UF Herbert Wertheim College of Engineering
UNIVERSITY of FLORIDA

Developing AI-based Wildfire Evacuation Behavior (AI-WEB) Model

Xilei Zhao, Ph.D.
Department of Civil and Coastal Engineering, University of Florida

POWERING THE NEW ENGINEER TO TRANSFORM THE FUTURE

Our research team

Xilei Zhao (PI)

Erica Kuligowski
(Co-PI)

Rino Lovreglio
(Co-PI)

Daniel Nilsson
(Co-PI)

Tom Cova
(Collaborator)

Xiang ‘Jacob’ Yan
(Collaborator)

Yiming Xu
(Ph.D. Student)

Yepeng Liu
(M.S. Student)

Ruoyang Xiong
(M.S. Student)

Overview

- Motivation
- Goals
- Analyzing wildfire evacuation decisions and departure timing with GPS data
- Forecasting real-time travel demand during wildfire evacuations
  - *Situational-Aware Multi-Graph Convolutional Recurrent Network (SA-MGCRN)*
- Key take-aways

Motivation

Insider (2019)

Kent Porter / The Press Democrat (2019)

Research Needs:

- Understand household behavior and movements in wildfires;
- Provide real-time decision support for emergency managers.

Goals

Goal 1:
Improve understanding of people's evacuation decision-making using large-scale GPS data.

Goal 2:
Advance methodology of forecasting real-time travel demand during wildfire evacuations.

Analyzing wildfire evacuation decisions and departure timing with GPS data

Zhao, X., Xu, Y., Lovreglio, R., Kuligowski, E., Nilsson, D., Cova, T., Wu, A., & Yan, X. (2022). Estimating wildfire evacuation decision and departure timing using large-scale GPS data. Transportation Research Part D: Transport and Environment, 107, 103277.

Study site exploration

2019 Kincade fire, Sonoma County, CA:

- Started at 9:27 pm on October 23, 2019 and was fully contained at 7:00 pm on November 6, 2019.

- Burned 77,758 acres, destroyed 374 structures, damaged 60 structures, and caused 4 injuries.

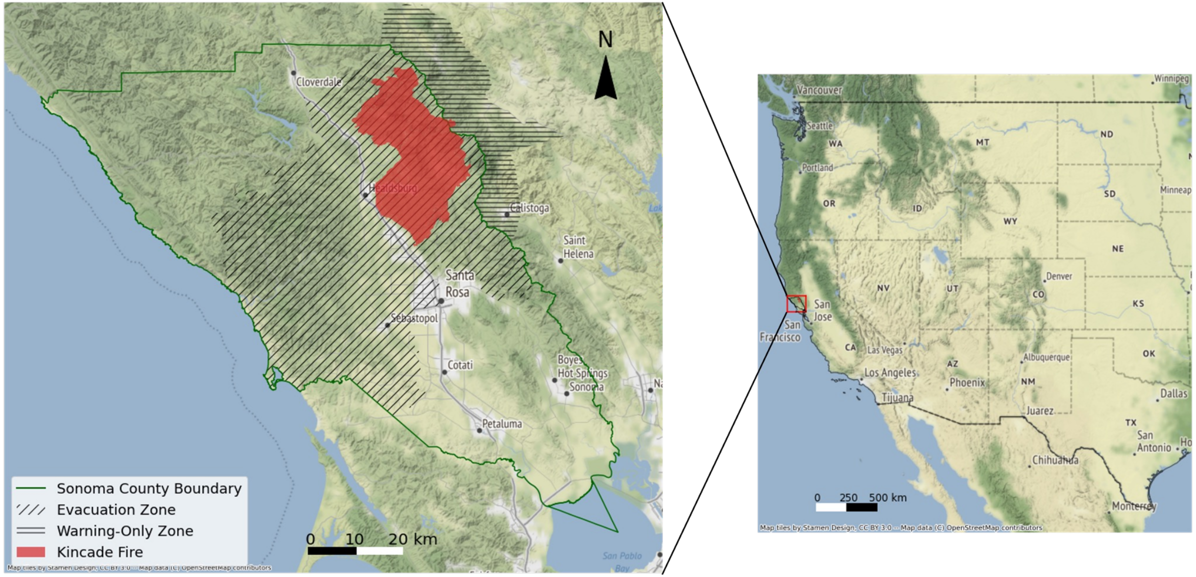


Figure. Sonoma County and the Kincade Fire perimeter

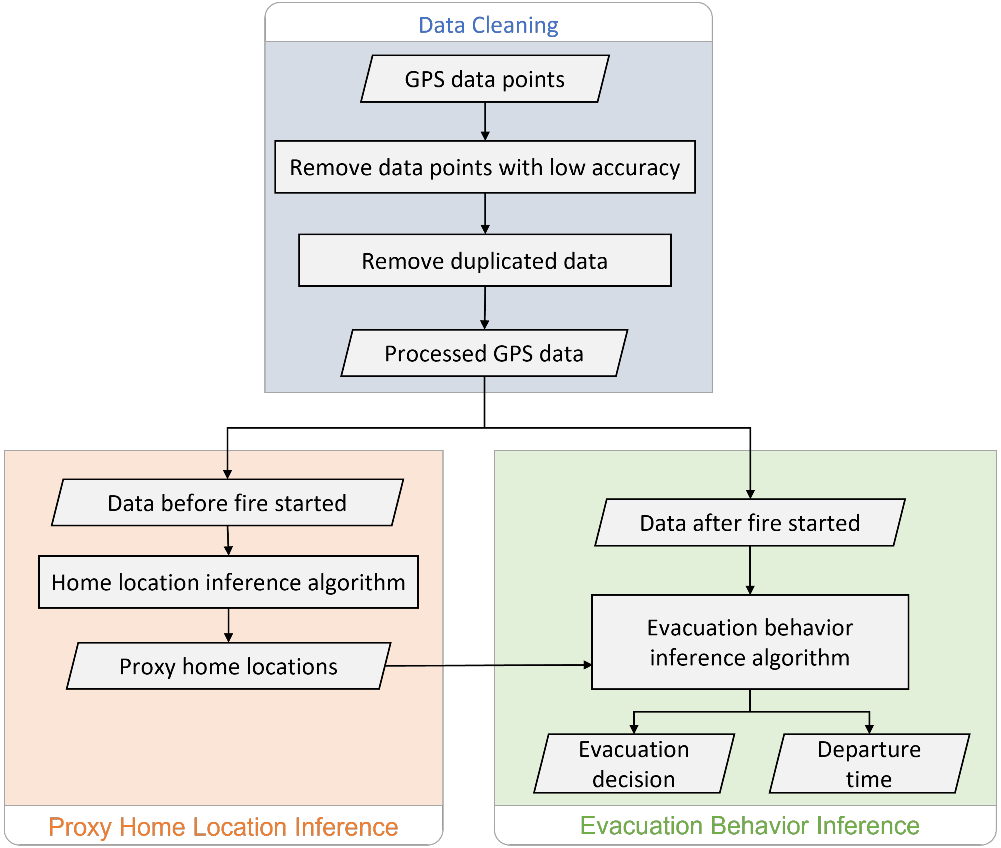
Data description & cleaning

- The GPS data was provided by Gravy Analytics and built on privacy-friendly mobile location data.
- After the data cleaning process, we retained 44,211,050 records, or a total of 5,338 residents for analysis.

