Heatwave Risk Analysis of Maharashtra: A Geospatial Approach

ES 216: Term Project Report

Shaikh Adnan Matiyoddin (23b4207)

Anika Saanvi Sahoo (23b4227)

Devanshi Khandelwal (23b4233)

Group: Group 2

Supervisor: Prof. Srinidhi Balasubramanian and TAs

November 5, 2025

Abstract

Heatwaves pose a growing threat to public health in India. This project presents a high-resolution (1km) geospatial heatwave risk assessment for Maharashtra, a state uniquely vulnerable due to its dense, humid coastal population. Using Google Earth Engine, the analysis follows a standard **Risk = Hazard** × **Exposure** × **Vulnerability** framework.

The **Hazard** component was modeled using a health-based Heat Index (HI) from 11km ERA5-Land data, interpolated to 1km. **Exposure** was mapped using 1km GPWv4 population density, and **Vulnerability** was a composite index of MODIS NDVI (vegetation) and Nighttime LST (Urban Heat Island proxy).

The results show that risk is highly concentrated. The final risk model identifies a critical hotspot in the **Mumbai Metropolitan Region (MMR)**, where all three risk components peak simultaneously. A district-level ranking confirms this, placing **Mumbai Suburban**, **Mumbai City**, and **Thane** as the top three most at-risk districts, providing actionable data for policy.

Environmental Science and Engineering Department

Indian Institute of Technology Bombay

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Background

Heatwaves, or prolonged spells of extreme heat, are becoming increasingly frequent and widespread due to **global warming**. These **extreme weather events** pose a serious threat to human well-being, affecting health, livelihoods, and critical infrastructure. While generally defined as periods of exceptionally high temperatures lasting at least two to three days, the specific criteria can vary.

In the Indian context, the **Indian Meteorological Department (IMD)** defines a heatwave based on specific thresholds. For the Plains, a heatwave is declared when the maximum temperature reaches at least 40°C and is 4.5°C to 6.4°C above the normal temperature for that location. Critically for this study, a separate criterion exists for **coastal stations**, where a heatwave can be declared if the temperature departure is 4.5°C or more, provided the actual maximum temperature reaches 37°C or more.

The increasing frequency of these events is directly linked to **human-induced climate change**. The Sixth Assessment Report by the **Intergovernmental Panel on Climate Change (IPCC)** highlights that as global temperatures rise, heatwaves will become even more frequent and severe. This is particularly evident in urban environments due to the **Urban Heat Island effect**, where buildings and concrete absorb and slowly release heat, keeping temperatures elevated even at night.

The risk for India is particularly acute. A 2022 World Bank report warned that India could become one of the first places in the world to experience heatwaves that break the **human survivability limit**. This rising heat stress puts a massive portion of the population and economy at risk. Up to 75% of India's workforce, or 380 million people, depend on **heat-exposed labor**. By 2030, India may account for 34 million of the projected 80 million global job losses from heat-stress-related productivity decline, putting up to 4.5% of India's GDP at risk by the end of the decade.

1.2 Problem Statement

While India faces a nationwide threat from rising temperatures, the state of **Maharashtra** presents a unique and critical case study. The risk here is amplified by a specific combination of geographic and demographic factors that are not uniform across the country.

First, the state is home to one of the world's largest urban agglomerations: the **Mumbai Metropolitan Region (MMR)**. This coastal area concentrates an immense population of over 25 million people into a very dense space, creating an unparalleled level of **Exposure** to any climate hazard.

Second, this dense population lives along the long, **humid Konkan coast**. This geography creates a distinct **"humid heat" risk**, which is often more dangerous to human health than the "dry heat" found inland. This factor alone justifies a specialized risk model that goes beyond simple temperature to include humidity, aligning with the IMD's own criteria for coastal stations.

Finally, the selection of Maharashtra as the study area is motivated by its direct **local relevance**. As researchers based at IIT Bombay, we possess a first-hand understanding of the regional climate and social context. This proximity to the subject reinforces the practical importance of developing a targeted risk assessment.

Despite this clear combination of **high exposure** and **unique hazard**, a granular, **district-level risk map** that combines these factors is lacking. This project directly addresses that gap by creating a targeted risk index to identify where and why specific areas within Maharashtra are most vulnerable.

1.3 Aim and Objectives

The primary **Aim** of this project is *"To conduct a geospatial heatwave risk assessment for Maharashtra at the district level."*

To achieve this aim, the following specific **Objectives** were established:

1. To quantify the heatwave **hazard** for the 2023 pre-monsoon season using a health-based **Heat Index (HI)**.
2. To map the **exposed population** using 1km GPWv4 density data.
3. To model environmental and urban **vulnerability** using NDVI and Nighttime LST as proxies.
4. To develop a final **Heatwave Risk Index** by combining these three components.
5. To identify and **rank all districts** in Maharashtra based on their average risk score.

2 Data and Methodology

This section details the data sources, conceptual framework, and computational steps used to create the heatwave risk index for Maharashtra. All data processing and analysis were conducted using the **Google Earth Engine (GEE)** platform.

2.1 Study Area

The chosen study area is the state of **Maharashtra**, located in the western peninsular region of India. The state's geography is highly diverse, featuring the 720 km-long humid **Konkan coastal plain** (home to the Mumbai Metropolitan Region), the **Western Ghats** mountain range, and the semi-arid **Deccan Plateau**. This climatological and demographic diversity makes it a critical and complex region for a heatwave risk analysis.

