ChatGPT

Datasets for Kathmandu Valley Flood Prediction

1. OpenStreetMap Road Network

- **Access**: Download the Nepal OSM extract (roads/bridges) from Geofabrik (e.g. nepal-latest-free.shp.zip) [1]. (Alternatively use OSM API/Overpass.)
- **Format**: ESRI Shapefile or OSM PBF (binary XML) containing road line geometries.
- **License**: OpenStreetMap data are under the Open Database License (ODbL) 1.0 [2] (free with attribution and share-alike).
- **Key content**: Vector lines for every road/bridge, with tags (attributes) such as *highway* (road type), *surface*, *name*, *maxspeed*, *oneway*, *layer*, *bridge*, *tunnel*, etc. [3]. These cover all major and minor roads in Kathmandu.
- **Relevance**: Provides the detailed road network needed to model connectivity and impacts: e.g. which roads or bridges flood, routing/evacuation analysis, and linking flood extents to specific streets and infrastructure.

2. Digital Elevation Model (DEM)

- **Access**: High-resolution DEMs are available from multiple sources. For example: Copernicus Global DEM (30 m) via Copernicus Data Space or AWS (public STAC at `copernicus-dem-30m` bucket) [4]; SRTM (30 m) via NASA/USGS (Earthdata or CGIAR/SRTM); ALOS AW3D30 (30 m) via JAXA EORC.
- **Format**: GeoTIFF (Cloud-Optimized GeoTIFF for Copernicus DEM); HGT/GeoTIFF for SRTM; GeoTIFF for ALOS.
- **License**: Copernicus DEM tiles are **free** to use under the Copernicus license [4]. SRTM data are public-domain (US Government) [5]. ALOS AW3D30 is free for any use under its terms (no fee) [6].
- **Key content**: Raster of elevation (meters above sea level). From this one can derive slope, aspect, flow direction/accumulation, watershed delineations, etc. Tiles cover Kathmandu Valley at ~30 m resolution.
- **Relevance**: Elevation data is critical for flood modeling – it defines water flow paths, floodplain topography, and road elevations (which affect which roads flood). High-resolution DEMs improve accuracy of slope and accumulation computations.

3. Hydrology Layers (Rivers and Streams)

- **Access**: Use global hydrology datasets. For example, **HydroRIVERS** (vector streamlines) and **HydroBASINS** (sub-basin polygons) from HydroSHEDS (based on SRTM) [7] [8]. Download from the HydroSHEDS website (http://hydrosheds.org) in Shapefile or Geodatabase format. OSM's own *waterway*= tags may also be extracted for local streams if needed.
- **Format**: ESRI Shapefiles (or geodatabase) of river/stream lines and basin polygons.
- **License**: HydroSHEDS products are freely available for any use (scientific, educational, commercial) under the HydroSHEDS license [9] (effectively open). OSM waterways are ODbL (like roads).

- **Key content**: River network geometry (lines of streams/rivers) and catchment boundaries (nested basins with Pfafstetter codes) [7] [8]. Attributes include river reach length, catchment area, flow accumulation indices, etc.
- **Relevance**: Knowing river locations and watershed areas is essential to predict where floodwaters come from and how they spread. These layers allow modeling riverine flooding and linking road segments to nearby streams (e.g. lowest road elevations next to rivers).

4. Rainfall Forecast Data (Short-term)

- **Access**: Use open weather APIs for near-real-time forecasts. Examples: OpenWeatherMap (global forecasts including hourly rain, JSON output; free tier up to 1,000 calls/day [10]) and Open-Meteo (free global API, no key required) [11]. Other services (Weatherbit, Meteomatics, etc.) exist but may be commercial.
- **Format**: Web API (JSON/REST). Forecast data (precipitation, temperature, etc.) typically returned as JSON time series.
- **License/Restrictions**: OpenWeatherMap has a free access tier (attribution required) [10]. Open-Meteo is free for non-commercial use under CC BY 4.0 [11] (open attribution license).
- **Key content**: Short-term forecasts (hourly or daily) of rainfall intensity (mm/h or mm) and other meteorological variables for locations in Kathmandu.
- **Relevance**: These forecasts provide the input rainfall over the next hours/days for prediction. A road-level flood alert system can poll such APIs to get localized rain forecasts and trigger flood warnings accordingly.

5. Historical Rainfall Data

- **Access**: Gridded precipitation archives. Two common ones: (a) **NASA GPM IMERG** – half-hourly and daily global precipitation (0.1° grid) from 2000–present. Obtain from NASA GES DISC (requires EarthData login) or AWS Open Data. (b) **CHIRPS** – daily rainfall (0.05° grid) from 1981–present, for 50°S–50°N. Download CHIRPS from UC Santa Barbara (https://data.chc.ucsb.edu) [12] or via Google Earth Engine.
- **Format**: HDF5 or NetCDF (IMERG), GeoTIFF or NetCDF (CHIRPS). Files contain precipitation (mm/hr or mm/day) on regular grids.
- **License**: IMERG data are US Government (public domain) [13]. CHIRPS data have been released into the public domain (no copyright) [12].
- **Key content**: Time series of past rainfall over Kathmandu (satellite-gauge merged estimates). IMERG provides very fine temporal (30-min) data; CHIRPS gives long-term daily totals. Variables include precipitation amount and sometimes quality flags.
- **Relevance**: Historical rain records are used to train or calibrate flood models. For example, feeding multi-year rainfall into a simulation to learn how past rainfalls produced floods at road locations. They also serve to compute climatological statistics.

