Wildfire Prediction Using Spatio-Temporal Knowledge Graphs

Martin Böckling[1,2,*,†], Heiko Paulheim[1,*,†] and Sarah Detzler[2,*,†]

[1]*University of Mannheim, Data and Web Science Group, B6, 26, 68161 Mannheim*
[2]*SAP SE, Dietmar-Hopp-Allee 16, 69190 Walldorf*

Abstract

The frequency of wildfires increases yearly and poses a constant threat to the environment and human beings. Different factors, for example surrounding infrastructure to an area (e.g., campfire sites or power lines) contribute to the occurrence of wildfires. In this paper, we propose using a Spatio-Temporal Knowledge Graph (STKG) based on OpenStreetMap (OSM) data for modeling such infrastructure. Based on that knowledge graph, we use the RDF2vec approach to create embeddings for predicting wildfires, and we align different vector spaces generated at each temporal step by partial rotation. In an experimental study, we determine the effect of the surrounding infrastructure by comparing different data composition strategies, which involve a prediction based on tabular data, a combination of tabular data and embeddings, and solely embeddings. We show that the incorporation of the STKG increases the prediction quality of wildfires.

Keywords

Knowledge Graph Embeddings, Spatio-Temporal Knowledge Graph, RDF2vec, Wildfire Prediction

1. Introduction

Wildfires as a natural phenomenon have increased in their frequency as well as in their damage to the territory of the United States. Wildfires are unplanned fires which can have natural causes, human-related causes as also have their origin in planned fires. When analyzing the period from 1960 to 2021, the top five highest burned areas by wildfires per year were all after 2006. For the year 2021 in total, 58,968 wildfires have been detected over the complete territory of the United States with a burned area around 7,100,000 acres [9]. The prediction of wildfires relies on factors which can be characterized as complex systems, in which different influence factors and surrounding factors are present in contributing to the occurrence of wildfires [20]. A base argumentation for complex systems in geography was determined by Tobler's first law, for which he stated that "Everything is related to everything else, but near things are more related than distant things" [27]. Our approach, which builds the base for the wildfire prediction

in this paper, builds up on this premise and follows the hypothesis that an involvement of a Knowledge Graph (KG) is beneficial to predicting wildfires. Wildfire prediction in the context of this paper represents the prediction of a wildfire occurrence for the given input types of one given spatial grid cell to the specified time step.

2. Related Work

For the topic of wildfire prediction, a variety of different research papers exist. Nami et al. [17] calculate the probability of wildfires for the Hyrcanian ecological region. For the probability prediction of wildfires temperature variables, landscape information and spatial property data are incorporated into the final dataset. All variables are modeled on a regular spatial grid to integrate them into a unified dataset. The spatial property data is organized by using distance bins, unfolding into a category for each bin [17]. A different study of wildfire prediction is conducted for the territory of South Korea. It includes landscape data, weather data as also social-economic data and the distance to spatial property data. Each dataset is transformed together by using a regular spatial grid. The modeling is performed by using a Maximum Entropy and Random Forest classifier [11].

Different research papers have used OSM data to construct a Spatial Knowledge Graph or enrich other Knowledge Graphs with data from OSM. Dsouza et al. [6] have constructed a Knowledge Graph called WorldKG, globally integrating the OSM data into a Knowledge Graph. For the construction of WorldKG, the data is derived from OSM and inserted into a hierarchical structure using the key-value pairs from OSM. The Ontology of WorldKG uses the keys as classes and the key-value pairs as subclasses. Furthermore, the spatial object structures are captured regarding the spatial geometry types and the exact structure of the geometry types stored in the Well Known Text (WKT) format. In total, the published Knowledge Graph consists of 828,550,751 triples and 113,444,975 total entities [5]. Johnson et al. [10] have created a concept for different Knowledge Graphs consisting of OSM data, social-economic data and flood data. For the different datasets modeled, the structure of the Knowledge Graph varies in the structure. For the social-economic dataset the data is divided into fixed regions, for the OSM data the overall geometry structure of single buildings [10]. The paper of Wu et al. [30] uses National Oceanic and Atmospheric Administration (NOAA) weathers data to model it into a knowledge graph. For each weather station, the data type together with the location is modeled into a Knowledge Graph. Furthermore, the position of each weather station is encoded into two separate entities separated between longitudes and latitudes. Furthermore, the Knowledge Graph is enriched together with the data derived from OSM and Wikidata [30].