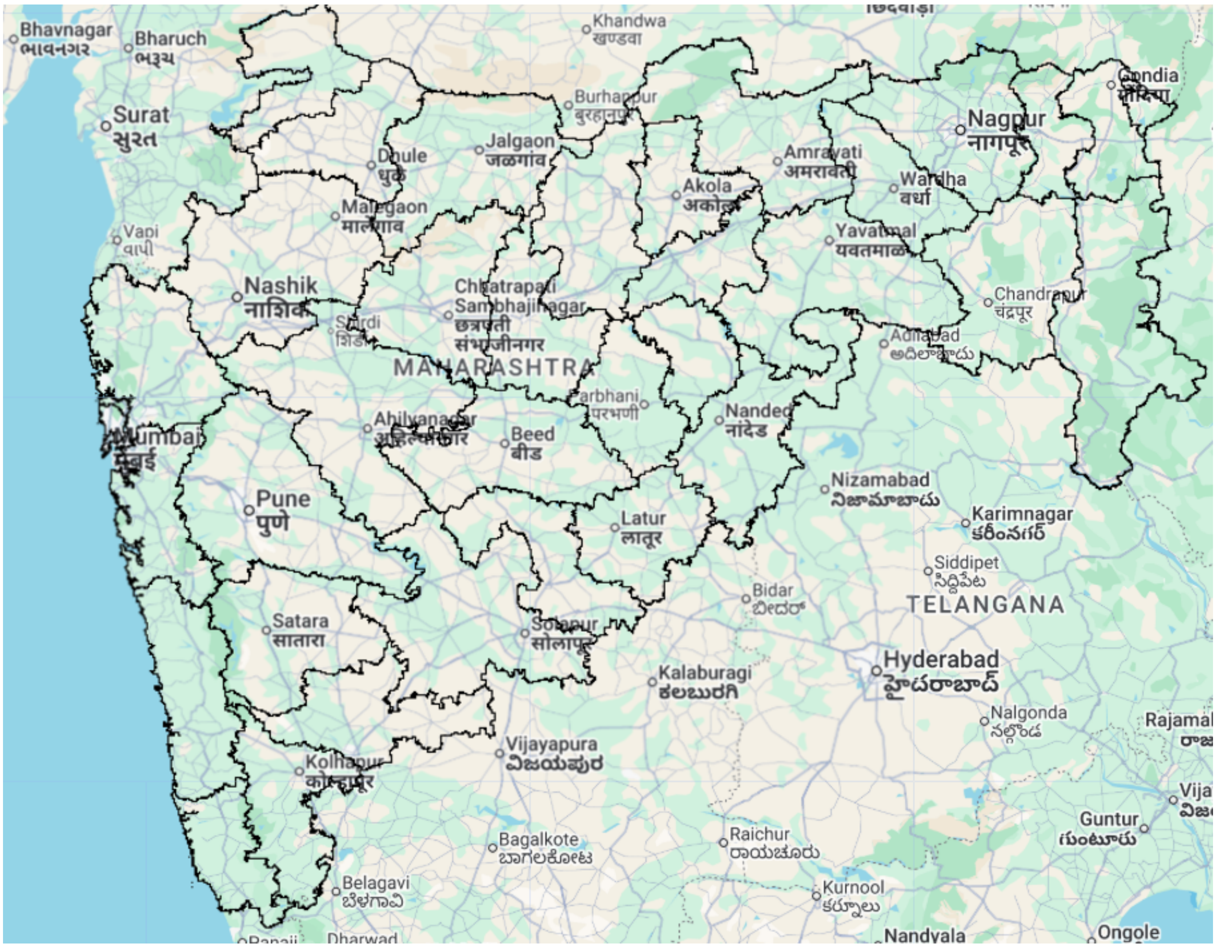

Figure 1: Map of Maharashtra showing district boundaries.

2.2 Conceptual Framework

The project is built on the standard, established framework for climate risk assessment, where **Risk** is a function of three components:

$$\text{Risk} = \text{Hazard} \times \text{Exposure} \times \text{Vulnerability}$$

- **Hazard (H):** The physical heatwave event itself. We defined this as the cumulative count of extreme heat days.
- **Exposure (E):** The population in the path of the hazard. We used population density as a direct measure.
- **Vulnerability (V):** The pre-existing conditions that make a population more susceptible to harm from the hazard. We modeled this using environmental and urban proxies.

