Dynamic Site Suitability for Emergency Services:

Hyderabad Hospital Analysis

Addressing "Healthcare Deserts" through advanced spatial analysis for optimal emergency medical service placement.

Github Link: https://github.com/AkiBatra25/GIS_Project

Team Number 21

Kunal Mittal (IMT2023533)
Lakshya Sachan (IMT2023612)
Kaushalraj Puwar (IMT2023063)
Akshat Batra (IMT2023025)

Citations, Data Sources & Acknowledgments

This analysis was made possible through the integration of high-quality open data sources and powerful geospatial tools. We acknowledge the following contributors to the spatial data infrastructure and analytical framework.

Spatial Data Infrastructure (Roads & POIs)

Data Owner: © OpenStreetMap contributors.

Road Network Distributor: Geofabrik GmbH (Karlsruhe, Germany) - India/Southern Zone Extracts.

Live Data Querying: QuickOSM Plugin (Etienne Trimaille).

Demographic Data

Source: WorldPop (School of Geography and Environmental Science, University of Southampton).

Dataset: 100m Population Counts (India).

Routing & Analysis Engines

Routing Technology: OpenRouteService (ORS) provided by the Heidelberg Institute for Geoinformation Technology (HeiGIT).

Software: QGIS (Open Source Geospatial Foundation).

Reference Imagery

Validation: Google Maps / Satellite Imagery used as a reference basemap for the manual digitization of the GHMC/ORR boundary extension.

The Challenge

Mitigating Healthcare Deserts in Rapidly Urbanising Hyderabad

Hyderabad is experiencing unprecedented urban growth, leading to a significant imbalance between population density and access to emergency medical services. This phenomenon, termed "Healthcare Deserts," creates critical gaps in healthcare provision, particularly for timely emergency response.

Why It Matters: The Imperative of Rapid Response

In emergency medicine, every second counts. Delayed response times directly correlate with increased morbidity and mortality rates. Traditional infrastructure planning often overlooks data-driven spatial analysis, resulting in an ad-hoc approach that fails to address these critical service gaps effectively.

Our objective is to employ a robust, data-driven approach to identify and strategize the optimal placement of emergency medical facilities, ensuring equitable access across Hyderabad's dynamic urban landscape.

Our Analytical Foundation: Tech Stack, Data, and Methodology

Core GIS Environment

QGIS (v3.4): The central hub for all vector cleaning, raster processing, and visualization tasks, ensuring robust spatial analysis capabilities.

- QuickOSM (Data Extraction)
- ORS Tools (Routing)
- Processing Toolbox (Automation)

Primary Data Sources

- **Road Network:** OpenStreetMap (OSM) via Geofabrik (Manual .shp.zip download for offline reliability).
- **Healthcare & Constraints:** OSM via QuickOSM API (Live querying for amenity=hospital and natural=water).
- **Demographics:** WorldPop Global Project (100m resolution unconstrained mosaic).
- **Boundaries:** Hybrid approach using OSM Admin Boundaries and Manual Digitization to ensure full Outer Ring Road (ORR) coverage.

Computational Engines & Algorithms

- **Routing Engine:** Local OpenRouteService (ORS) instance for offline, reproducible Network Analysis (Matrix API).
- **Vector Operations:** Fix Geometries, Centroid Extraction, Clip Vector by Extent.
- **Raster Operations:** Zonal Statistics (Sum), Min-Max Normalization, Weighted Overlay (Raster Calculator).
- **Machine Learning:** K-Means Clustering (QGIS Native) for identifying spatial "Desert" clusters.

Spatial Reference System

EPSG:32644 (WGS 84 / UTM Zone 44N): A projected coordinate system chosen to ensure accurate metric distance calculations and realistic travel-time modeling for precise results.

Road Network Acquisition & Processing: Laying the Foundation for Connectivity

1

Data Acquisition & Subsetting

The complete India road network dataset was downloaded from **Geofabrik** as a .shp.zip, ensuring offline reliability and comprehensive coverage. Subsequently, the "Clip Vector by Extent" tool in QGIS was utilised to extract only the Hyderabad study area, significantly reducing file size and optimising processing load.

2

Geometric Cleaning & Reprojection

Topological errors within the dataset were resolved using the "Fix Geometries" function. Crucially, the data was reprojected from WGS84 to **UTM Zone 44N (EPSG:32644)** to facilitate accurate, meter-based routing calculations, essential for realistic travel time modelling.