Given the presented papers that were available for our research, the methodology used in this paper is the first to create a STKG with OSM data modeled on a spatial grid over continuous time steps. Especially, other papers and approaches do not model their KG using a spatial grid to capture dynamic changes over continuous time steps. The use of KGs for wildfire prediction is, to the best of our knowledge, the first time used to model spatial relationships between geographic entities.

3. Approach

For this paper, the approach for our research results in three different research datasets that we use for predicting wildfires. The first dataset, named BaseCase dataset, includes tabular data from a variety of different data sources, which we integrate by using a spatial grid. This approach is consistent with the current state of the art and is used similarly by the presented work in section 2 [17, 11]. A general overview of our approach can be found in figure 1. The second dataset, which is named HybridCase dataset, uses as a base for the dataset structure the BaseCase dataset. In addition, the embeddings derived from the STKG using RDF2vec are introduced to the dataset as new columns. The third dataset is using solely the embeddings derived from the constructed STKGs by integrating every dataset into the KG. In the following paragraphs, the procedure for our paper is outlined in detail based on figure 1.
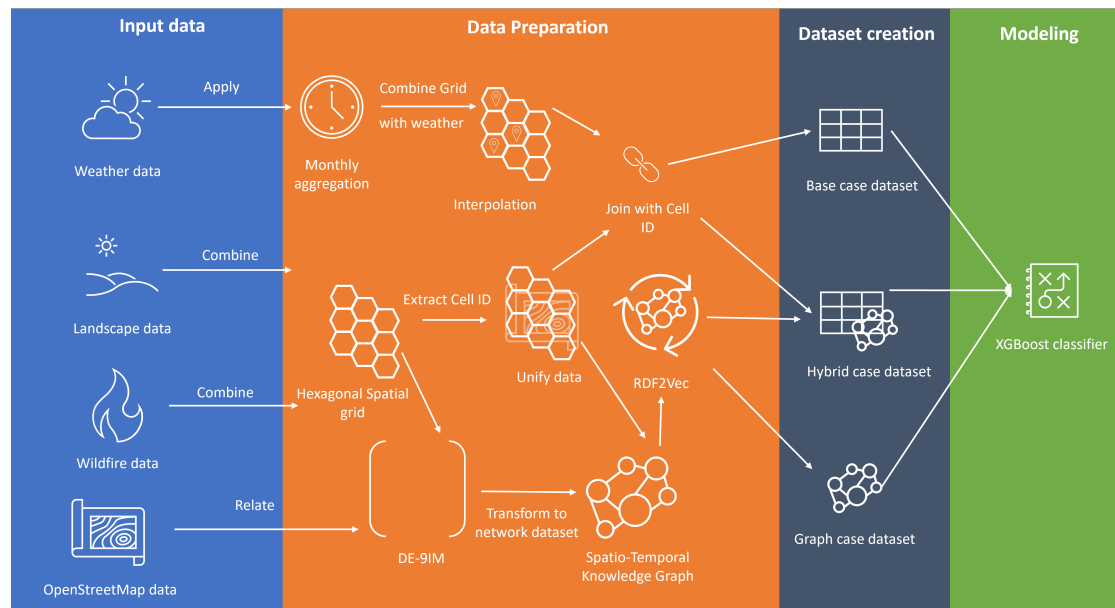

Figure 1: Approach for the wildfire detection performed in this paper

Besides the wildfire data [7] which is used to mark a wildfire occurrence, weather data [16], landscape data [31] as also a selection of OSM data [19][1] is used for the wildfire prediction datasets. The wildfire prediction in this paper is based on monthly data. If not all datasets cover the monthly periods, we either aggregate the data to a monthly basis or expand the data to a monthly basis if it does not cover all months in the period of 2010-2021. For the expansion of the datasets, we duplicate the data associated to a specific year onto the monthly basis. The extraction of the specified data can be found on the first layer in figure 1.