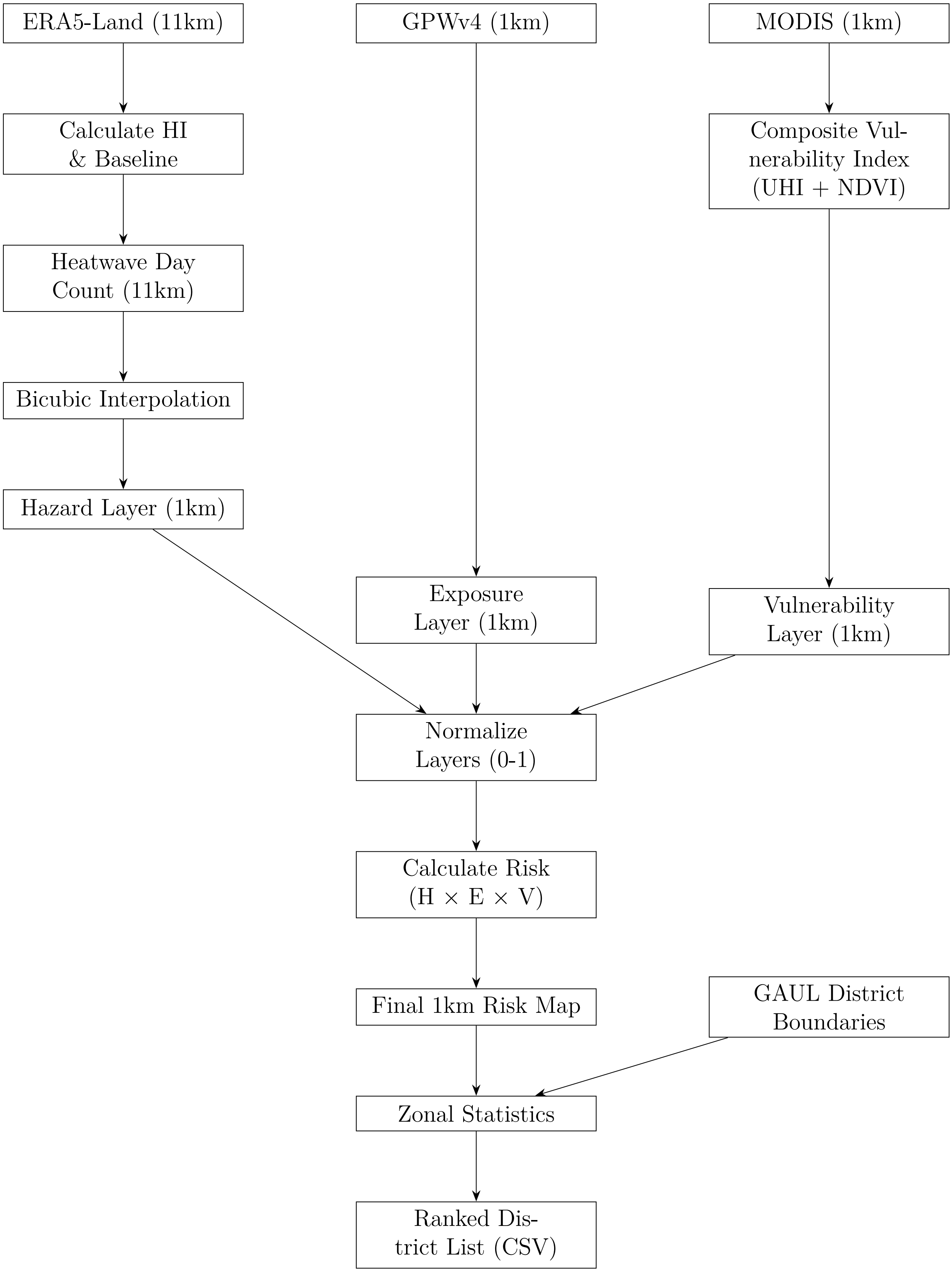

Figure 2: Methodology flowchart showing the workflow from data inputs to the final risk map.

2.3 Data Acquisition

The datasets for each component of the risk model were sourced from open-access geospatial repositories and accessed via Google Earth Engine. The data is summarized in Table 1.

Table 1: Datasets used for the Heatwave Risk Analysis

Parameter	**Data Source**	**Resolution**	**Purpose**
Hazard	ERA5-Land	11km (native)	Temp. & Dewpoint for HI
Exposure	GPWv4	1km	Population Density
Vulnerability	MODIS NDVI (MOD13A3)	1km	Vegetation (Greenness)
Vulnerability	MODIS LST (MYD11A1)	1km	Nighttime Temp. (UHI)
Boundaries	FAO GAUL (L1 & L2)	Vector	AOI & District Shapes

2.4 Parameter Calculation

2.4.1 Hazard (H) Calculation and Interpolation

The Hazard layer was the most complex to create, involving four distinct processes.

1. **Heat Index (HI) Computation:**
   First, we loaded the 11km ERA5-Land daily data (Step 2 in the GEE script). We could not use temperature alone, as this would ignore the severe health risk from high humidity on Maharashtra's coast. Therefore, we calculated the Heat Index (HI), a "feels-like" temperature, using the full Steadman biophysical formula.

```
var calculateHI = function(image) {
  var t = image.select('temperature_2m');
  var rh = image.select('RH');

  var hi = t.expression(
    'c1 + (c2*T) + (c3*RH) + (c4*T*RH) + (c5*T*T) + (c6*RH*RH) +
    ↪  (c7*T*T*RH) + (c8*T*RH*RH) + (c9*T*T*RH*RH)', {
    'T': t, 'RH': rh,
    'c1': -8.78469475556, 'c2': 1.61139411, 'c3': 2.33854883889,
    'c4': -0.14611605, 'c5': -0.012308094, 'c6': -0.0164248277778,
    'c7': 0.002211732, 'c8': 0.00072546, 'c9': -0.000003582
```

```
  });

  var simpleHI = t.expression('0.5 * (T + 16.92 + abs(T - 16.92) + 0.18 *
  ↪  RH)', {'T': t, 'RH': rh});
  var finalHI = hi.where(t.lt(26.7), simpleHI);

  return image.addBands(finalHI.rename('HI'));
};
```