3

Attribute Engineering

The Field Calculator, employing SQL logic (`CASE`

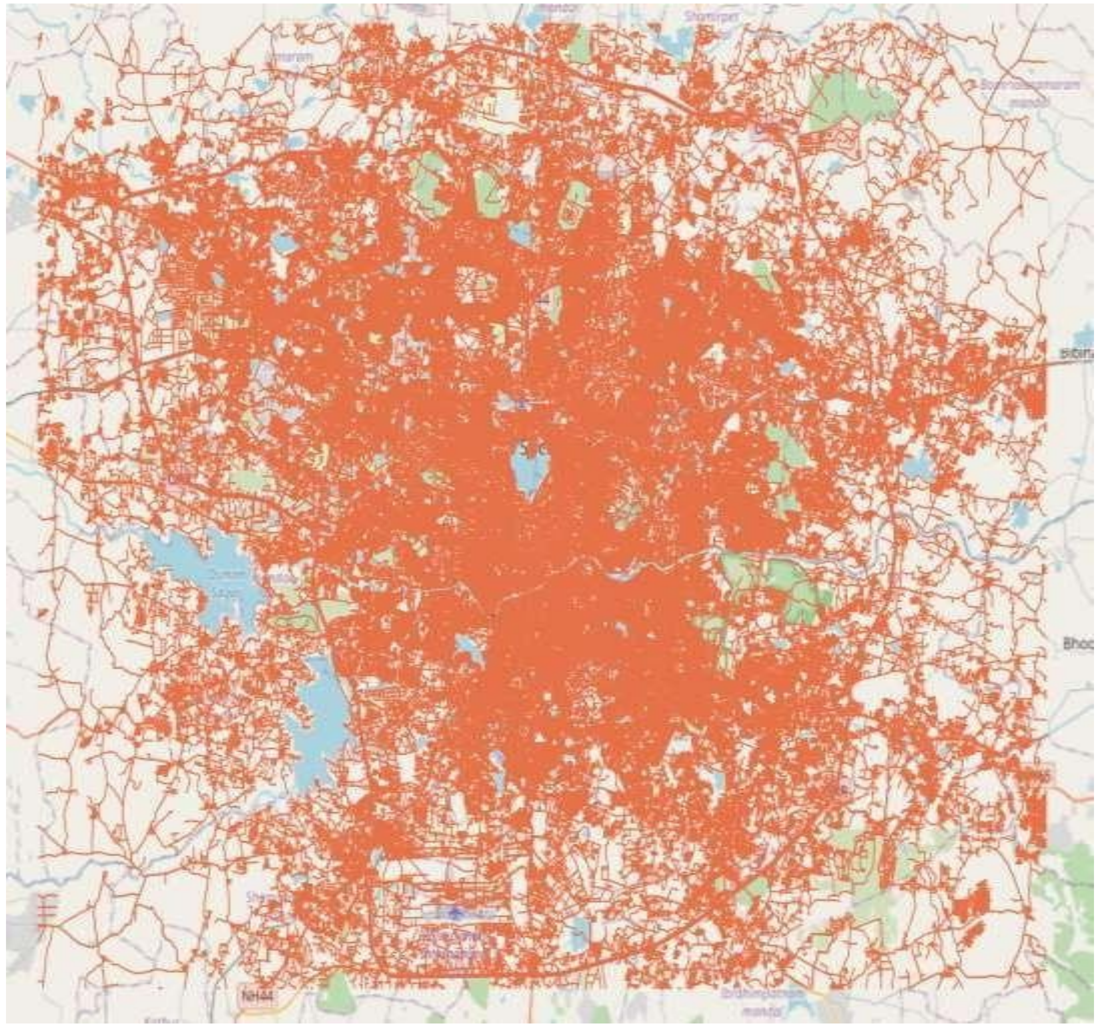
```
WHEN "fclass" IN ('motorway','trunk','primary') THEN
'primary' WHEN "fclass" = 'secondary' THEN 'secondary'
WHEN "fclass" = 'tertiary' THEN 'tertiary'
WHEN "fclass" IN ('residential','service', 'unclassified') THEN
'residential' ELSE 'other'
```

`END`), was used to streamline complex OSM tags into a standardised `road_class` hierarchy, categorising roads into Primary, Secondary, Tertiary, and Residential classifications for enhanced analytical utility.

4

Final Output

The meticulously processed vector layer, now clean and reclassified, was exported as `roads_cleaned.gpkg`, ready for integration into the network routing analysis.



Hospital Data Extraction & Processing: Pinpointing Healthcare Resources

Accurate identification and spatial representation of existing healthcare facilities are paramount for effective site suitability analysis. This stage focused on leveraging OpenStreetMap data and robust GIS processing to create a precise hospital dataset.

1

Data Acquisition (OSM)

Utilising the **QuickOSM** plugin, we extracted all features tagged `amenity=hospital` within the Hyderabad administrative boundary. This comprehensive approach captured hospitals mapped as points, polygons (building footprints), and complex relations.

2

Geometry Harmonization

To ensure compatibility with routing algorithms, which require single destination nodes, complex hospital polygons were converted to precise single-point coordinates through "Fix Geometries" and **Centroid Extraction** processes. This was critical for accurate travel time modelling.

3

Data Cleaning & Validation

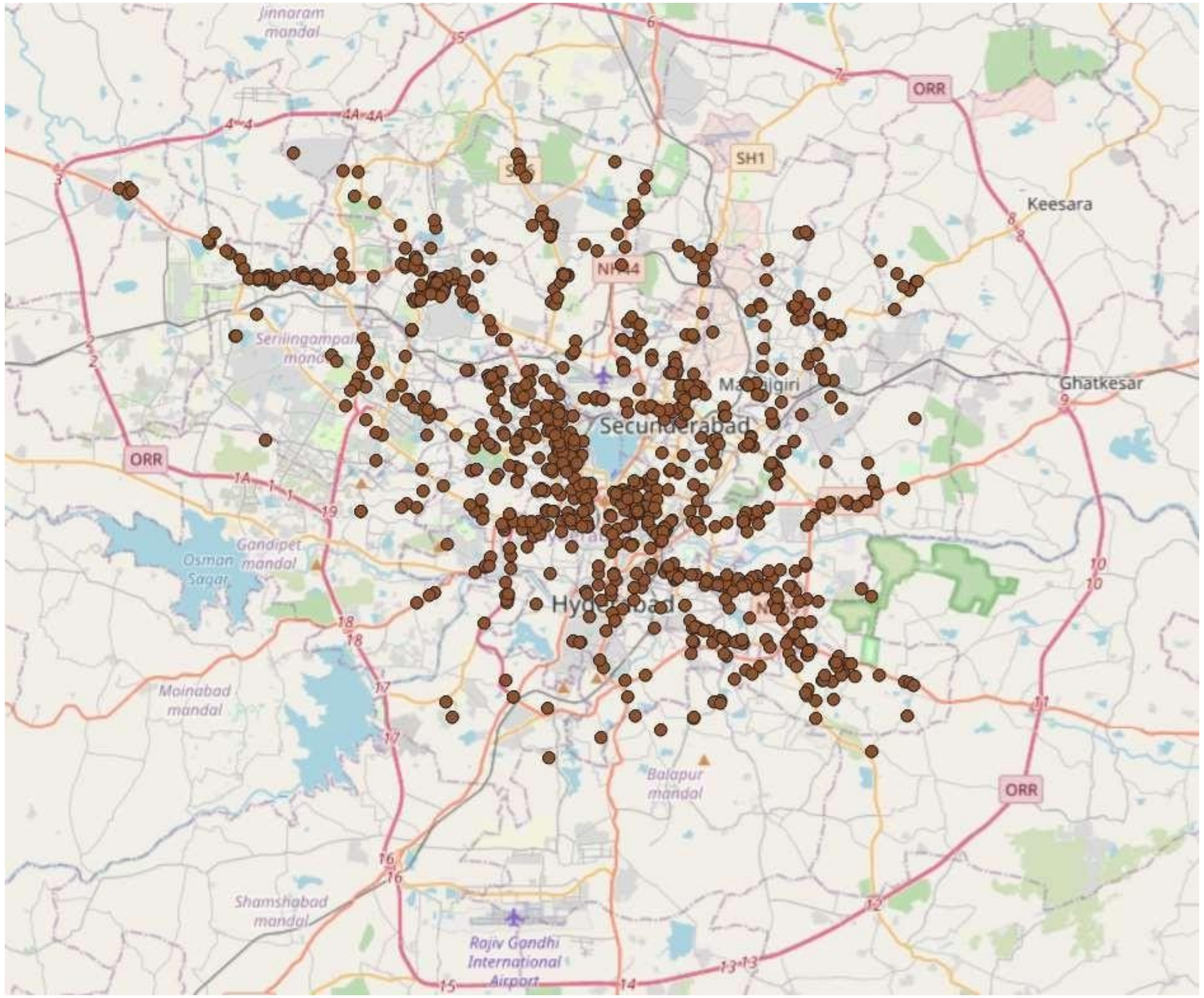
Extraneous metadata was removed, retaining only key analytical fields such as name, operator (Govt/Private), and bed capacity. A robust spatial de-duplication logic was implemented to eliminate duplicate entries where hospitals were mapped as both points and polygons.