As we want to integrate data sources of different structure, we use spatial grids to unify the data and divide the research region into different singular cells [22]. For the created spatial grid,

---

[1]Details of OSM data selection can be found under the following page https://github.com/MartinBoeckling/WildfirePredictionSTKG/blob/main/wildfirearea/OSMTags.md.

we use a regular geometric structure for our grid, as those are best suited for the data types we are using [22]. For the selection of the geometric base structure, we use hexagons, as those allow us to aggregate data more accurately [28] and provide an equal distance to each of its neighbors [2]. By tessellating the territory of California, we construct a hexagonal spatial grid which is spanned over the complete territory of California. Each of the individual grid cells within the spatial grid is assigned with an unique Identifier (ID).

For the individual datasets, a different data preparation is needed to align the datasets for a monthly wildfire prediction. The weather dataset captures the daily weather for the period of 2010 to 2021. Therefore, the weather data needs to be aggregated on a monthly basis so that the period is matching to the required period. The selection of the aggregation method is derived by the meaning of the individual variables. If the average temperature is captured in a variable, a mean aggregation is applied to the individual station for the complete month. This paradigm is used across all variables for the weather dataset. After the aggregation of the weather data on a monthly basis, we interpolate the weather to expand the point-wise measurement over the territory of California. For the interpolation of the weather data, we rely on a combination of Inverse Distance Weighting (IDW) and Kriging. The Kriging interpolation showcased to have an accurate data representation for the interpolation [21]. However, Kriging produces also values which are above the global extreme values [15]. For our paper, the assumption is made that the simple limitation would perform worse than the IDW performed interpolation. Therefore, the replacement of invalid values violating the global limit is performed by an IDW interpolation.

For the landscape data, the dataset has data for the years 2011, 2013, 2016 and 2019. As the landscape data is on a more fine granular level, we aggregate the values of the landscape data to our created grid using the majority vote aggregation, for which the landscape category with the highest fraction in the grid cell is assigned to the respective grid cell. Afterward, the data is expanded on a monthly basis to capture each individual year for the covered period.

For the creation of our STKG, we incorporate three different data preparations depending on the structure of the incorporated datasets. The base for our STKG is the neighbor relation between each spatial grid cell. For each of the individual grid cells, the neighboring relation is modeled. Between the OSM dataset with its geometries and the individual grid cells, the relationship is determined by using the Dimensionally Extended 9-Intersection Model (DE−9IM) method [4]. From the methodology, in total seven different relations are derived: Contains, covered by, covers, crosses, intersects, overlaps and touches. An explanatory situation for the modeled relation is displayed in figure 2.

For values within the tabular dataset including all used variables, the values are transposed into individual vertices, in which the respective object ID is the subject, the variable of the value builds the predicate and the value itself is the object. Those individual transformations are aligned on the monthly basis, so that the STKG cover the necessary monthly period for the years 2010 to 2021. The outlined approach enables the possibility to model relationships not only locally, but also between neighbors. This provides the possibility to realize the first law of geography.

To make the constructed STKG applicable for the wildfire prediction, we transform the data into a vector representation. For the vector generation, we use the RDF2vec method [23]. The created STKGs differ for the HybridCase and NetworkCase dataset. For the HybridCase dataset,
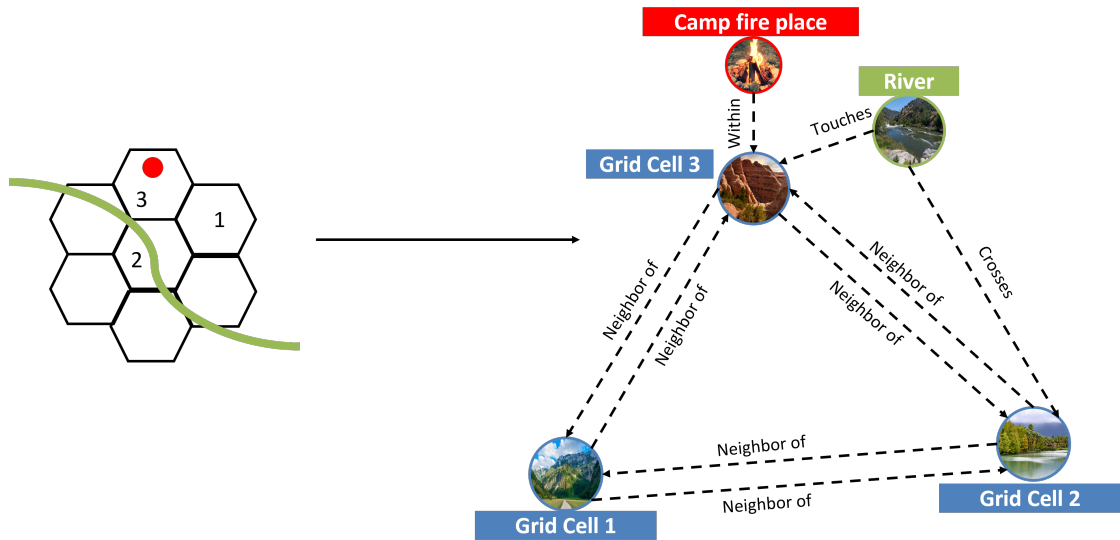
Figure 2: KG relation modeled from datasets