2. **Decadal Baseline Definition:**
   Second (Step 3 in the script), to define what constitutes an "extreme" heat day, we created a robust 30-year (1991-2020) baseline. This was an advanced method where the 90th percentile HI was calculated for each month, averaged across three separate decades (1991-2000, 2001-2010, 2011-2020). This "Decadal Baseline" represents the "normal" threshold for extreme heat for each month.

```
function percentileByDecade(start, end) {
  var subset = baselineHI.filterDate(start, end);
  return ee.ImageCollection.fromImages(
    ee.List.sequence(1, 12).map(function(m) {
      var monthImgs = subset.filter(ee.Filter.calendarRange(m, m,
      ↪  'month'));
      return monthImgs.reduce(ee.Reducer.percentile([90]))
                      .rename('HI_threshold')
                      .set('month', m);
    })
  );
}
// Compute decadal baselines
var baseline1 = percentileByDecade('1991-01-01', '2000-12-31');
var baseline2 = percentileByDecade('2001-01-01', '2010-12-31');
var baseline3 = percentileByDecade('2011-01-01', '2020-12-31');
// Merge and average
var merged = baseline1.merge(baseline2).merge(baseline3);
var dailyThresholds = ee.ImageCollection.fromImages(
```

```
  ee.List.sequence(1, 12).map(function(m) {
    var monthly = merged.filter(ee.Filter.eq('month', m));
    return monthly.mean().set('month', m);
  })
);
```

3. **Heatwave Day Count:**

   Third (Step 3 in the script), the coarse 11km `hazardImage_coarse` was generated by counting the total number of days in our 2023 study period (Mar-May) where the daily HI exceeded this decadal baseline.

4. **Interpolation to 1km:**

   Finally (Step 4 in the script), this coarse 11km image was resampled to 1km using **bicubic interpolation**. This methodological choice was made to spatially align the hazard data with the 1km Exposure and Vulnerability layers, creating a visually "smooth" and high-resolution final map.

```
var hazardImage = hazardImage_coarse
  .resample('bicubic')  // Smooth interpolation
  .reproject({
    crs: 'EPSG:4326',
    scale: 1000  // Force to 1km
  });
```

2.4.2 Exposure (E) Calculation

The Exposure layer was modeled using the GPWv4 population density dataset for 2020. This layer was already at our target 1km resolution and was clipped to the AOI and unmasked.

2.4.3 Vulnerability (V) Calculation

The Vulnerability layer is a composite index (1km resolution) created by combining two proxies:

- **Environmental Vulnerability (`v_env`):** Derived from inverted MODIS NDVI. Low NDVI values (less vegetation) indicate higher vulnerability.

- **Urban Heat Island Vulnerability (v_uhi):** Derived from Nighttime MODIS LST. High nighttime temperatures are a strong proxy for the UHI effect, indicating dense built-up areas that retain heat and thus have higher vulnerability.

These two proxies were combined using a weighted average, giving the UHI proxy (Nighttime LST) a 60% weight and the environmental proxy (NDVI) a 40% weight.

```
var vulnerabilityImage = v_env.multiply(0.4)
  .add(v_uhi.multiply(0.6))
  .unmask(0);
```

2.5 Risk Index Generation

This final phase combined the three parameters into the final risk index.

- **Harmonization:** This step was simplified because all three layers (Hazard, Exposure, and Vulnerability) were processed to a uniform 1km resolution. No further spatial aggregation (like `reduceResolution`) was needed.

- **Normalization:** All three 1km layers were normalized to a common 0-to-1 scale using the `unitScale()` function. This step is essential to make the layers comparable before combining them.

- **Final Model:** The Final Risk Index was calculated using the standard multiplicative model: $\text{Risk} = \text{normalizedHazard} \times \text{normalizedExposure} \times \text{normalizedVulnerability}$. This model ensures that risk is only high in areas where all three components are present.

```
var finalRiskMap = normalizedHazard
  .multiply(normalizedExposure)
  .multiply(normalizedVulnerability)
  .clip(aoi)
  .rename('risk');
```

3 Results and Discussion

This section presents the findings of the heatwave risk analysis. It begins by comparing the two primary methods for harmonizing the data (the **11km "blocky" map** vs. the

1km "smooth" map) to justify the final methodological choice. It then analyzes the individual components (H, E, V) and the final risk map.

3.1 Methodological Choice: 11km vs. 1km Analysis

The core challenge of this project was the resolution mismatch between the **11km Hazard** layer (from ERA5-Land) and the **1km Exposure** and Vulnerability layers (from GPWv4 and MODIS). Two methods were tested to resolve this.

1. **The 11km "Blocky" Model (Statistically Robust):**
   The first method aggregates the 1km population and vulnerability data up to the coarse 11km grid using a `reduceResolution` function. The result, seen in Figure 3, is a "blocky" map that is statistically robust, as each 11km pixel represents the total sum of people and average vulnerability within it. However, this method produced a null error during the normalization step (`getStats`) that proved unresolvable within the interactive GEE environment, as the `exposureHarmonized` calculation was too computationally complex.