4

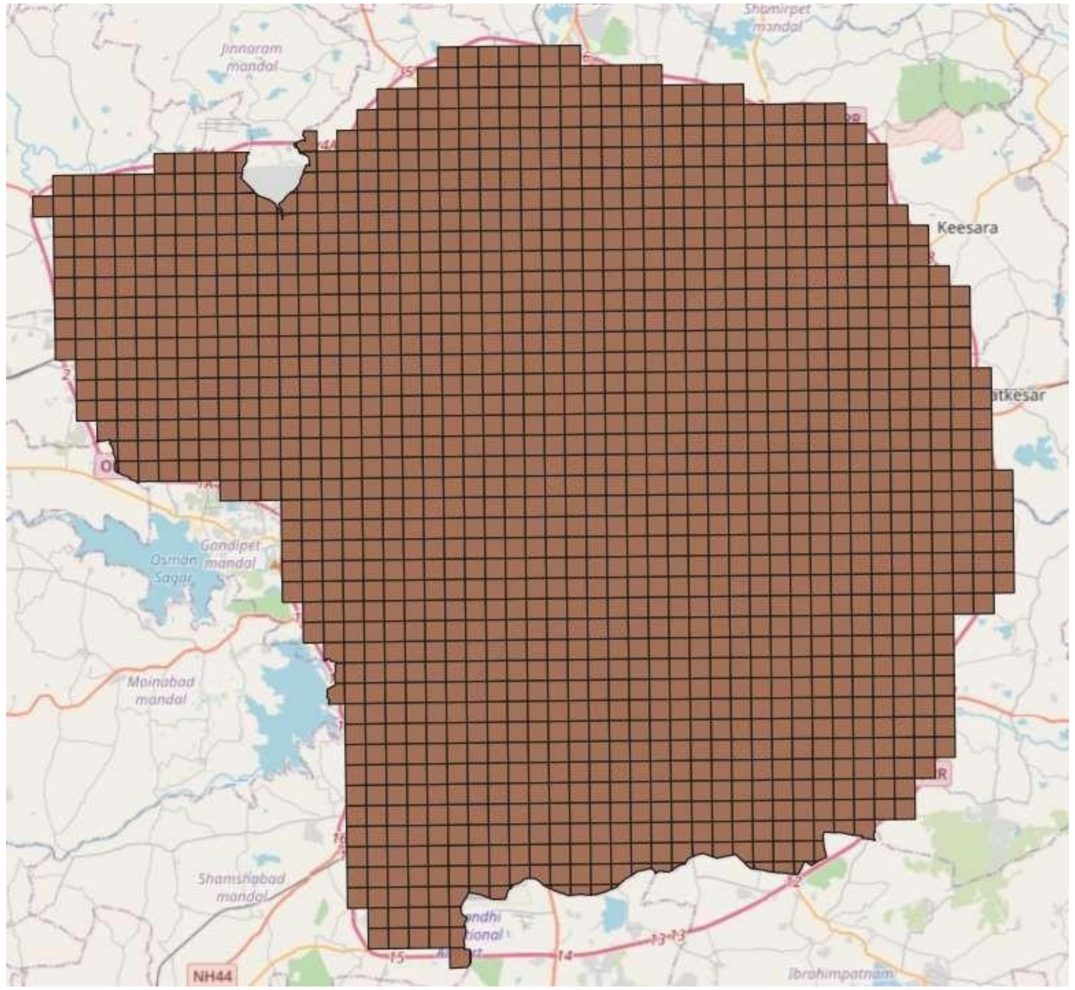
Final Subsetting

A final **Clip** operation, using the Hyderabad administrative mask, ensured the `hospitals.gpkg` layer precisely aligned with the study area, removing any bounding-box artefacts and providing a clean, accurate dataset.

Existing hospital locations in Hyderabad



Population Grid Analys is : Mapping Human Distribution and Demand



01

Data Acquisition & Boundary Creation

Raw population density was sourced from **WorldPop (100m resolution)**. The **GHMC boundary** was manually digitised using Google Maps as a reference, ensuring full coverage of the Outer Ring Road (ORR) area and addressing known gaps in standard datasets.

02

Spatial Processing & Grid Construction

All data was reprojected to **EPSG:32644 (UTM Zone 44N)** for accurate spatial measurement. A standardised **1km × 1km vector grid** was then constructed, spanning the entirety of the newly defined GHMC extent.