only the prepared OSM data is used, for the NetworkCase dataset the STKG with each dataset except the wildfire dataset is used for the vector generation. The RDF2vec has two important parameters, i.e., the maximum walk length and the maximum number of walks per entity. For the walk length, the maximum steps necessary to reach all entities related to the neighboring grid cell from one spatial grid cell is a walk distance of four. Due to the limit in available memory storage, the maximum number of walks extracted per entity is set to 1,024. As we use the RDF2vec approach for each temporal interval, we need to align the produced embeddings to a common projection. For the alignment of the embeddings, we use a methodology relying on the orthogonal Procrustes problem, where a mapping between two matrices is estimated and then applied to the matrix, which needs to be aligned [8].

Each individual dataset is joined using the grid cell ID and the month. As discussed earlier in this paragraph, we differentiate between three different dataset constellations. For the BaseCase dataset, this data consists of the prepared tabular datasets from the landscape data, weather data and wildfire data. The HybridCase dataset uses the BaseCase dataset as a base and expands the dataset with the generated embeddings for the OSM based STKG. The NetworkCase dataset uses the embeddings derived from the combined STKG consisting of landscape data, weather data and OSM data.

After the creation of the datasets, we classify wildfire events using the XGBoost (XGB) algorithm [3]. As XGB comes with a variety of different parameters, we use a Bayesian Optimization technique to tune a selection of XGB hyperparameters using the BayesianOptimization python package [18]. Bayesian Optimization for hyperparameter tuning showed promising results in comparison to other Hyperparameter Tuning methods regarding the speed of finding the optimal solution [24]. For our hyperparameter tuning, we perform ten random iterations and afterward perform 50 iterations using the Bayesian Optimization.

4. Evaluation

To evaluate the performance for all different datasets, we determine the effect of the classification model by comparing the performance with the F1 score, as it provides a better representation than the accuracy [14]. In subsection 4.1 we determine the exact testing procedure for all dataset constellations.

4.1. Evaluation Protocol

For the evaluation of the different datasets, we define the train split for all data values from 2010 to 2019. The test dataset contains all rows related to the years 2020 and 2021. As the occurrences of wildfires are with only 0.58% present in the train dataset, we consider a Random Oversampling to rebalance the train dataset. As we tune our hyperparameters, we perform a cross validation of the data. Due to the time dependency in the data, we cannot directly use a cross validation function, as we could base our prediction on future values for the past. Therefore, we split the dataset based on the time for our cross validation. For each training iteration, we only rebalance the dataset that is used for training, whereas the datasets for validation and test are not rebalanced.

For the hyperparameter tuning, we use a subset of the available parameters within XGB. The selection of the hyperparameter variables with its ranges is based on the recommendation from the official XGB website. Under the XGB website, the recommended variables are outlined. The XGB classifier uses a fixed seed to make the results reproducible. As an objective for the algorithm, we use the binary:logistic configuration together with the parameter tree_method and value hist. For the overall validation routine, we use the F1 score to determine the best parameter combination, to determine the optimal parameters for the different dataset combinations. In table 1 the overview of the tuned parameters with the value range can be found.

