, thereby, addressing the limitations of the currently adopted approaches such as FWI. From a social perspective, this paper aims to add to the set of new solutions that can potentially reduce the gap between developed and developing countries in terms of efficient and costless wildfire prediction.

As future work, we plan to further enhance the prediction of wildfire events by introducing health, demographic, and non-environmental factors, and address the explainability and fairness of assessing wildfire severity [38]. Furthermore, we will explore solutions to address the problem of fusing multi-source data, especially, in the absence of labelled fire indices. This can be approached, for example, with the concepts of contrastive learning [39]. Another important direction for developing pro-active and reliable wildfire prediction is associated with a better understanding of the impact of data resolution on decision making tasks [40], and TDA tools may be a promising approach to tackle this fundamental problem [41], [42].

ACKNOWLEDGMENTS

This work has been supported in part by the NASA AIST grant 21-AIST21_2-0059. In addition, Y.C., H.L., and N.L. have been supported by the NSF Grant DMS-2335846/2335847. Part of this material is also based upon work supported by (while serving at) the NSF. The views expressed in the article do not necessarily represent the views of NSF and NASA.

REFERENCES

[1] M. W. Jones, J. T. Abatzoglou, S. Veraverbeke, N. Andela, G. Lasslop, M. Forkel, A. J. Smith, C. Burton, R. A. Betts, G. R. van der Werf *et al.*, "Global and regional trends and drivers of fire under climate change," *Rev. Geophys.*, vol. 60, no. 3, p. e2020RG000726, 2022.

[2] C. Hoffman, J. Canfield, R. Linn, W. Mell, C. Sieg, F. Pimont, and J. Ziegler, "Evaluating crown fire rate of spread predictions from physics-based models," *Fire Technology*, vol. 52, pp. 221–237, 2016.

[3] J. Muñoz-Sabater, E. Dutra, A. Agustí-Panareda, C. Albergel, G. Arduini, G. Balsamo, S. Boussetta, M. Choulga, S. Harrigan, H. Hersbach *et al.*, "Era5-land: A state-of-the-art global reanalysis dataset for land applications," *Earth System Science Data*, vol. 13, no. 9, pp. 4349–4383, 2021.

[4] C. Van Wagner and T. Pickett, *Equations and FORTRAN program for the Canadian forest fire weather index system*. Government of Canada, 1985, vol. 33.

[5] J. T. Abatzoglou, A. P. Williams, L. Boschetti, M. Zubkova, and C. A. Kolden, "Global patterns of interannual climate–fire relationships," *Global change biology*, vol. 24, no. 11, pp. 5164–5175, 2018.

[6] F. Di Giuseppe, C. Vitolo, B. Krzeminski, C. Barnard, P. Maciel, and J. San-Miguel, "Fire weather index: the skill provided by the european centre for medium-range weather forecasts ensemble prediction system," *Nat. Hazards and Earth System Sci.*, vol. 20, no. 8, pp. 2365–2378, 2020.

[7] P. F. Moore, "Global wildland fire management research needs," *Current Forestry Reports*, vol. 5, pp. 210–225, 2019.

[8] P. Jain, S. C. Coogan, S. G. Subramanian, M. Crowley, S. Taylor, and M. D. Flannigan, "A review of machine learning applications in wildfire science and management," *Environmental Reviews*, vol. 28, no. 4, pp. 478–505, 2020.

[9] D. Shadrin, S. Illarionova, F. Gubanov, K. Evteeva, M. Mironenko, I. Levchunets, R. Belousov, and E. Burnaev, "Wildfire spreading prediction using multimodal data and deep neural network approach," *Scientific Reports*, vol. 14, no. 1, p. 2606, 2024.

[10] Z. Xu, J. Li, S. Cheng, X. Rui, Y. Zhao, H. He, and L. Xu, "Wildfire risk prediction: A review," *arXiv:2405.01607*, 2024.

[11] M. Naderpour, H. M. Rizeei, and F. Ramezani, "Forest fire risk prediction: A spatial deep neural network-based framework," *Remote Sensing*, vol. 13, no. 13, p. 2513, 2021.

[12] D. Radke, A. Hessler, and D. Ellsworth, "Firecast: Leveraging deep learning to predict wildfire spread." in *IJCAI*, 2019, pp. 4575–4581.

[13] S. Kondylatos, I. Prapas, M. Ronco, I. Papoutsis, G. Camps-Valls, M. Piles, M.-Á. Fernández-Torres, and N. Carvalhais, "Wildfire danger prediction and understanding with deep learning," *Geophys. Res. Lett.*, vol. 49, no. 17, p. e2022GL099368, 2022.

[14] J. Chen, Y. Yang, L. Peng, L. Chen, and X. Ge, "Knowledge graph representation learning-based forest fire prediction," *Remote Sensing*, vol. 14, no. 17, p. 4391, 2022.

[15] S. Singla, A. Mukhopadhyay, M. Wilbur, T. Diao, V. Gajjewar, A. Eldawy, M. Kochenderfer, R. Shachter, and A. Dubey, "Wildfiredb: An open-source dataset connecting wildfire spread with relevant determinants," in *NeurIPS*, 2021.