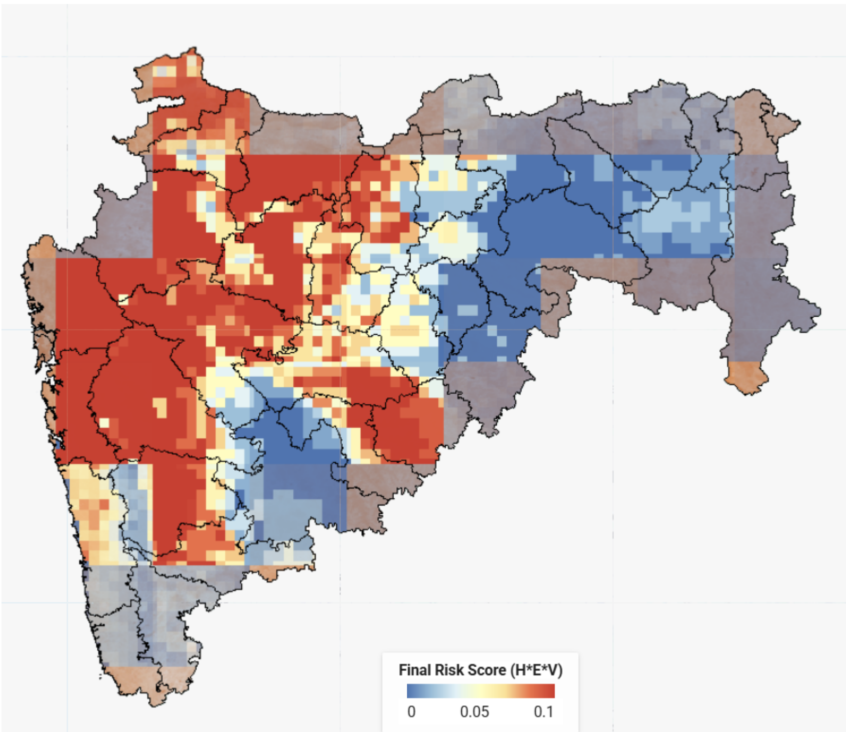

Figure 3: Final Risk Map (11km Harmonization). Note the coarse, 'blocky' pixels.

2. **The 1km "Smooth" Model (Visually Interpolated):**
   To solve this computational barrier and create a higher-resolution product, a second

methodology was adopted. Instead of downsampling the high-resolution data, we upsampled the 11km Hazard layer to 1km using **bicubic interpolation** (`.resample('bicubic')`). This created a "smooth" 1km hazard map (see Figure 4), which could then be directly multiplied with the 1km Exposure and Vulnerability layers.

This 1km "smooth" method was chosen for the final analysis as it successfully ran and produced a visually detailed, district-level map. **All following results are based on this 1km interpolated methodology.**

3.2 Analysis of 1km Risk Components (H, E, V)

3.2.1 Finding 1: The Interpolated Hazard (H) Map

The final `hazardImage` (Figure 4) shows the interpolated count of "heatwave days" (days where HI > 90th percentile).

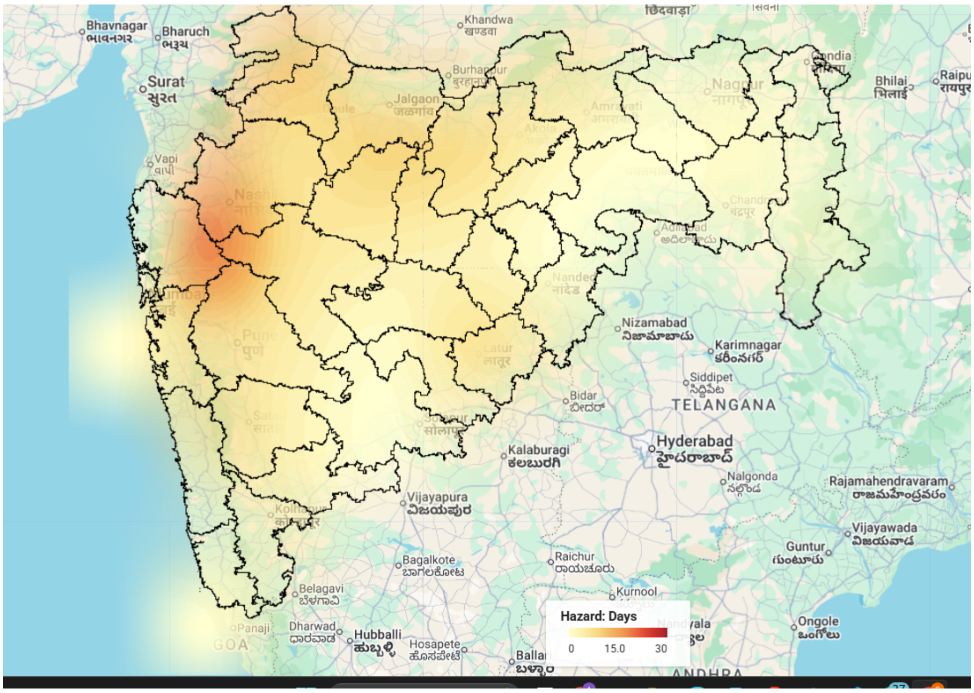

Figure 4: Interpolated Hazard (H) Map (1km).