Table 1
Hyperparameter selection together with parameter range

Parameter	Parameter range
min_child_weight	[0, 100]
max_depth	[0, 50]
subsample	[0.01, 1]
colsample_bytree	[0.01, 1.0]
gamma	[0, 50]
n_estimators	[50, 1000]

4.2. Results

For the comparison between the different dataset constellations, we compare in total four different datasets. The dataset BaseCase does involve the tabular data as described in subsection 3. For the hybrid data constellation, we combine the embeddings together with the tabular data represented in the HybridCase dataset. The third data constellation is based solely on the

generated embeddings derived from the constructed STKG using all datasets. In table 2 the results of the wildfire prediction after the hyperparameter tuning are outlined.

Table 2
Model results after hyperparameter tuning, the highest F1 score is marked bold

Dataset ID	True Negative	False Positive	False Negative	True Positive	F1 Score
BaseCase	718,695	82,178	13,538	25,523	0.3478±0.0010
HybridCase	729,614	71,259	13,160	25,901	**0.3803**±0.0011
NetworkCase	797,829	3,068	38,834	227	0.0107±0.0002

4.3. Discussion

For the selection of all datasets used in this paper, a bias regarding the Machine Learning (ML) results is present, as not all possible influence factors of wildfires are included. For example, long-term drought factors or social economic data, which might have a long-term effect on wildfire occurrence, are not present in the datasets [13]. Therefore, each result derived from the modeling phase needs to be critically viewed, as not all necessary influence factors for wildfires are present within the data.

The validity of the datasets used is only partially assessed. For the weather variables, the State Climate Extremes Committee (SCEC)provides extreme values for a selection of weather variables [25]. A filtering of anomaly-based methods is not conducted, as it could have filtered out patterns of weather extremes contributing to wildfire occurrences.

For our STKG we have involved a sample of OSM data. Compared to WorldKG created by [26] we only use a subset of data types specifically related to wildfire causes in California. This selection proposes the situation that not all possible geographic relations captured in OSM have been determined within the territory of California. Furthermore, the OSM data is derived on a yearly basis due to the limited performance of the OSM Application Programming Interface (API). As OSM is a community-based platform similar to Wikipedia, its data is not guaranteed as complete and correct. Based on the curation process of OSM, inconsistent data can be present as well as missing information for existing spatial properties in reality [1].

Our results are strictly restricted to the territory of California. In general, the constructed artifacts and performance metrics cannot be directly compared to other regions in the world, as the hexagonal grid and different other data sources are not directly transferable to other regions within the United States. An alternative solution for the spatial grid would be by using globally spanned hierarchical grids. Examples of spatial hierarchical grids are the h3 grid created by Uber and OpenEAGGR, which both would allow furthermore a consistent global integration of spatial data. Due to the limited area we investigated for this paper, both approaches have not been evaluated regarding their influence on the wildfire prediction.

Within the paper, different dataset constellations have been combined and used. For different data preparation combinations within the datasets, a change in the metrics for the prediction can be observed. However, it is important to state that due to the focus of the paper and experimental design, it is not the goal to assess which data preparation techniques help to boost the wildfire

prediction generally. For the different dataset constellations, the dataset HybridCase performs significantly better than all other datasets. However, the dataset NetworkCase involving the STKG embeddings showcased a significant lower performance compared to the BaseCase dataset.

5. Conclusion and Outlook

This paper has showcased that, dependent on the data preparation, the incorporation of STKG embeddings has a significant effect on the prediction of wildfires. This effect has been showcased in the hybrid data constellation, where the tabular data is combined with the STKG embeddings. Furthermore, the performance of the data constellation where the pure Knowledge Graph embedding has been used showed a significant lower performance compared to the BaseCase data constellation. Therefore, it can be derived that the construction of the STKG for the network data constellation could be improved. Instead of pivoting the tabular datasets into an edge dataset, the values could be defined as node attributes, which would describe the node more accurately. Furthermore, instead of modeling a separate STKG, the performance could be compared to the created WorldKG created by Tempelmeier and Demidova [26].

For further research, additional vectorization techniques can be tested regarding the constructed knowledge graphs and compared towards the performance of a downstream task like the wildfire prediction. The different vectorization techniques could also be compared regarding their appliance to a dynamic Knowledge Graph. Examples for alternative Knowledge Graph embedding methods could be TransE, RESCAL or neural network-based embedding approaches like Semantic Matching Energy [29]. Furthermore, the Navi approach, proposed by Krause [12], could be compared to the alignment used in this paper regarding its performance on wildfire detection [12]. Alternatively to the used XGB classifier, other classification algorithms like a Random Forest or Graph Neural Networks could be used to compare the results derived from the datasets.