03

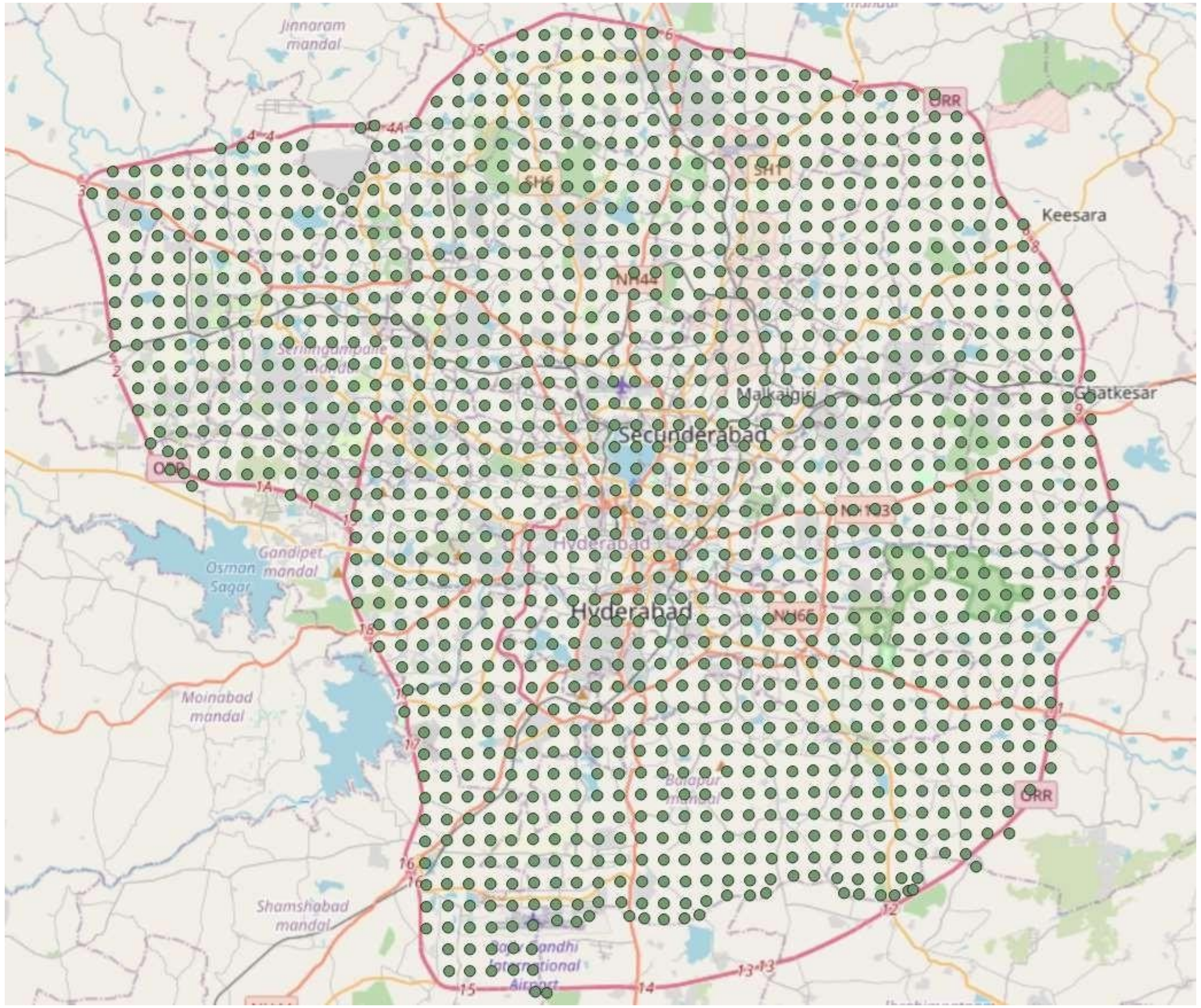
Zonal Statistics & Normalization

Zonal Statistics (SUM) aggregated the underlying 100m raster pixels to derive a total population count for each 1km grid cell. This was followed by **Min-Max Normalisation** to scale population values between 0 and 1, creating a relative density score for the suitability model.

> Formula: (Cell_Pop - Min_Pop) / (Max_Pop - Min_Pop)

Road-Network Travel Time Analysis: Real-World Accessibility

To accurately assess emergency access, we moved beyond simplistic Euclidean distances to simulate realistic ambulance travel times using network analysis. This critical step provides a foundation for identifying underserved areas.



Offline Engine Deployment

A **Local OpenRouteService (ORS)** routing engine was deployed to ensure fully offline and reproducible results, critical for consistent and verifiable analysis.

Matrix Calculation

The **ORS Matrix API** was utilised to compute travel times from every Population Grid Centroid to the nearest 20 hospitals, providing a comprehensive accessibility matrix.