Table 1. Synthetic GPS Data Samples

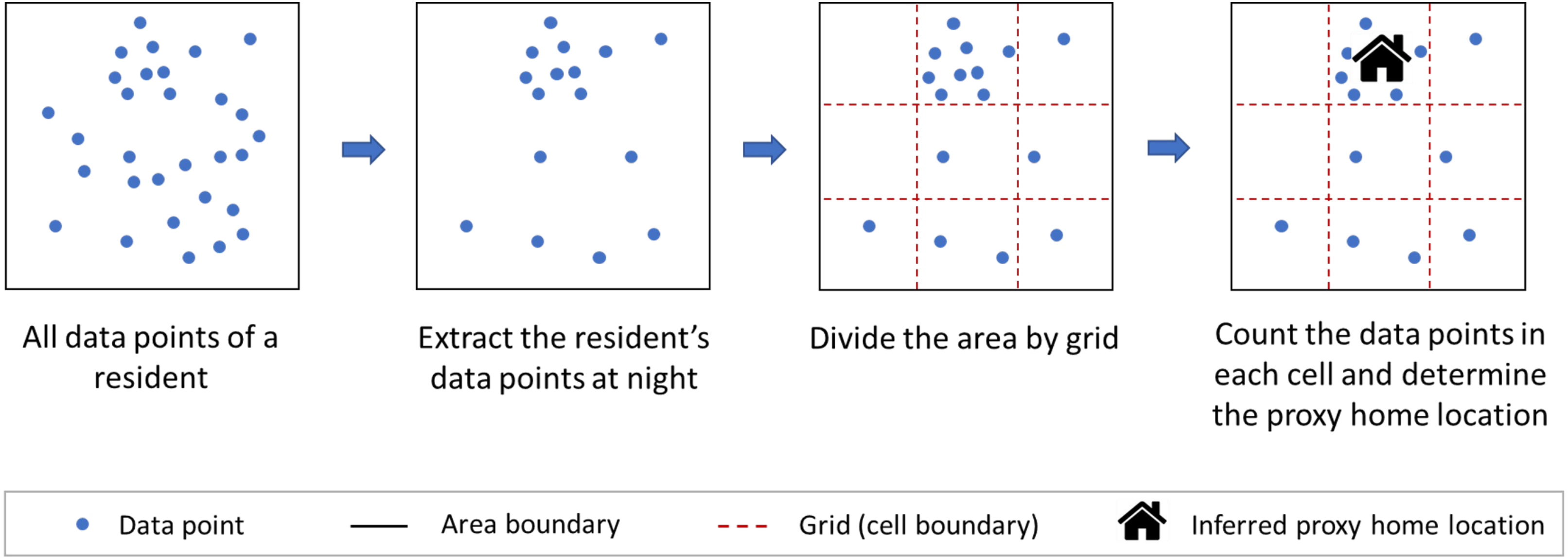
ID	LATITUDE	LONGITUDE	GEOHASH9	TIMESTAMP_EPOCH	TIMEZONE	FLAG
00001	y_1	x_1	9qbd*****	15715********	TZ1	0
00002	y_2	x_2	9qbc*****	15715********	TZ1	0
00003	y_3	x_3	9qbs*****	15712********	TZ1	0
00003	y_4	x_4	9qbe*****	15726********	TZ1	0
00004	y_5	x_5	9qbd*****	15713********	TZ1	0
00004	y_6	x_6	9qbd*****	15714********	TZ1	0

Methodological framework



Proxy-home-location inference

Apply time-space heuristics method accompanied by clustering.



Evacuation-behavior inference

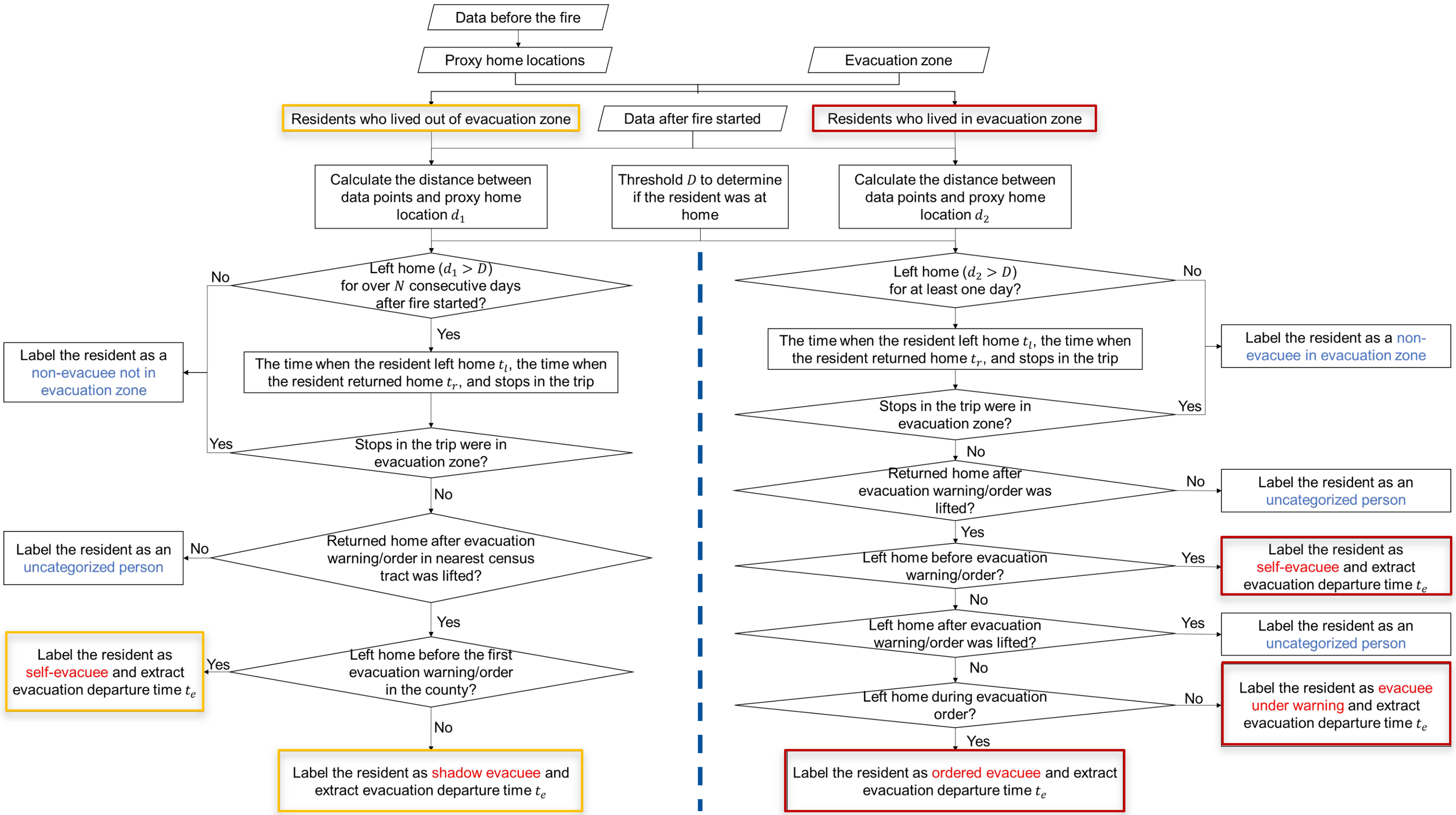
Note that we only analyze the evacuation behavior of people who resided in or near the evacuation zones (within 5 miles of the evacuation zones' boundaries).

