An interdisciplinary data-science approach to managing natural hazards risk

MARTA C. GONZALEZ
CRISTOBAL PAIS
MINHO KIM
JOHN RADKE
UNIVERSITY OF CALIFORNIA, BERKELEY
High-level Flow diagram

Data layers

Optimization model
- Decisions
- Constraints
- Objective

Outputs & analysis
Fire Hazards

Oakland firestorm of 1991
3,280 structures; 6 km²
Losses near 3 billion (in 2020 $s)
October 19–23, 1991;
Caused by an incompletely extinguished grass fire

Camps Fire (Paradise), 2018,
18,804 structures; 621 km², October 19–23, 1991; caused by a
Costs: $16.65 billion
November 8 – November 25, 2018
Caused by PG&E power line failures during high winds. Drought was a factor: Paradise, which typically sees 5” inches of autumn rain by November 12, had only received 1/7” by that date in 2018
Input Layers
Data: Fire Behavior Model by FlamMap

<table>
<thead>
<tr>
<th>Driver</th>
<th>Data Layer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topography</td>
<td>Elevation</td>
</tr>
<tr>
<td></td>
<td>Slope</td>
</tr>
<tr>
<td></td>
<td>Aspect</td>
</tr>
<tr>
<td>Vegetation</td>
<td>Fuel Model</td>
</tr>
<tr>
<td></td>
<td>Canopy Cover</td>
</tr>
<tr>
<td></td>
<td>Canopy Base Height</td>
</tr>
<tr>
<td></td>
<td>Canopy Bulk Density</td>
</tr>
<tr>
<td></td>
<td>Canopy Stand Height</td>
</tr>
<tr>
<td>Weather</td>
<td>Precipitation, temperature, relative density, wind speed, wind direction</td>
</tr>
</tbody>
</table>

[Maps showing elevation, slope, aspect, and fuel model data for California]
Data: Median house values

- **Dataset:** Median house value across California 2020 (source: Zillow).
- **Location:** Multiple sample points from zip level analysis.
- **Total area covered:** State level.
Data: Population

- **Dataset:** Population California (source: Census).
- **Total area covered:** State level. Granularity at a block level.
Data: Street network and expected travel times

- **Dataset**: State level road network. Total of 2,456,671 nodes (including intersections and dead ends) and 6,039,371 edges.

- **Traveling times**: OpenStreetMap and publicly available traveling times from ride sharing companies.

- **Total area covered**: State level.
Data: Wildfire and related location of fire stations (3150) across the state.

CALFIRE: https://wifire-data.sdsc.edu/es_AR/dataset/cal-fire-facilities-for-wildland-fire-protection
Homeland Infrastructure Foundation-Level Data (HIFLD): https://hifld-geoplatform.opendata.arcgis.com/
Tessellation: Resolution and aggregation patterns

- We use **h3 library** from Uber
- **Dynamic and adaptable** tessellation depending on the data resolution
- Effective for **statistic** and comparison: the distance between the centroids of the neighbors is always the same (in contrasts, e.g., to squares, triangles, and other schemes).
**RI: Risk index definition**

\[
RI = 0.5 \times (FB + SD) \times STTFS
\]

RI ∈ [0, 1]

**FB:** Fire behavior depends on the Rate of Spread and Fire intensity

**SD:** Takes as input Population, Median house values

**STTFS:** Shortest travel time from Fire Station
Example:

LA County AS - IS Distribution

Risk index equal

Risk index Fire dom

Risk index Pop dom
Optimal distribution of facility

• Given a city, if we know the distribution of population and road network, what’s the optimal distribution to deploy Fire station facilities?

• Problem statement:
  • Minimize the total travel time of the Fire station capabilities.
  • N blocks in a city with population.
  • k of N blocks are assigned with Fire station capabilities.
  • People choose facility per the free flow travel time from their residential block to facility block.
Optimization model

Objective functions

Minimize maximum traveling times from fire stations to nodes/hexs centroids; taking into account risk index given by demographics and fire behavior

\[(IP_{RI}) \quad \min U = \sum_{i,j \in I} RI_i \times x_{(i,j)} \times t_{(i,j)}\]

\[\text{s.t.} \quad \sum_{i \in I} x_{(i,j)} = 1 \quad \forall j \in I\]

\[x_{(i,j)} \leq y_i \quad \forall i, j \in I\]

\[\sum_{j \in I : j \neq i} x_{(i,j)} \geq y_i \quad \forall i \in I\]

\[\sum_{i \in I} y_i = S\]

\[x_{(i,j)}, y_i \in \{0, 1\} \quad \forall i, j \in I\]

where \(I\) := set of potential locations (i.e., hexagons) connected to the street network; \(S\) := the total number of existing stations; \(t_{(i,j)}\) := the shortest traveling time between \(i\) and \(j\); \(RI_i\) := the value of the outcome variable of interest at hexagon \(i\); and the decision variables:

\[x_{(i,j)} = 1 \text{ if station in location } i \text{ covers } j\]

\[y_i = 1 \text{ if a station is located in } i\]
Example: LA County optimization
California: Optimizing RI

Δ improvement % points (+ is better)

Δ improvement % distribution

Δ% intervals

# Hexagons
A core-based statistical area (CBSA) is a U.S. geographic area defined by the Office of Management and Budget (OMB) that consists of one or more counties (or equivalents) anchored by an urban center of at least 10,000 people plus adjacent counties that are socioeconomically tied to the urban center by commuting.
What can we generalize about an interdisciplinary data-science approach to managing natural hazards risk?
Optimal deployment of Fire station Facilities

Average distance $L$ vs. number of facilities

$$\lambda = \frac{1.833}{N_{occ}}$$

$N_{occ}$: number of blocks with population above 500 (urban area)

$N$: number of facilities

Fire station density versus population density in the optimal scenario.