The analysis shows a clear spatial pattern: the highest hazard (dark red) is concentrated along the **Konkan coast**. This is a key finding, highlighting that the high humidity in this region—factored into our Heat Index—creates a more severe health hazard than the high dry temperatures of the interior **Vidarbha** region.

3.2.2 Finding 2: The Exposure (E) and Vulnerability (V) Maps

The `exposureImage` (Figure 5) and `vulnerabilityImage` (Figure 6) show the human and environmental factors.

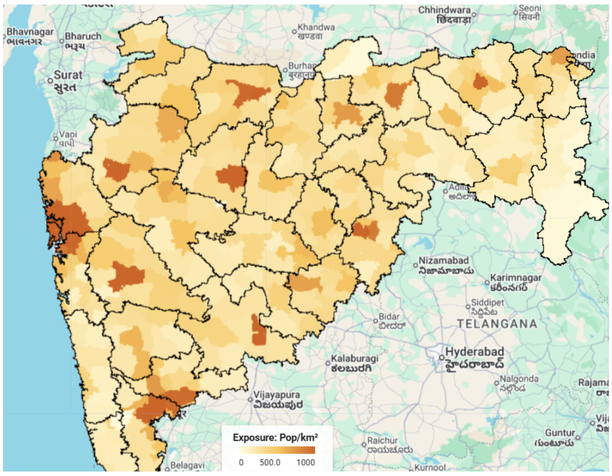

Figure 5: Exposure (E) - Population Density (1km).

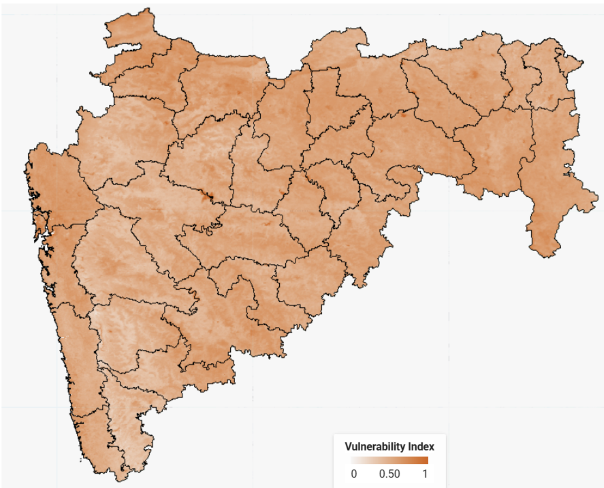

Figure 6: Vulnerability (V) - Composite Index (1km).

The Exposure map (Figure 5) clearly shows that population is concentrated in major urban centers, primarily the **Mumbai Metropolitan Region (MMR)**, **Pune**, and **Nagpur**. The Vulnerability map (Figure 6) shows that these same urban areas suffer from high vulnerability, driven by the **Urban Heat Island (UHI) effect** (high nighttime LST) and low vegetation (inverted NDVI).

3.2.3 Finding 3: The Final 1km Risk Map

The combined effect of Hazard, Exposure, and Vulnerability is clearly visible in the Final 1km Risk Map (Figure 7). This map shows the regions within Maharashtra most susceptible to heatwave impacts.