- **Assumption 1**: All evacuees departed from home.
- **Assumption 2**: If the distance between the resident's current location and the resident's proxy home location was larger than D, the resident has left home.
- **Assumption 3**: A resident is considered as an evacuee, if they left the evacuation zone during the evacuation process.
- **Assumption 4**: The evacuation departure time is when the evacuee left home to evacuate.

Evacuation-behavior inference

Definitions of evacuee groups:

- ***Self-evacuee***: The evacuee, located in or near the evacuation zone, left after the fire started but before any evacuation warning/order was issued.
- ***Shadow evacuee***: The evacuee, located outside but near the evacuation zone, left after an evacuation warning/order was issued.
- ***Evacuee under warning***: The evacuee was in the evacuation warning zone and evacuated after the warning was issued and before an order was issued (if any).
- ***Ordered evacuee***: The evacuee lived in the evacuation order zone and evacuated after the order was issued.


Data before the fire
Proxy home locations
Evacuation zone
Residents who lived out of evacuation zone
Data after fire started
Residents who lived in evacuation zone
Calculate the distance between data points and proxy home location d_1
Threshold D to determine if the resident was at home
Calculate the distance between data points and proxy home location d_2
Left home ($d_1 > D$) for over N consecutive days after fire started?
No
Yes
Label the resident as a non-evacuee not in evacuation zone
The time when the resident left home t_l, the time when the resident returned home t_r, and stops in the trip
Stops in the trip were in evacuation zone?
Yes
No
Returned home after evacuation warning/order in nearest census tract was lifted?
No
Label the resident as an uncategorized person
Yes
Left home before the first evacuation warning/order in the county?
Yes
Label the resident as self-evacuee and extract evacuation departure time t_e
No
Label the resident as shadow evacuee and extract evacuation departure time t_e
Left home ($d_2 > D$) for at least one day?
No
Label the resident as a non-evacuee in evacuation zone
The time when the resident left home t_l, the time when the resident returned home t_r, and stops in the trip
Stops in the trip were in evacuation zone?
Yes
No
Returned home after evacuation warning/order was lifted?
No
Label the resident as an uncategorized person
Yes
Left home before evacuation warning/order?
Yes
Label the resident as self-evacuee and extract evacuation departure time t_e
No
Left home after evacuation warning/order was lifted?
Yes
Label the resident as an uncategorized person
No
Left home during evacuation order?
No
Label the resident as evacuee under warning and extract evacuation departure time t_e
Yes
Label the resident as ordered evacuee and extract evacuation departure time t_e

An example to illustrate the algorithm

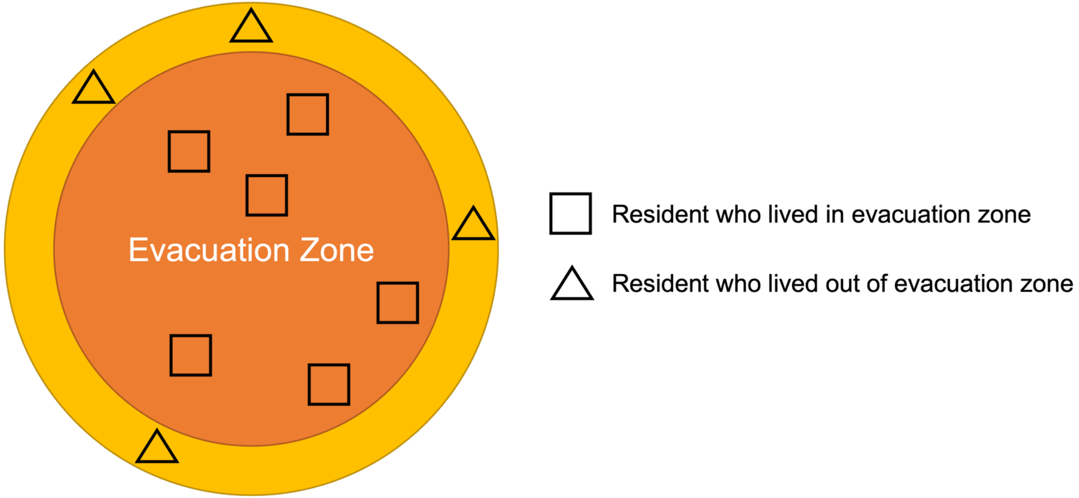
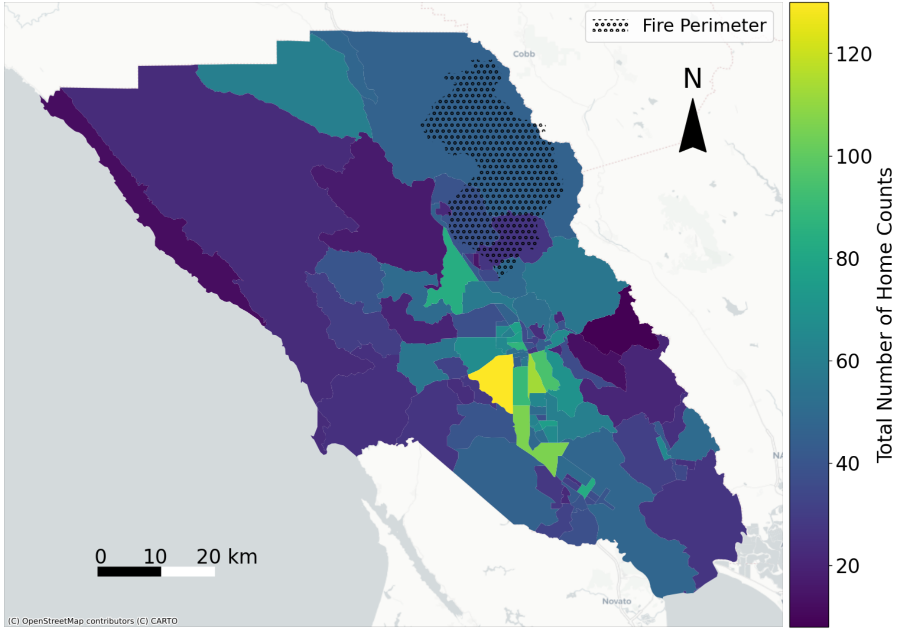
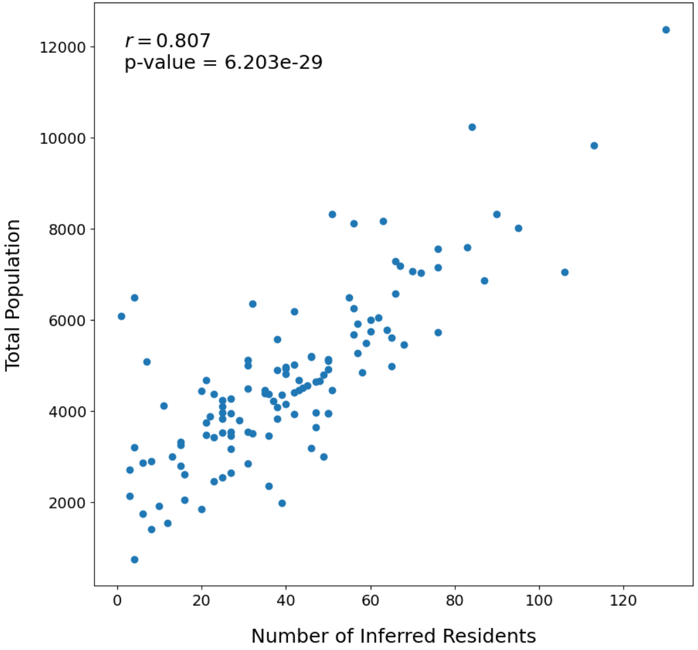