Optimised Routing

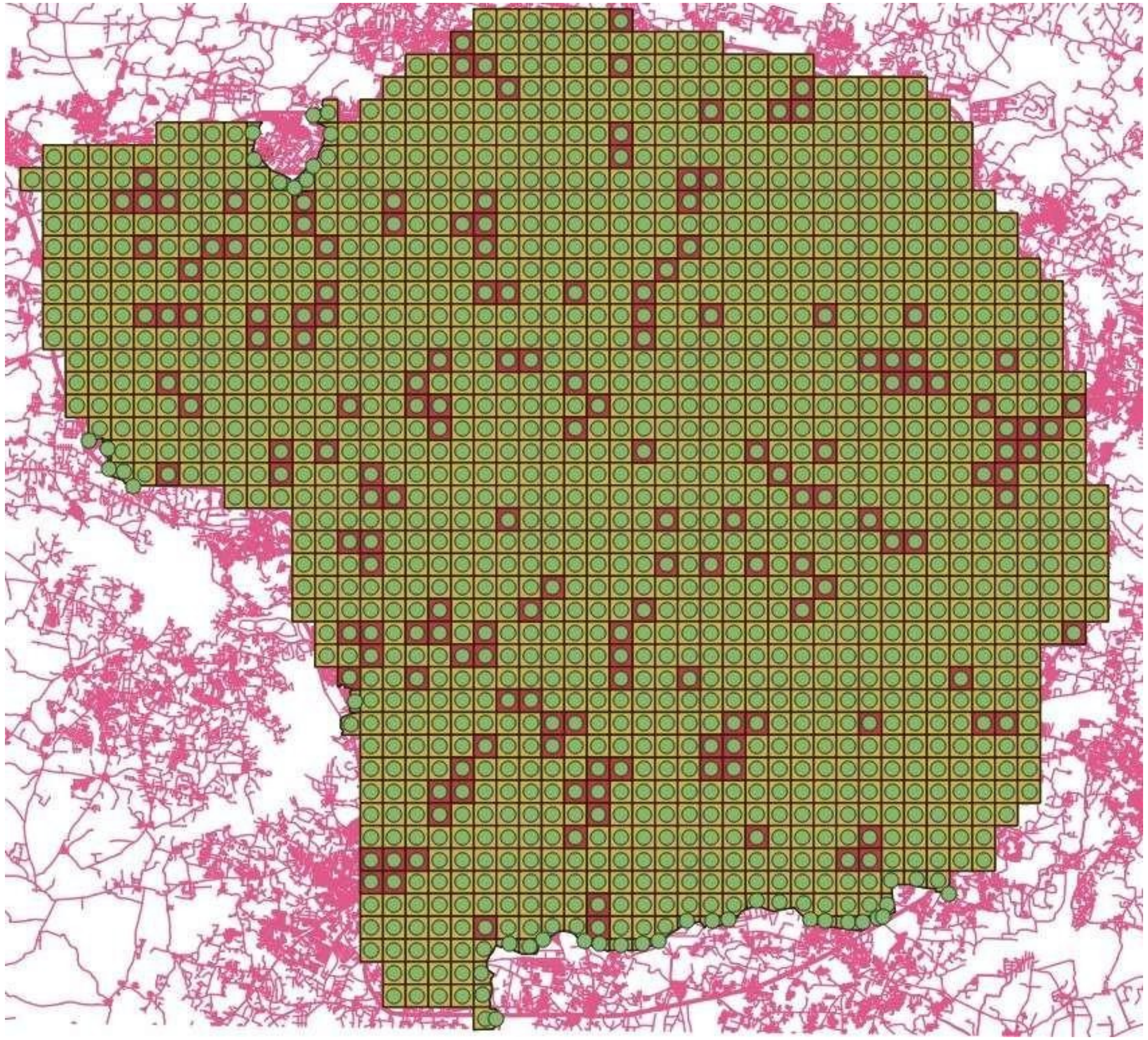
Population points were snapped to the **OSM-derived road graph** to account for actual road networks, intersections, and turn constraints, ensuring a high degree of real-world accuracy.

Automated Analysis & Output

Batch processing automated the analysis for all centroid-hospital pairs, selecting the **minimum computed travel time** for each grid cell to represent best-case emergency access. The output was a spatial travel-time surface, ready for integration into the suitability model.

K-Means Clustering: Identifying Healthcare Desert Clusters

Integrating population density with travel time data allowed us to identify "Healthcare Desert" clusters—areas simultaneously experiencing high population density and extended emergency response times.



Data Integration

Normalised Population and Travel Time datasets were integrated, preparing them for unsupervised machine learning.

Feature Engineering: The "Desert Index"

A composite metric, **Desert Index = Pop_Norm * Time_Norm**, was calculated. This index was then scaled (**Index * 100,000**) to convert small decimals into substantial continuous values, optimising separation during clustering.

K-Means Clustering

The **QGIS Processing Toolbox (K-Means)** was utilised to segment the grid based on the scaled Desert Index, grouping cells into **3 distinct classes**: High, Medium, and Low Priority.

Target Identification

The "High" class was definitively identified as the **"Healthcare Desert Cluster"**. A binary attribute, `is_desert`, was created, marking these high-priority cells as '1' to serve as a dominant factor in the final suitability model.

Brown Cells – Desert Locations
Green Cells – Non desert locations

Weighted Suitability Modelling: Identifying Optimal Locations

The final stage involved integrating all derived factors through a weighted overlay to produce a comprehensive suitability map, highlighting optimal locations for new healthcare facilities.

Constraint Modeling: "No-Go" Mask

Water bodies (natural=water) were extracted via QuickOSM. A binary raster mask was created where Water = 0 (Unsuitable) and Land = 1 (Suitable). This mask was then clipped to the city grid, ensuring only viable land areas were considered.

Factor Normalisation (Scaling 0-1)

- **Road Proximity (20%):** Inverted Normalisation applied so areas closest to roads received higher scores.
- **Travel Time (30%):** Direct Normalisation, prioritising areas with longer travel times (underserved).
- **Desert Cluster (50%):** Binary raster of `is_desert` (0 or 1) served as the primary weighting factor.