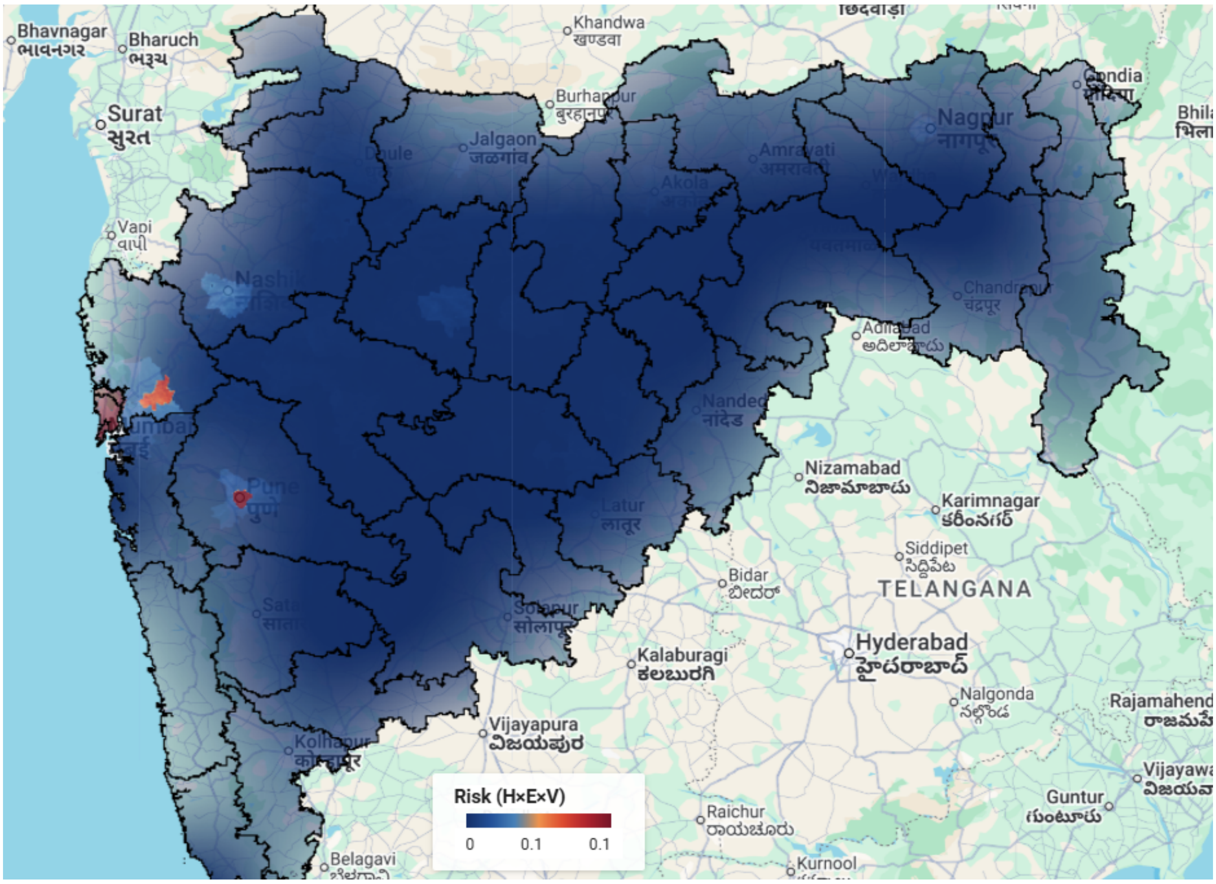

Figure 7: Final Heatwave Risk Map (1km). Risk shown in red.

The map visually confirms the earlier findings, with the **Mumbai Metropolitan Region (MMR)** standing out as the primary hotspot. This area exhibits the highest combined risk (red and orange colors), driven by its dense population, coastal humidity, and urban heat island effects. Other urban centers like Pune and Nagpur also show elevated risk, but not to the same critical extent as the MMR.

3.2.4 Finding 4: District Risk Ranking

Finally, the mean risk score was calculated for every district. The results, seen in Table 2, confirm the visual analysis.

Table 2: District-Level Risk Ranking

Rank	District	Risk_Max	Rank	District	Risk_Max
1.	Mumbai S	0.23014	19.	Amravati	0.003452
2.	Mumbai C	0.23014	20.	Solapur	0.003255
3.	Pune	0.224026	21.	Bid	0.00302
4.	Thane	0.207336	22.	Dhule	0.0029655
5.	Jalgaon	0.09368	23.	Sangli	0.002822
6.	Kolhapur	0.05674	24.	Osmanaba	0.002688
7.	Ahmedna	0.04963	25.	Parbhani	0.00212
8.	Satara	0.04702	26.	Hingoli	0.001533
9.	Buldana	0.04245	27.	Sindhudul	0.001449
10.	Nanded	0.04073	28.	Gondiya	0.001318
11.	Akola	0.04059	29.	Washim	0.001233
12.	Nashik	0.036634	30.	Wardha	0.001226
13.	Raigarh	0.029966	31.	Bhandara	0.001015
14.	Aurangaba	0.009833	32.	Nagpur	0.0010098
15.	Jalna	0.008404	33.	Yavatmal	0.000462
16.	Nandurba	0.005814	34.	Chandrapur	0.000453
17.	Latur	0.00452	35.	Garchiro	0.000265
18.	Ratnagiri	0.003592			

The zonal statistics from the 1km map show that Mumbai Suburban, Mumbai city, and Thane are, by a significant margin, the top three most at-risk districts. This provides a clear, actionable ranking for policymakers to prioritize mitigation efforts.

4 Conclusion and Limitations

This project successfully developed and implemented a high-resolution (1km) geospatial heatwave risk index for the state of Maharashtra for the 2023 pre-monsoon season. The analysis confirms that heatwave risk is not uniform and is, in fact, highly concentrated.

The main finding of this study is the identification of a critical risk hotspot in the **Mumbai Metropolitan Region (MMR)**. This is the logical outcome of the Risk = Hazard × Exposure × Vulnerability model. The MMR is the only region where all three risk components peak simultaneously:

- A high **Hazard** (driven by the high-humidity Heat Index).
- A massive **Exposure** (the densest population).
- A high **Vulnerability** (driven by the Urban Heat Island effect and low vegetation).

The final district-level ranking confirms this, identifying **Mumbai**, **Mumbai Suburban**, and **Thane** as the most at-risk districts by a significant margin. This model provides an actionable tool for policymakers to identify risk hotspots and demonstrates the power of **Google Earth Engine** for complex, large-scale climate analysis.

4.1 Limitations and Future Work

While the 1km risk map provides a significant improvement over coarser national-level models, this study has three main limitations that should be addressed in future work.

1. **Hazard Data Interpolation (The 11km-to-1km Process)**
   The primary limitation is that the core Hazard data (Heat Index) was derived from the **11km ERA5-Land** dataset. To match our 1km Exposure and Vulnerability layers, this 11km data was "upsampled" using **bicubic interpolation**. It is critical to note that this improves the visualization and creates a "smooth" map, but it **does not add new 1km-scale climate information**. The underlying hazard data is still coarse, and the interpolation can create visual artifacts, such as the transparency seen at the state's borders.