Table 1: Definitions of Different Evacuee Groups

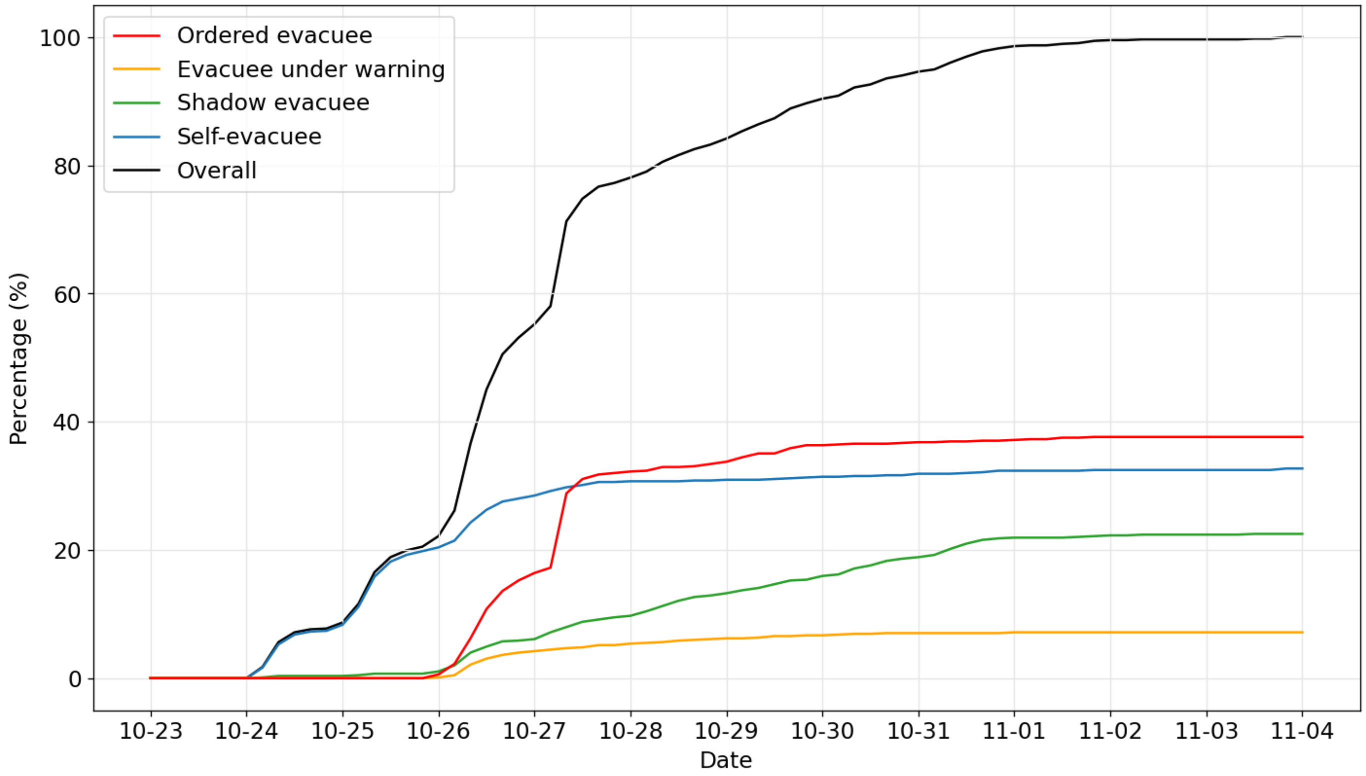
	Day 1	**Day 2**	**Day 3: Warning**	**Day 4**	**Day 5: Evacuation order**
□	Self-evacuee	Self-evacuee	Evacuee under warning	Evacuee under warning	Ordered evacuee
△	Self-evacuee	Self-evacuee	Shadow evacuee	Shadow evacuee	Shadow evacuee