Acknowledgments

References

[1]   Anahid Basiri et al. "Quality assessment of OpenStreetMap data using trajectory mining". In: 19.1 (2016), pp. 56–68. DOI: 10.1080/10095020.2016.1151213.

[2]   Colin P.D. Birch, Sander P. Oom, and Jonathan A. Beecham. "Rectangular and hexagonal grids used for observation, experiment and simulation in ecology". In: *Ecological Modelling* 206.3-4 (2007), pp. 347–359. ISSN: 03043800. DOI: 10.1016/j.ecolmodel.2007.03.041.

[3] Tianqi Chen and Carlos Guestrin. "XGBoost: A Scalable Tree Boosting System". In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Ed. by Balaji Krishnapuram et al. New York, NY, USA: ACM, 2016, pp. 785–794. ISBN: 9781450342322. DOI: 10.1145/2939672.2939785.

[4] Eliseo Clementini, Paolino Di Felice, and Peter van Oosterom. "A small set of formal topological relationships suitable for end-user interaction". In: *Advances in Spatial Databases*. Ed. by G. Goos et al. Vol. 692. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer Berlin Heidelberg, 1993, pp. 277–295. ISBN: 978-3-540-56869-8. DOI: 10.1007/3-540-56869-7_16.

[5] Gianluca Demartini et al., eds. *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. New York, NY, USA: ACM, 2021. ISBN: 9781450384469. DOI: 10.1145/3459637.

[6] Alishiba Dsouza et al. "WorldKG: A World-Scale Geographic Knowledge Graph". In: *Proceedings of the 30th ACM International Conference on Information & Knowledge Management*. Ed. by Gianluca Demartini et al. New York, NY, USA: ACM, 2021, pp. 4475–4484. ISBN: 9781450384469. DOI: 10.1145/3459637.3482023.

[7] Louis Giglio et al. *MCD64A1 MODIS/Terra+Aqua Burned Area Monthly L3 Global 500m SIN Grid V006*. 2015. DOI: 10.5067/MODIS/MCD64A1.006.

[8] William L. Hamilton, Jure Leskovec, and Dan Jurafsky. "Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change". In: *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Ed. by Katrin Erk and Noah A. Smith. Stroudsburg, PA, USA: Association for Computational Linguistics, 2016, pp. 1489–1501. DOI: 10.18653/v1/P16-1141.

[9] Katie Hoover and Laura A. Hanson. "Wildfire statistics". In: *Congressional Research Service* 2 (2021), pp. 1–3.

[10] J. Michael Johnson et al. "Knowledge graphs to support real–time flood impact evaluation". In: *AI Magazine* 43.1 (2022), pp. 40–45. ISSN: 0738-4602. DOI: 10.1002/aaai.12035.

[11] Sea Jin Kim et al. "Multi-Temporal Analysis of Forest Fire Probability Using Socio-Economic and Environmental Variables". In: *Remote Sensing* 11.1 (2019), pp. 86–105. DOI: 10.3390/rs11010086.

[12] Franz Krause. "Dynamic Knowledge Graph Embeddings via Local Embedding Reconstructions". In: *The Semantic Web: ESWC 2022 Satellite Events*. Ed. by Paul Groth et al. Vol. 13384. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2022, pp. 215–223. ISBN: 978-3-031-11608-7. DOI: 10.1007/978-3-031-11609-4_36.

[13] Jeremy S. Littell. "Drought and Fire in the Western USA: Is Climate Attribution Enough?" In: *Current Climate Change Reports* 4.4 (2018), pp. 396–406. DOI: 10.1007/s40641-018-0109-y.

[14] Oded Z. Maimon and Lior Rokach. *Data mining and knowledge discovery handbook*. 2nd ed. New York and London: Springer, 2010. ISBN: 978-0-387-09823-4.

[15] Georges Matheron. "Principles of geostatistics". In: *Economic Geology* 58.8 (1963), pp. 1246–1266. ISSN: 0361-0128. DOI: 10.2113/gsecongeo.58.8.1246.