[16] F. Chazal, V. De Silva, and S. Oudot, "Persistence stability for geometric complexes," *Geometriae Dedicata*, vol. 173, no. 1, pp. 193–214, 2014.

[17] H. Kim and C. Vogel, "Deciphering active wildfires in the southwestern usa using topological data analysis," *Climate*, vol. 7, no. 12, p. 135, 2019.

[18] R. Bendick and Z. H. Hoylman, "Topological data analysis reveals parameters with prognostic skill for extreme wildfire size," *Env. Research Letters*, vol. 15, no. 10, p. 104039, 2020.

[19] G. Carlsson and M. Vejdemo-Johansson, *Topological data analysis with applications*. Cambridge University Press, 2021.

[20] G. Carlsson and V. De Silva, "Zigzag persistence," *Foundations of computational mathematics*, vol. 10, pp. 367–405, 2010.

[21] A. Tausz and G. Carlsson, "Applications of zigzag persistence to topological data analysis," *arXiv preprint arXiv:1108.3545*, 2011.

[22] G. Carlsson, V. De Silva, S. Kališnik, and D. Morozov, "Parametrized homology via zigzag persistence," *Algebraic & Geometric Topology*, vol. 19, no. 2, pp. 657–700, 2019.

[23] H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, and L. Ziegelmeier, "Persistence images: A stable vector representation of persistent homology," *JMLR*, vol. 18, 2017.

[24] Y. Chen, I. Segovia, and Y. R. Gel, "Z-GCNETs: Time zigzags at graph convolutional networks for time series forecasting," in *ICML*, 2021, pp. 1684–1694.

[25] Y. Chen, Y. Gel, and H. V. Poor, "Time-conditioned dances with simplicial complexes: Zigzag filtration curve based supra-hodge convolution networks for time-series forecasting," *NeurIPS*, vol. 35, pp. 8940–8953, 2022.

[26] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[27] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph wavenet for deep spatial-temporal graph modeling," *arXiv:1906.00121*, 2019.

[28] Y. Chen, I. Segovia-Dominguez, B. Coskunuzer, and Y. Gel, "Tamps2gcnets: coupling time-aware multipersistence knowledge representation with spatio-supra graph convolutional networks for time-series forecasting," in *ICLR*, 2022.

[29] J. Xu, X. Sun, Z. Zhang, G. Zhao, and J. Lin, "Understanding and improving layer normalization," *NeurIPS*, vol. 32, 2019.

[30] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," in *ICLR*, 2010.

[31] C. Cammalleri, J. V. Vogt, B. Bisselink, and A. de Roo, "Comparing soil moisture anomalies from multiple independent sources over different regions across the globe," *Hydrology and Earth System Sciences*, vol. 21, no. 12, pp. 6329–6343, 2017.

[32] A. J. Tatem, "Worldpop, open data for spatial demography," *Scientific data*, vol. 4, no. 1, pp. 1–4, 2017.

[33] G. Büttner, "Corine land cover and land cover change products," *Land use and land cover mapping in Europe: practices & trends*, pp. 55–74, 2014.

[34] F. Huot, R. L. Hu, M. Ihme, Q. Wang, J. Burge, T. Lu, J. Hickey, Y.-F. Chen, and J. Anderson, "Deep learning models for predicting wildfires from historical remote-sensing data," *arXiv:2010.07445*, 2020.

[35] L. Alonso, F. Gans, I. Karasante, A. Ahuja, I. Prapas, S. Kondylatos, I. Papoutsis, E. Panagiotou, D. Mihail, and F. Cremer, "Seasfire cube: A global dataset for seasonal fire modeling in the earth system," 2022.

[36] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.

[37] X. Shi, Z. Chen, H. Wang, D.-Y. Yeung, W.-K. Wong, and W.-c. Woo, "Convolutional lstm network: A machine learning approach for precipitation nowcasting," in *NeurIPS*, vol. 28, 2015.

[38] A. Abdollahi and B. Pradhan, "Explainable artificial intelligence (xai) for interpreting the contributing factors feed into the wildfire susceptibility prediction model," *Science of the Total Environment*, vol. 879, p. 163004, 2023.

[39] J. W. Choi, N. LaHaye, Y. Chen, H. Lee, and Y. R. Gel, "Self-supervised contrastive learning for wildfire detection: utility and limitations," in *Advances in Machine Learning and Image Analysis for GeoAI*. Elsevier, 2024, pp. 153–163.

[40] H. Fan, X. Yang, C. Zhao, Y. Yang, and Z. Shen, "Spatiotemporal variation characteristics of global fires and their emissions," *Atmospheric Chemistry and Physics*, vol. 23, no. 13, pp. 7781–7798, 2023.

[41] D. Ofori-Boateng, H. Lee, K. M. Gorski, M. J. Garay, and Y. R. Gel, "Application of topological data analysis to multi-resolution matching of aerosol optical depth maps," *Frontiers in Environmental Science*, vol. 9, p. 684716, 2021.

[42] L. Ver Hoef, H. Adams, E. J. King, and I. Ebert-Uphoff, "A primer on topological data analysis to support image analysis tasks in environmental science," *Artificial Intelligence for the Earth Systems*, vol. 2, no. 1, 2023.