Results: Home location inference

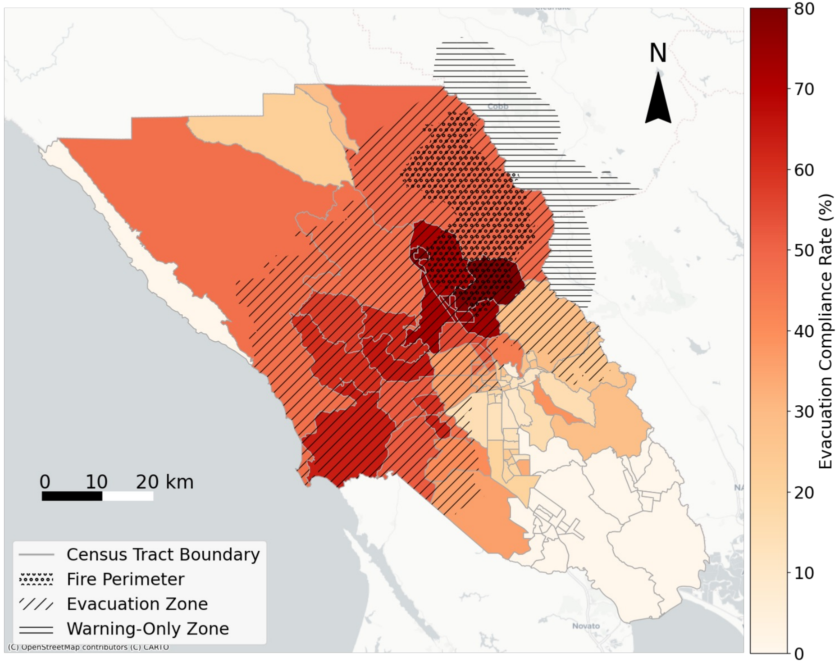




Results: Temporal patterns



Results: Spatial patterns



GPS data: 46% evacuated v.s. Survey data: 80% evacuated

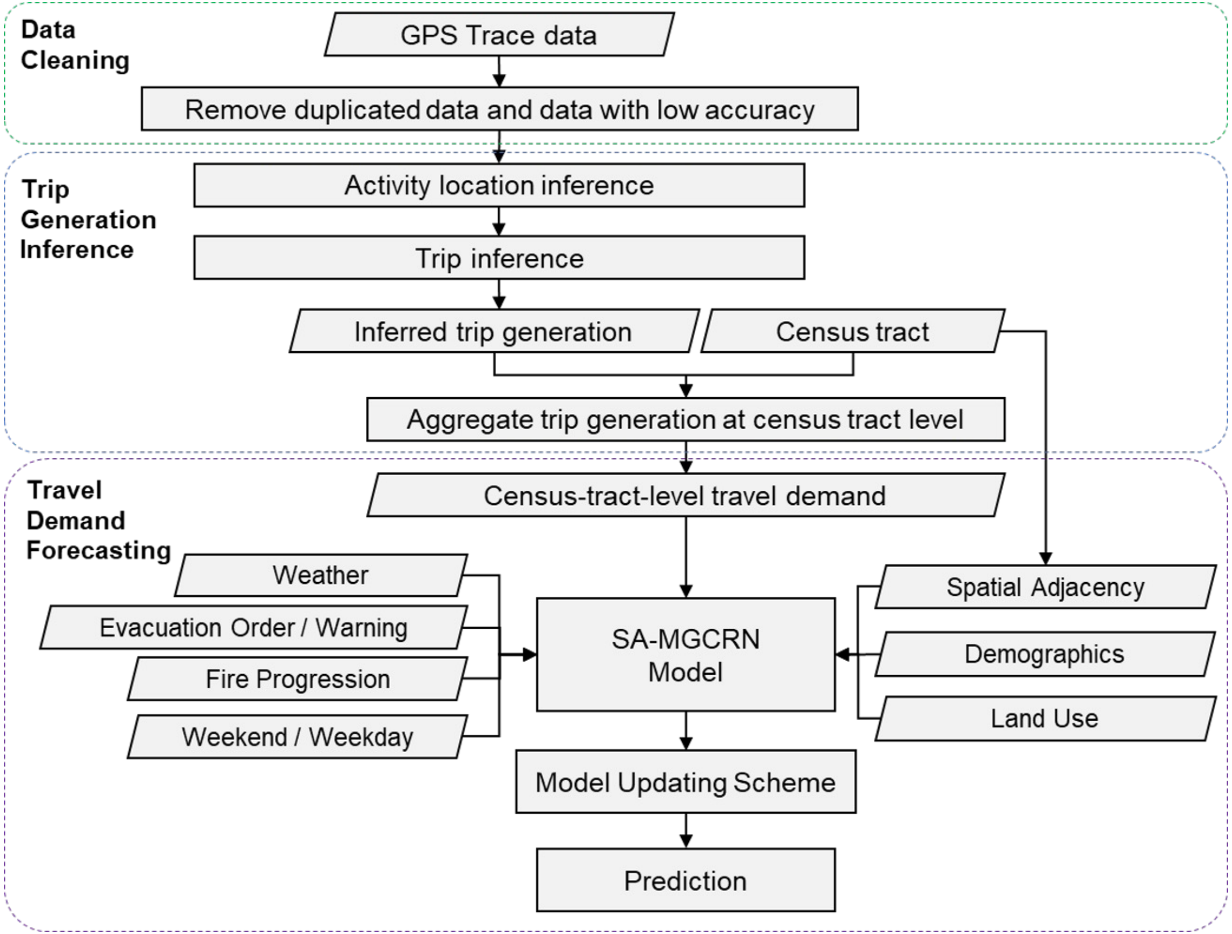
Key take-aways

- A set of novel methodologies are developed to systematically analyze wildfire evacuation process and identify different groups of evacuees.
- Self-evacuees and shadow evacuees consisted of more than half of evacuees during the Kincade Fire.
- The total evacuation compliance rate is around 50%, which shows some discrepancy from the results obtained from the separate survey study for the same fire conducted by our team (Kuligowski et al., 2022).

Kuligowski, E. D., Zhao, X., Lovreglio, R., Xu, N., Yang, K., Westbury, A., ... & Brown, N. (2022). Modeling evacuation decisions in the 2019 Kincade fire in California. *Safety Science, 146,* 105541.

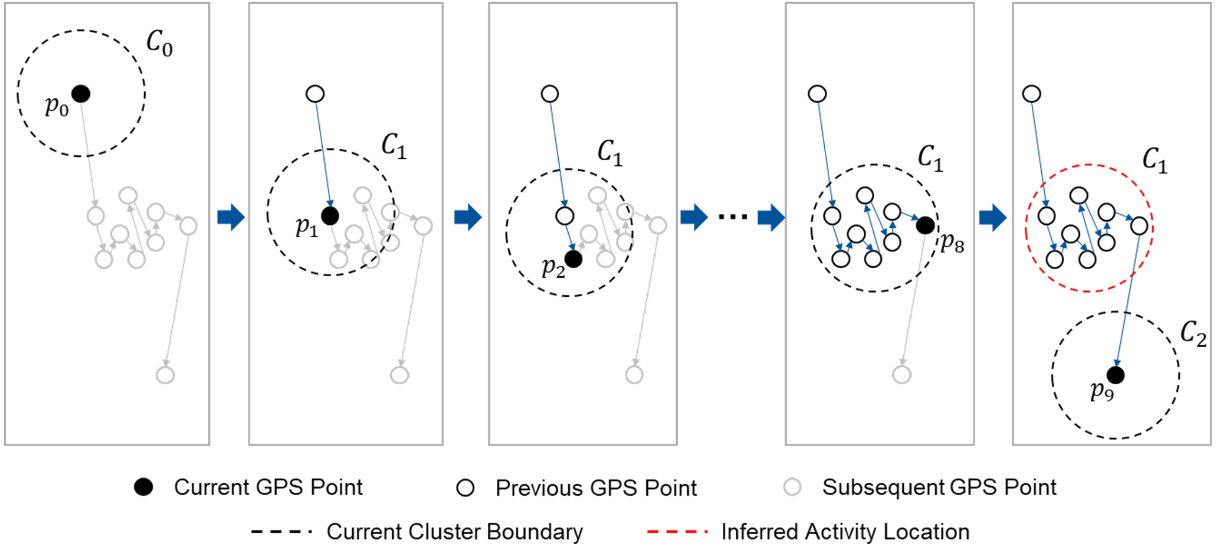
Forecasting real-time travel demand during wildfire evacuations