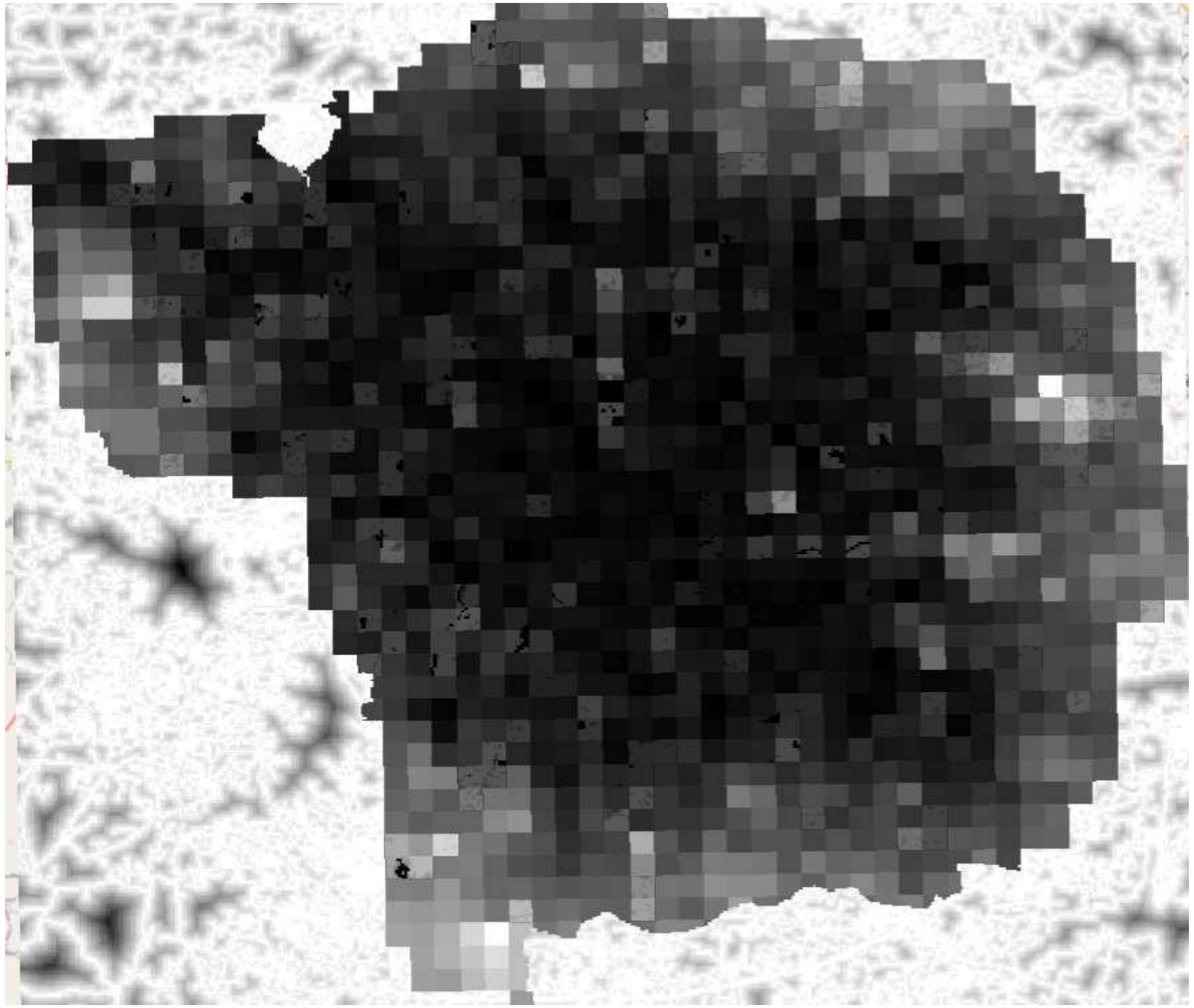
Weighted Overlay Equation

The QGIS Raster Calculator was used to execute the final decision logic:

```
Final_Score = Constraints *
((0.50 * Desert) + (0.30 * Time)
+ (0.20 * Roads))
```

This yielded a final suitability raster where brighter pixels indicated optimal sites.

Suitability Map for optimal hospital locations



- **Bright White Pixels**:

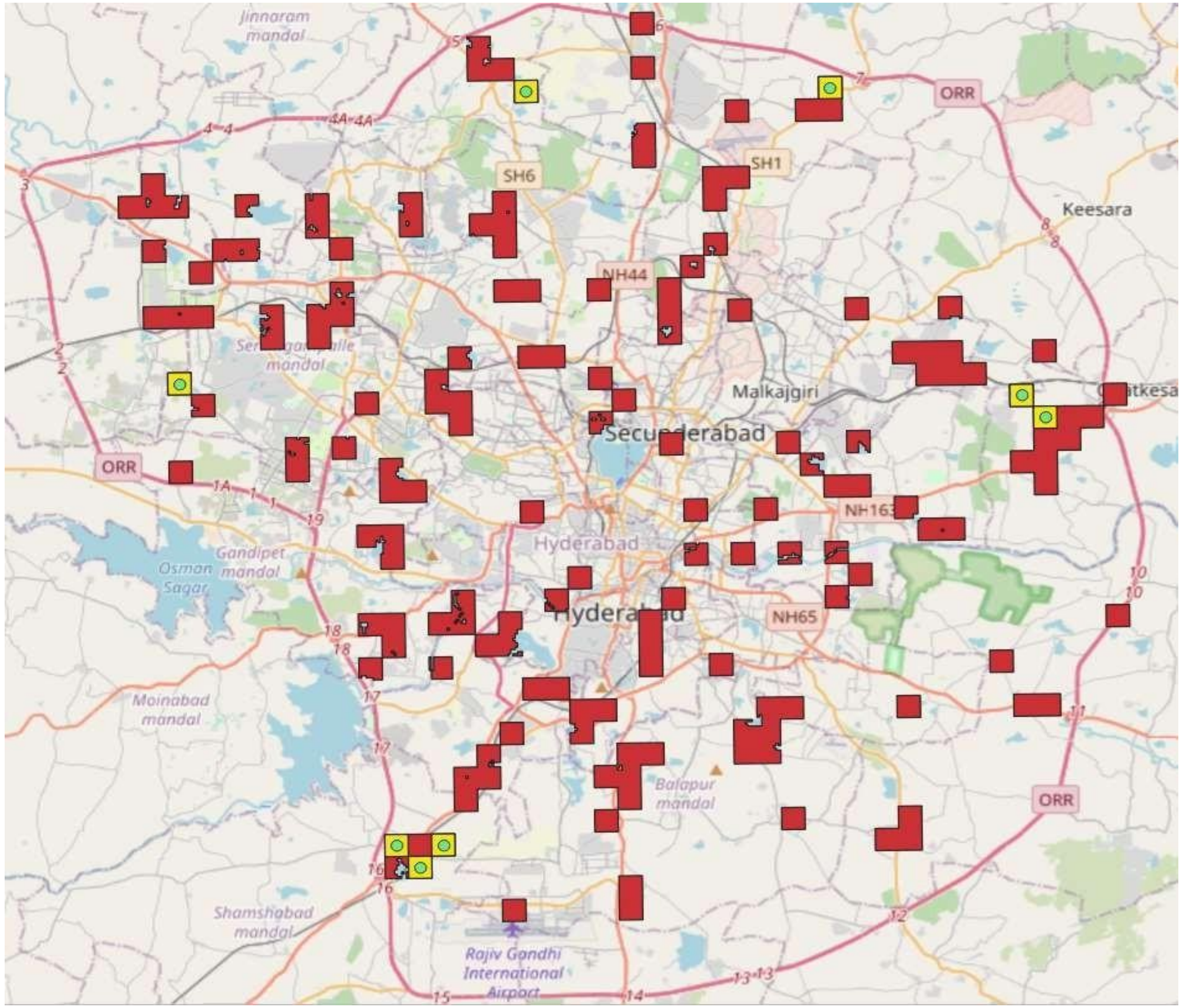
  These are the Best Locations. They are safe (not water), inside a healthcare desert, have high travel needs, and are reasonably close to a road.