[16] Matthew J. Menne et al. "An Overview of the Global Historical Climatology Network-Daily Database". In: *Journal of Atmospheric and Oceanic Technology* 29.7 (2012), pp. 897–910. ISSN: 0739-0572. DOI: 10.1175/JTECH-D-11-00103.1.

[17] M. H. Nami et al. "Spatial prediction of wildfire probability in the Hyrcanian ecoregion using evidential belief function model and GIS". In: *International Journal of Environmental Science and Technology* 15.2 (2018), pp. 373–384. DOI: 10.1007/s13762-017-1371-6.

[18] Fernando Nogueira. *Bayesian Optimization: Open source constrained global optimization tool for Python.* 2014. URL: https://github.com/fmfn/BayesianOptimization.

[19] OpenStreetMap contributors. *Planet dump.* 2017. URL: https://www.openstreetmap.org (visited on 08/28/2022).

[20] Elsa Pastor et al. "Mathematical models and calculation systems for the study of wildland fire behaviour". In: *Progress in Energy and Combustion Science* 29.2 (2003), pp. 139–153. ISSN: 03601285. DOI: 10.1016/S0360-1285(03)00017-0.

[21] Timothy Pede and Giorgos Mountrakis. "An empirical comparison of interpolation methods for MODIS 8-day land surface temperature composites across the conterminous Unites States". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 142 (2018), pp. 137–150. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2018.06.003.

[22] Philippe Rigaux, Michel Scholl, and Agnès Voisard. *Spatial Databases.* 1st edition. Boston, MA: Morgan Kaufmann and Safari, 2001. ISBN: 1558605886.

[23] Petar Ristoski and Heiko Paulheim. "RDF2Vec: RDF Graph Embeddings for Data Mining". In: *The Semantic Web – ISWC 2016.* Ed. by Paul Groth et al. Vol. 9981. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2016, pp. 498–514. DOI: 10.1007/978-3-319-46523-4_30.

[24] Jasper Snoek, Hugo Larochelle, and Ryan P. Adams. "Practical Bayesian Optimization of Machine Learning Algorithms". In: *Advances in neural information processing systems 25.* Ed. by P. Bartlett et al. Vol. 25. Advances in neural information processing systems. Red Hook, NY: Curran Associates, Inc, 2012, pp. 1–9. ISBN: 9781627480031.

[25] State Climate Extremes Committee. *Records.* 2022. URL: https://www.ncdc.noaa.gov/extremes/scec/records/ca/all (visited on 08/28/2022).

[26] Nicolas Tempelmeier and Elena Demidova. "Linking OpenStreetMap with knowledge graphs — Link discovery for schema-agnostic volunteered geographic information". In: *Future Generation Computer Systems* 116 (2021), pp. 349–364. ISSN: 0167739X. DOI: 10.1016/j.future.2020.11.003.

[27] Waldo R. Tobler. "A Computer Movie Simulating Urban Growth in the Detroit Region". In: *Economic Geography* 46 (1970), pp. 234–240. ISSN: 00130095. DOI: 10.2307/143141.

[28] Lu Wang and Tinghua Ai. "The comparison of drainage network extraction between square and hexagonal grid-based dem". In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XLII-4 (2018), pp. 687–692. DOI: 10.5194/isprs-archives-XLII-4-687-2018.

[29] Quan Wang et al. "Knowledge Graph Embedding: A Survey of Approaches and Applications". In: *IEEE Transactions on Knowledge and Data Engineering* 29.12 (2017), pp. 2724–2743. ISSN: 1041-4347. DOI: 10.1109/TKDE.2017.2754499.

[30] Jiantao Wu et al. "LinkClimate: An interoperable knowledge graph platform for climate data". In: *Computers & Geosciences* 169 (2022). ISSN: 00983004. DOI: 10.1016/j.cageo.2022.105215.

[31] Limin Yang et al. "A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies". In: *ISPRS Journal of Photogrammetry and Remote Sensing* 146 (2018), pp. 108–123. ISSN: 09242716. DOI: 10.1016/j.isprsjprs.2018.09.006.

A. Online Resources

The sources for the paper are available via the following GitHub repository: https://github.com/MartinBoeckling/WildfirePredictionSTKG.