Xu, Y., Xiong, R., Lovreglio, R., Nilsson, D., & Zhao, X. (In Preparation). Forecasting real-time travel demand during wildfire evacuations: A Situational-Aware Multi-Graph Convolutional Recurrent Network (SA-MGCRN) approach. In *Proceedings of Transportation Research Board 102nd Annual Meeting,* Washington, D.C.



evacuation trips + other types of trips (e.g., background trips, intermediate trips)

Trip generation inference – Incremental clustering

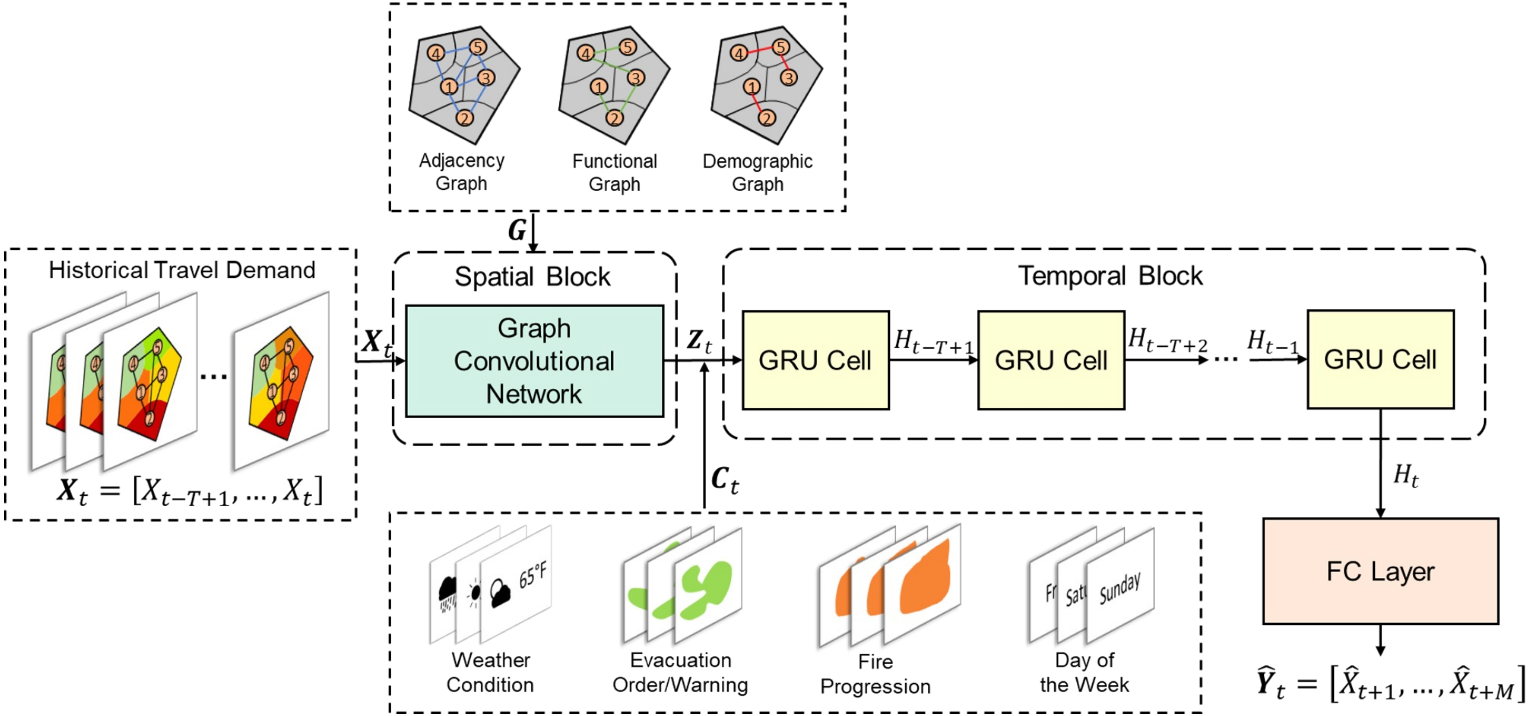


Cluster radius: 500 meters
Time threshold: 5 minutes

1. Wang, F., & Chen, C. (2018). On data processing required to derive mobility patterns from passively-generated mobile phone data. Transportation Research Part C: Emerging Technologies, 87, 58-74.
2. Wang, F., Wang, J., Cao, J., Chen, C., & Ban, X. J. (2019). Extracting trips from multi-sourced data for mobility pattern analysis: An app-based data example. *Transportation Research Part C: Emerging Technologies*, *105*, 183-202.
3. Chen, C., Bian, L., & Ma, J. (2014). From traces to trajectories: How well can we guess activity locations from mobile phone traces?. Transportation Research Part C: Emerging Technologies, 46, 326-337.
4. Calabrese, F., Diao, M., Di Lorenzo, G., Ferreira Jr, J., & Ratti, C. (2013). Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. Transportation research part C: emerging technologies, 26, 301-313.

Architecture of the SA-MGCRN model

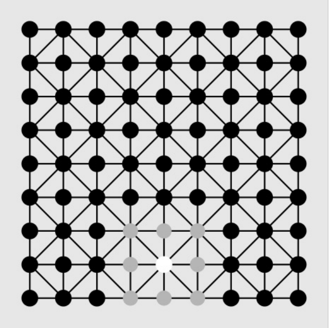
Situational-Aware Multi-Graph Convolutional Recurrent Network (SA-MGCRN)



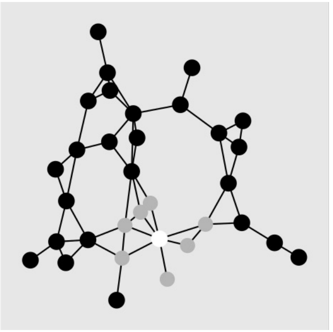
Graph Convolutional Network (GCN)

Why GCN? GCN v.s. CNN

- CNN can only be performed in Euclidean space, while GCN can handle graph-structured data.
- The census tracts do not have a regular spatial structure but can be represented by a graph.



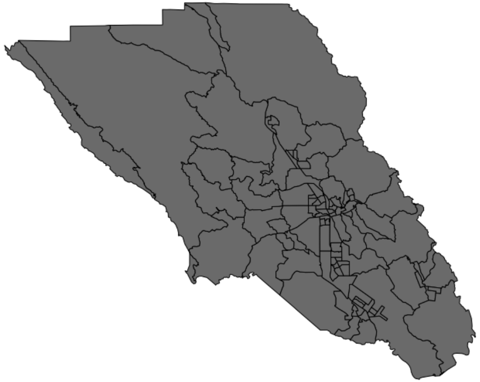
Euclidean data



Graph-structured data

How does GCN work?