- **Black Pixels**:

  These are bad locations.

Hospital Site Suitability & Priority Zone Selection

Our objective was to pinpoint high-priority geographic zones for new hospitals, integrating accessibility, population demand, and spatial suitability into a consistent, cell-based decision framework for actionable planning.



Standardisation & Aggregation

The continuous suitability raster was converted into standardised **1 km × 1 km polygon planning units**, aligning with population datasets. Zonal Statistics then calculated the mean suitability score for each grid cell.

Ranking & Selection

All valid cells were ranked by their suitability score, and the **Top 5%** were selected as high-priority zones, ensuring data-driven prioritisation.

Constraint Handling

Physically infeasible zones, such as airport restricted areas, were systematically removed and replaced with the next highest-ranked buildable zones, ensuring pragmatic recommendations.

Output & Key Advantage

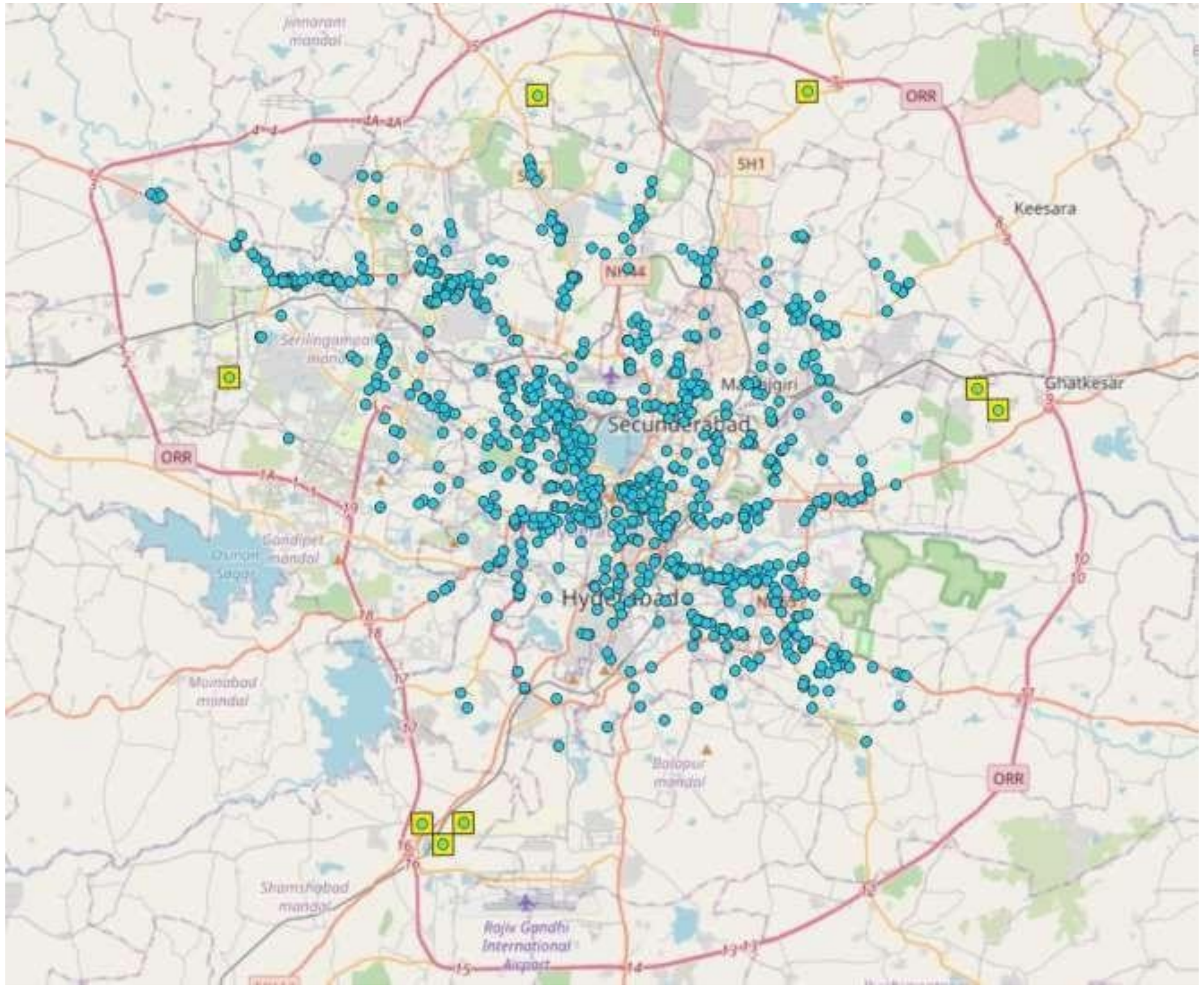
The output is a vector layer of **High-Priority 1km Planning Zones**, representing scientifically justified areas for future development. This delivers transparent, policy-relevant "Planning Zones" rather than arbitrary points, offering flexibility for detailed field assessments by urban planners.

Red Cells – All suitable sites
Yellow Cells – Candidate Sites

Final Maps: Healthcare Deserts vs. Recommended Sites

Visualizing the Problem and the Solution

The culmination of our analysis reveals both the challenge and the path forward: identifying Healthcare Deserts where access is critically limited, and pinpointing optimal sites that maximize coverage for underserved populations.



Blue Dots – Current Hospitals
Yellow Cell – Candidate Sites for hospitals

Optimized Hospital Sites - Strategic locations that maximize coverage and minimize travel time for underserved populations

Future Scope & Scalability

Expanding Impact Through Modularity, Open Data, and Enhanced Resolution

This framework's true power lies in its adaptability and scalability. Built on open data and modular design principles, it can be extended across services, geographies, and resolutions to transform emergency infrastructure planning globally.