2. **Vulnerability Proxies**
   The Vulnerability index was based on environmental proxies (NDVI for greenness and Nighttime LST for the UHI effect). These are effective, but indirect. The model would be significantly more robust if it incorporated direct **socio-economic data** from the Census of India. Future analysis should aim to include district-level statistics on poverty rates, literacy, and the percentage of the elderly (over 65) and young (under 5) populations, as these groups are known to be the most vulnerable.

3. **Masked Data in Hazard Layer**
   The ERA5-Land dataset, which forms the basis of our Hazard layer, masks large inland water bodies. In our script, the `.unmask(0)` function fills these "holes" with a zero. This is a necessary step to prevent the script from crashing, but it means the final risk score for pixels over large lakes or reservoirs is **artificially zero**. This is an expected artifact of a land-based analysis but should be noted.

5 Acknowledgements

We would like to express our gratitude to the organizations that provide the open-access data, without which this research would not be possible: the **ECMWF** for the ERA5-Land dataset, **CIESIN at Columbia University** for the GPWv4 population data, and **NASA** for the MODIS (NDVI and LST) products.

This entire analysis was made possible by the **Google Earth Engine (GEE)** platform, which was used for all data acquisition, processing, and analysis. This report itself was typeset using **LaTeX**.

We also acknowledge the use of Generative AI tools to assist in the development of this project. **Gemini (Google)** was used for methodological brainstorming, GEE code generation, and for assistance with LaTeX code, English grammar, and report formatting. **ChatGPT (OpenAI)** was used for minor discussions regarding code syntax and debugging.

6 References

1. Muñoz-Sabater, J., et al. (2021). ERA5-Land: A new, high-resolution land surface reanalysis data set. *Earth System Science Data*, 13(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

2. Center for International Earth Science Information Network (CIESIN), Columbia University. (2018). *Gridded Population of the World, Version 4 (GPWv4): Population Density, Revision 11.* NASA Socioeconomic Data and Applications Center (SEDAC). https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-rev11

3. Didan, K. (2021). *MODIS/Terra Vegetation Indices Monthly L3 Global 1km SIN Grid V061.* NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MOD13A3.0

4. Wan, Z., Hook, S., & Hulley, G. (2021). *MODIS/Aqua Land Surface Temperature/Emissivity Daily L3 Global 1km SIN Grid V061.* NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/MODIS/MYD11A1.061

5. Loo, S., K. V., A., Kumar, P., S, S., J, K., M, V., & S, S. (2024). Geospatial analysis of heatwaves and associated health risks in India: A systematic review. *Heliyon*, 10(13), e33762. https://doi.org/10.1016/j.heliyon.2024.e33762

6. Food and Agriculture Organization of the United Nations (FAO). (2015). *Global Administrative Unit Layers (GAUL) 2015.* Accessed via Google Earth Engine.

7. Indian Meteorological Department (IMD). (2024). *Annual Statement on Climate of India 2023.* IMD, Ministry of Earth Sciences, Government of India.

8. Perkins, S. E., & Alexander, L. V. (2013). On the measurement of heatwaves. *Journal of Climate*, 26(13), 4500–4517.

9. Robinson, P. J. (2001). On the definition of a heat wave. *Journal of Applied Meteorology*, 40(4), 762–775.

10. Hunt, A. G. (2007). Moist heat waves. *Journal of Applied Meteorology and Climatology*, 46(11), 1810–1816.

A Supplementary Information (SI)

A.1 Final GEE Script (1km Interpolated Model)

A.2 Full District-Level Risk Ranking

Table 3: Full District-Level Risk Ranking

District	$\text{Risk}_{M}ax$	$\text{Risk}_{M}ean$
Mumbai Suburban	0.230140312	0.132396112
Mumbai city	0.230140312	0.125877362
Thane	0.207336141	0.012813747
Pune	0.224026013	0.003776061
Nashik	0.036633751	0.003523649
Nandurbar	0.005813879	0.003283429
Raigarh	0.029655101	0.003006786
Dhule	0.004995893	0.002652976
Jalgaon	0.009368053	0.002555312
Aurangabad	0.009832555	0.001956127
Buldana	0.004245427	0.001598706
Akola	0.004059442	0.001506139
Latur	0.004051711	0.001421593
Kolhapur	0.005673919	0.001290049
Ahmednagar	0.004963471	0.001185348
Jalna	0.008402093	0.001088425
Ratnagiri	0.003592105	0.001015648
Amravati	0.003452031	8.44E-04
Satara	0.004702375	8.32E-04
Parbhani	0.002120085	7.56E-04
Bid	0.0028218	6.99E-04
Sangli	0.003020283	6.21E-04
Osmanabad	0.002688352	5.52E-04
Nanded	0.004072996	5.14E-04
Washim	0.001233182	4.85E-04
Nagpur	0.010097917	4.22E-04
Sindhudurg	0.001448918	3.00E-04
Hingoli	0.001533403	2.88E-04
Gondiya	0.001318348	2.65E-04
Bhandara	0.001015159	2.19E-04
Solapur	0.003254961	2.18E-04
Chandrapur	4.53E-04	1.53E-04
Garhchiroli	2.65E-04	9.01E-05
Wardha	0.001225949	7.74E-05
Yavatmal	4.62E-04	5.00E-05