- GCN performs convolutional operation using a filter.
- The filter is applied on each node of the graph, thus capturing spatial dependency between a node and its adjacent nodes.



Census Tracts (Sonoma County, CA)

Gated Recurrent Unit (GRU)

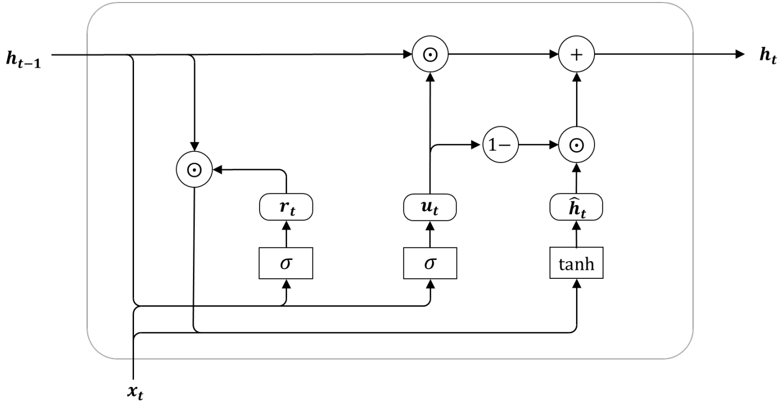
Why GRU? GRU v.s. LSTM

- Widely-used RNN model to capture temporal dependency.
- GRU is faster to compute but still offers comparable performance in prediction compared with LSTM.
- Using the memorize long-term information, GRU can deal with the vanishing gradient problem.

How does GRU work?

- GRU uses two gates to determine what information should be kept and passed to the output.
- Reset gate: how much of the previous state information to remember.
- Update gate: how much of the previous information to pass to the new state.

Structure of a GRU unit



x_t -- input at time t;
h_{t-1}, h_t -- hidden states;
r_t -- reset gate;
u_t -- update gate;
$\hat{h}_t$ -- candidate hidden state.

Model updating scheme

Case study: Kincade fire

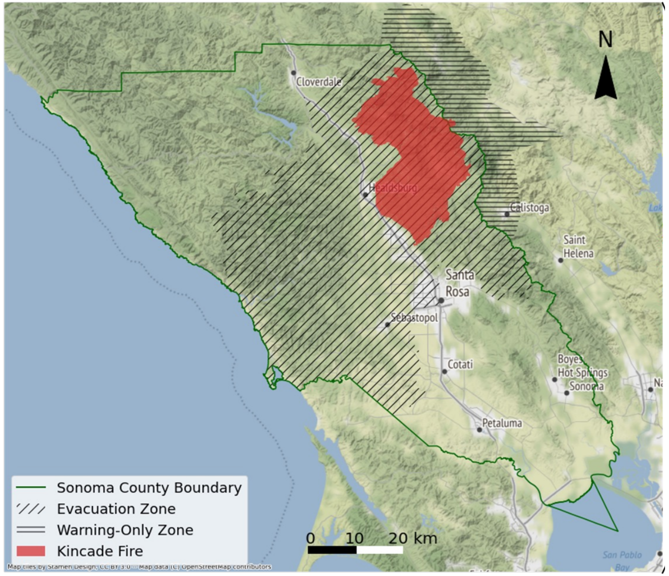


TABLE 1: Descriptive statistics of input variables

Variables	Mean	Std.	Min	Max	Category
Proportion of residential areas	38.15%	16.46%	5.23%	100%	Functional similarity graph
Proportion of commercial areas	22.40%	19.25%	1.82%	100%	
Proportion of agricultural areas	7.40%	16.32%	0.0%	100%	
Proportion of multi-family areas	26.74%	25.30%	0.0%	100%	
Population density (per sq. mile)	3339.79	3192.12	7.16	12474.63	Demographic similarity graph
Proportion of the young population	27.01%	7.68%	12.09%	54.98%	
Proportion of the white population	76.58%	12.90%	38.17%	95.54%	
Proportion of population with BA's degree and above	36.34%	12.65%	12.09%	63.62%	
Median household income (US dollar)	83823.67	20522.56	49856.0	145147.0	
Proportion of households own 0 car	7.46%	6.67%	0%	32.02%	
Proportion of households own 1 cars	36.77%	11.52%	3.65%	62.94%	
Proportion of households own 2 cars	35.60%	11.41%	10.66%	88.03%	
Proportion of households own 3 cars	13.86%	7.63%	0%	36.80%	
Employment rate	95.65%	2.08%	88.96%	99.45%	
Fire distance	185.65	14.27	176.46	323.5	Temporal Variables
Evacuation order/warning	0.10	0.29	0	1	
Day of the week	0.30	0.46	0	1	
Temperature	56.91	12.63	33.4	90.6	
Feels like temperature	56.69	12.41	33.4	86.4	
Wind speed	4.25	4.49	0.0	29.6	
Sea level pressure	1016.57	3.27	1005.9	1023.9	
Humidity	58.76	28.99	8.92	100	
Visibility	8.79	2.38	0.0	9.9	
Cloud cover	17.45	27.22	0.0	100	

Results: Trip generation inference

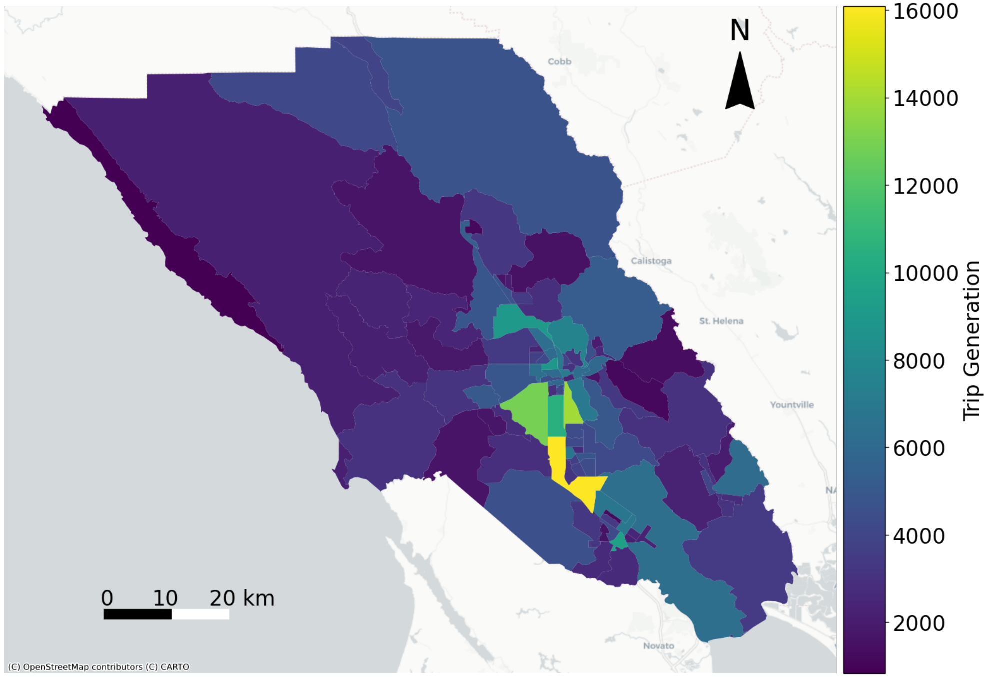


Figure. Distribution of total trip generation in Sonoma County (census-tract level)

The mean of hourly travel demand for each census tract is 6.14, the standard deviation is 6.07, the maximum value is 56, and the minimum value is 0.