Service Expansion (Multi-Domain Utility)

The workflow is highly modular—by simply swapping destination points from hospitals to police stations or fire brigades and recalibrating suitability weights, this framework optimizes coverage for any emergency service. The same analytical pipeline applies universally.

Geographic Scalability

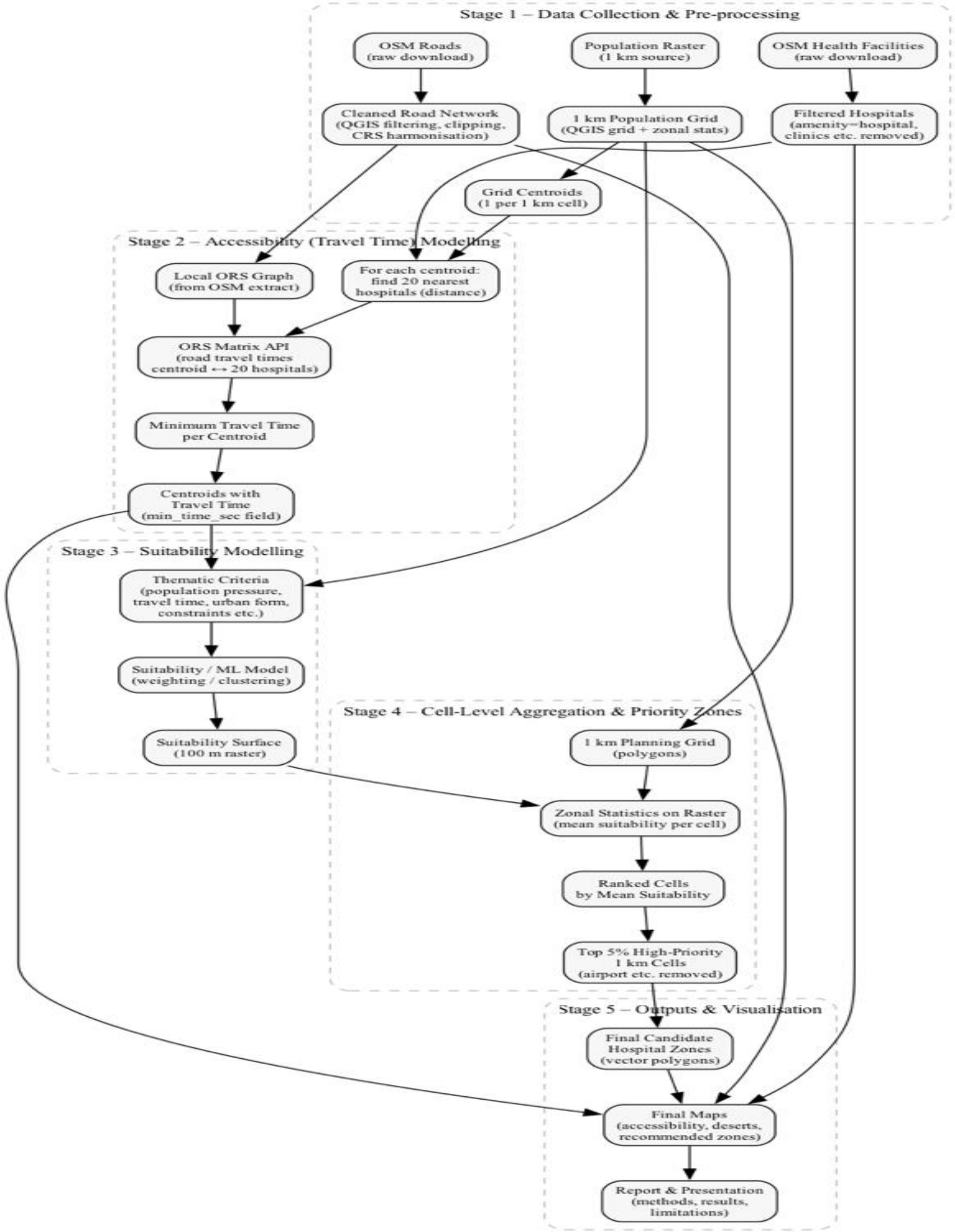
Built entirely on Open Data (OSM, WorldPop, Geofabrik), the methodology transcends proprietary restrictions. It can be immediately replicated for any major metropolitan city or developing region where data scarcity typically hinders planning, democratizing access to sophisticated spatial analysis.

Technical Optimization (High-Res Modeling)

Currently constrained by laptop processing to a 1km grid resolution. With cloud computing or GPU resources, routing resolution can increase to 100m or 50m grids, enabling hyper-accurate, block-level travel time estimations for precision planning at the neighborhood scale.

Policy & Governance Integration

Serves as a scientific decision-support tool for Government Welfare Schemes, moving planning from intuition to data-driven resource allocation. Ensures infrastructure investments are deployed exactly where demographic demand and accessibility gaps are most critical.



Stage 1 – Data Collection & Pre-processing
OSM Roads (raw download)
Population Raster (1 km source)
OSM Health Facilities (raw download)
Cleaned Road Network (QGIS filtering, clipping, CRS harmonisation)
1 km Population Grid (QGIS grid + zonal stats)
Filtered Hospitals (amenity=hospital, clinics etc. removed)
Grid Centroids (1 per 1 km cell)
Stage 2 – Accessibility (Travel Time) Modelling
Local ORS Graph (from OSM extract)
For each centroid: find 20 nearest hospitals (distance)
ORS Matrix API (road travel times centroid ↔ 20 hospitals)
Minimum Travel Time per Centroid
Centroids with Travel Time (min_time_sec field)
Stage 3 – Suitability Modelling
Thematic Criteria (population pressure, travel time, urban form, constraints etc.)
Suitability / ML Model (weighting / clustering)
Suitability Surface (100 m raster)
Stage 4 – Cell-Level Aggregation & Priority Zones
1 km Planning Grid (polygons)
Zonal Statistics on Raster (mean suitability per cell)
Ranked Cells by Mean Suitability
Top 5% High-Priority 1 km Cells (airport etc. removed)
Stage 5 – Outputs & Visualisation
Final Candidate Hospital Zones (vector polygons)
Final Maps (accessibility, deserts, recommended zones)
Report & Presentation (methods, results, limitations)

Flow Chart