Decrease in Risk Index % by adding facilities

$D$: Facility Density

$\rho$: Population Density

Deconstructing laws of accessibility and facility distribution in cities

Y Xu, LE Olmos, S Abbar, MC González Science advances (2020)
Environmental and Decision Sciences allowed us to target vulnerable locations at state level.

When defining a compound risk index that includes environmental behavior the universalities of optimal distribution of facilities disappear.

It is a must to combine disciplines in risk mitigation strategies.
Towards Resilient Critical Infrastructures: Understanding the Impact of Coastal Flooding on the Fuel Transportation Network in the San Francisco Bay

Y He, S Lindbergh, Y Ju, M Gonzalez, J Radke

ISPRS International Journal of Geo-Information (2021)
Research Question

How will **coastal flooding** impact **fuel transportation networks** under future **climate change** scenarios?

**STEP 01**
- The definition of **fuel transportation networks**
- Build a **network model** to represent the network
- **Network properties** and characteristics

**STEP 02**
- The definition of **coastal flooding**
- **Scenarios** of coastal flooding under future **climate change** (GCM, RCP, SLR percentiles, time horizons)

**STEP 03**
- **Regional impact** analysis of network properties
- **Local impact** analysis focusing on cascading effects and routing simulations
| STEP 01 |
Fuel Transportation Networks
120 Coastal Flooding Scenarios
General Circulation Models (GCMs) represent physical processes in the atmosphere, ocean, cryosphere and land surface, are the most advanced tools currently available for simulating the response of the global climate system to increasing greenhouse gas concentrations.

Representative Concentration Pathways (RCPs) are four greenhouse gas concentration (not emissions) trajectories adopted by the IPCC for its fifth Assessment Report (AR5) in 2014.
STEP 02
120 Coastal Flooding Scenarios
Note: for each time period, the flooding scenario with RCP 4.5, MICRO5 GCM and 95th percentile SLR is selected.
Impact Analysis: Regional

STEP 03

- RCP 4.5, 50th Percentile SLR
- RCP 4.5, 95th Percentile SLR
- RCP 4.5, 99.9th Percentile SLR
- RCP 8.5, 50th Percentile SLR
- RCP 8.5, 95th Percentile SLR
- RCP 8.5, 99.9th Percentile SLR

- Median inundation scenario under all RCP 4.5 scenarios
- Median inundation scenario under all RCP 8.5 scenarios

Number of Nodes vs Time Horizons (2000 - 2100)
Impact Analysis: Regional
| STEP 03 |

Impact Analysis: Regional

Before ‘Attack’

After ‘Attack’/Inundation

Number of Sub-networks

Time Horizons

- 2000 - 2020
- 2020 - 2040
- 2040 - 2060
- 2060 - 2080
- 2080 - 2100
$$GE = \frac{1}{N(N - 1)} \sum_{s \neq t}^{n} \frac{1}{Z_{st}}$$

Latora and Machori proposed Global Efficiency (GE) as a measure of the exchange of information within a network.

Osei-Asamoah et al. explained that GE quantifies how flow is exchanged between nodes in a transportation network.

- $Z_{st}$ is the length of the shortest path between node $s$ and node $t$
- $N$ is the total number of nodes in the network
- The GE value is normalized by dividing by the GE of an ideal network where all node pairs are connected
Conclusions

• The direct impact of coastal flooding on fuel transportation networks increases overtime across different climate scenarios. The impact under RCP 8.5 scenarios is larger than RCP 4.5 scenarios.

• The multimodal network is likely to become fragmented towards the end of century, breaking down into smaller sub-networks. The efficiency within the network will decrease as well.

• When considering cascading effects within the network, the real impact of coastal flooding will be larger. Some smaller hubs within the network could cause a bigger ripple effect than some of the biggest hubs.
We introduce 17 toy cities with different UCI (Ref. Pereira 2013), and another six real-word cities (Paris, Barcelona, London, Dublin, Mexico City, and Melbourne).

Deconstructing laws of facility distribution in cities

City-customized function

\[ L(N) = l_{\text{min}} \cdot (1 - e^{-\alpha N}) + A \cdot N^{-\lambda} \cdot e^{-\alpha N}, \]

Universal function

\[ L(N) = l_{\text{min}} \cdot (1 - e^{-\alpha N}) + 1.4443 \cdot (\alpha N)^{-\bar{\lambda}} \cdot e^{-\alpha N}, \]

We find collapse of the L among diverse cities by rescaling N with the urban area \( N_{occ} \).

The collapse may give us insights on the universal laws for facility planning in diverse cities. \( \bar{\lambda} = 0.382 \quad l_{\text{min}} = 0.5 \text{ km}. \)

\( E \) and \( F \) show \( L(D_{occ}) \quad D_{occ} = \frac{N}{N_{occ}} \)

\( N_{occ} \): number of blocks with population above 500 (urban area)

\( N \): number of facilities
Example: Actual vs Optimal Distribution

Gain index of block: \( r_i = \frac{\text{Travel distance to facility in optimal scenario}}{\text{Travel distance to facility in actual scenario}} \)

- \( r_i > 1 \): benefit from planning
- \( r_i < 1 \): affected by planning

A: Service communities and population in the actual scenario
B: Service communities and population in the optimal scenario
C: Gain index of each block

\[ p_j^S \]