Results: Model comparison

Performance Metrics

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|,$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2},$$

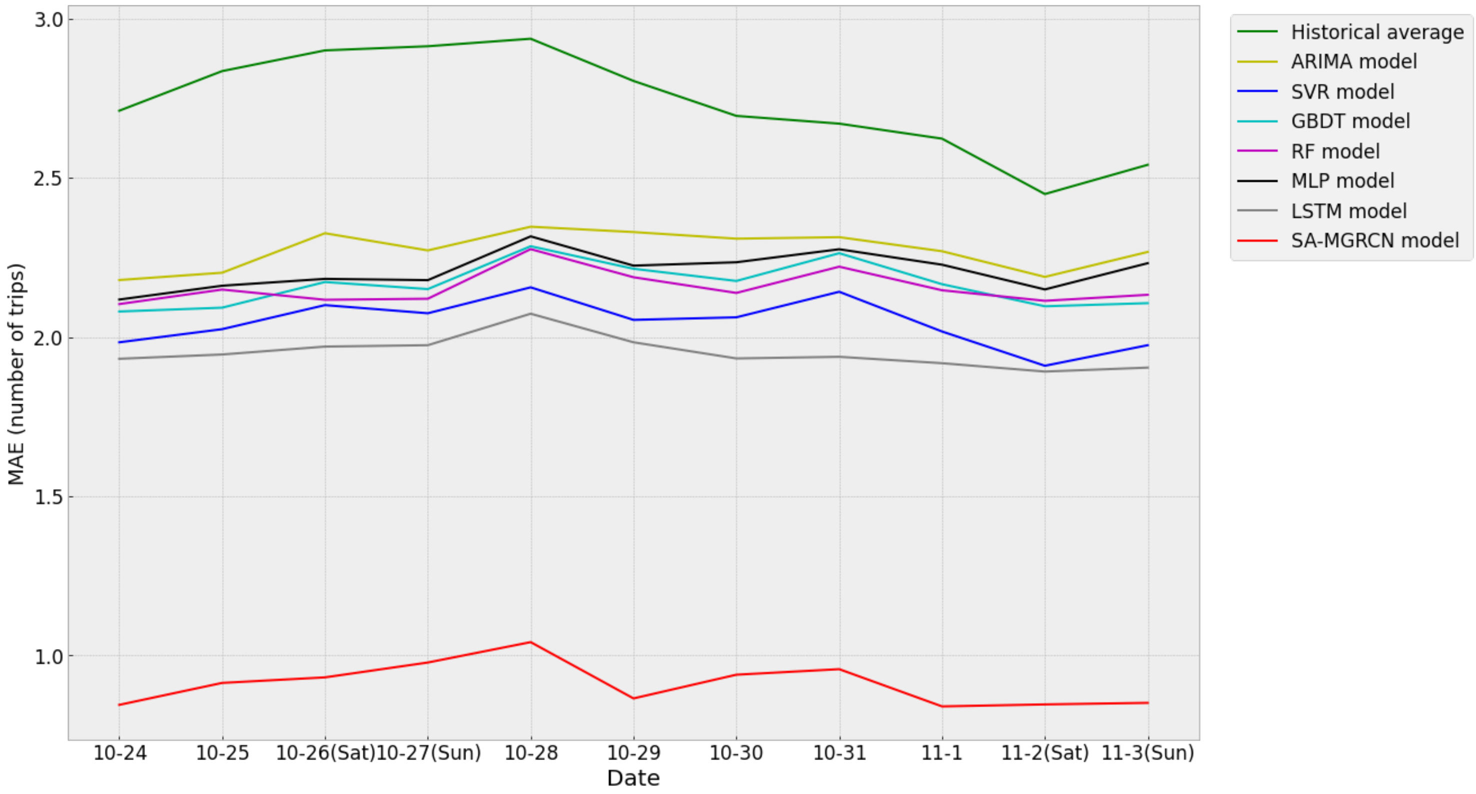
$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right|,$$

where y_i is the observed value and $\hat{y}_i$ is the predicted value.

TABLE 2: Prediction performance of SA-MGCRN and benchmark models

Methods	MAE	RMSE	MAPE	
HA	2.7347	3.5217	57.89%	Statistical models
ARIMA	2.2732	2.6110	48.12%	
SVR	2.0455	2.6441	43.48%	Classical machine learning models
GBDT	2.1641	2.3093	45.77%	
RF	2.1553	2.8008	45.63%	
MLP	2.2092	2.7356	46.77%	Deep learning models
LSTM	1.9512	2.1262	41.27%	
SA-MGCRN	**0.9095**	**1.1224**	**20.13%**	

Results: Model comparison



Results: Ablation study

The ablation study examines the performance of the model by removing certain components to see the contribution of the removed components.

TABLE 3: Results of ablation study

Methods	MAE	RMSE	MAPE
SA-MGCRN	**0.9095**	**1.1224**	**20.13%**
W/O whether the day is weekend	0.9368	1.1644	20.82%
W/O evacuation order/warning information	0.9583	1.2135	21.21%
W/O spatial adjacency	1.0463	1.3356	23.16%
W/O functional similarity	1.0032	1.3098	22.20%
W/O demographic similarity	1.0369	1.3273	22.95%
W/O weather information	1.3781	1.4256	30.50%
W/O fire distance information	1.6133	1.7215	35.71%

Key take-aways

- A new deep learning model, SA-MGCRN, along with a model updating scheme, is developed to accurately forecast real-time travel demand in wildfire evacuations.
- SA-MGCRN can be directly deployed to facilitate real-time emergency management and revolutionize the state-of-the-practice.
- Fire proximity is the most important component of SA-MGCRN. In future work, other fire cues (e.g., smoke and embers) need to be incorporated.

Acknowledgement

This work was performed under the following financial assistance award No. 60NANB20D182 and 60NANB21D180 from U.S. Department of Commerce, National Institute of Standards and Technology (NIST). Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of NIST. The authors would like to thank Dr. Nancy A. Brown, the Emergency Manager of Sonoma County, CA for providing valuable inputs for the study.

NIST
National Institute of Standards and Technology

Thank you for your attention!

Xilei Zhao (xilei.zhao@essie.ufl.edu)

SERMOS LAB