6. Historical Flood Incident Maps/Polygons

- **Access**: Event-based flood maps from agencies. For example, **UNOSAT** publishes satellite-derived flood assessments (e.g. Kathmandu Sep 2024 flood maps) – often as PDF/GeoTIFF products with delineated flood extents. (See example: *"Flood Impact assessment... Kathmandu"* from Spot-6 imagery [14].) **Copernicus EMS** (Emergency Mapping) provides rapid flood maps (vector polygons and

images) after major events; these can be downloaded (on request) from the EMS portal [15]. ReliefWeb/OCHA posts situational reports and maps (satellite water masks) for Nepal floods (e.g. "Satellite-detected water extent" infographics).
- **Format**: Shapefile or GeoJSON for flood polygons (EMS, UNOSAT AI outputs), GeoTIFF or PNG for inundation masks, PDF maps. Data from UNOSAT or EMS may include GIS vector layers. ReliefWeb maps are images.
- **License/Restrictions**: UNOSAT/UN products are typically available freely for humanitarian use (often CC BY-NC or no restrictive license). EMS products are free under Copernicus emergency rules (essentially open for crisis response) [15]. ReliefWeb content is public information (usually CC BY).
- **Key content**: Extent of areas inundated during past flood events (polygons), identified flooded streets/land from satellite, counts of affected buildings, etc. Example: UNOSAT reported ~8 km^2 flooded in Kathmandu (Sep 2024) [14]. These datasets often include metadata like date, data source, analysis notes.
- **Relevance**: Historical flood extents are valuable ground truth for model training and validation: they show which areas (and roads) actually flooded in known events. Polygons can be intersected with road network to label flooded road segments. They also help validate the model's predictions against real incidents.

7. Satellite Imagery and Flood Masks

- **Access**: Free copernicus imagery. **Sentinel-1 SAR** (C-band, 10 m) and **Sentinel-2 MSI** (optical, 10 m) are open via Copernicus SciHub or AWS. Sentinel-1 is crucial for flood mapping (works through clouds). Level-1 GRD products (amplitude images) are standard; Level-2 water masks (not routinely provided) must be derived. Also useful: the Sen1Floods11 dataset provides curated Sentinel-1/2 image chips with flood labels (Google Cloud Storage "gs://sen1floods11") under CC BY 4.0 [16].
- **Format**: Sentinel data are in SAFE (ZIP) or COG GeoTIFF (for GRD). Sen1Floods11 offers GeoTIFF chips of 512×512 pixels with water/non-water masks [16]. Other derived flood products (e.g. JRC Global Surface Water) come as GeoTIFF.
- **License**: Copernicus Sentinel data are open (no charge, usage permitted for any purpose). Sen1Floods11 (example flood dataset) is CC BY 4.0 [16].
- **Key content**: SAR backscatter (used to identify water) and optical reflectance (visual flood inundation). Derived flood masks or labels (e.g. binary water extent) are provided in Sen1Floods11. These data include metadata (acquisition date, polarization).
- **Relevance**: These images and masks can be used to detect and delineate floods, train ML models (as in Sen1Floods11), or fuse with road networks. Historical flood detection in Kathmandu (e.g. using Sentinel-1) can refine the model's mapping of floodwater to roads.

8. Crowdsourced Reports (Community Inputs)

- **Access**: Formal crowdsourced flood-report datasets for Nepal are scarce. One related effort is the **"PrakopAlert"** mobile app (Nepal Red Cross/ICIMOD) which delivers local flood forecasts to users [17], but it does not publish user-submitted flood reports. Some social platforms (Facebook/Twitter) or messaging groups may contain location-tagged flood observations, but no centralized open dataset. Humanitarian channels (OCHA/HDX) occasionally compile local reports in SitReps, but usually not geocoded.
- **Format**: N/A – no standard dataset. PrakopAlert is an Android app (data likely on a server). ReliefWeb postings are PDFs/images.

- **License**: The PrakopAlert app content is CC BY-NC-ND 4.0 [18]. Volunteer data (if any) would typically be shareable for humanitarian use, but no specific license data.
- **Key content**: Potential content could include user-reported flood locations, photos or text reports (city, road, time). In practice, we found no public feed of such data for Nepal. The PrakopAlert app provides authoritative forecasts (rainfall, streamflow) [17] rather than crowdsourced observations.
- **Relevance**: Crowdsourced flood reports could offer real-time local confirmation of flooding; if available, they could validate or supplement predictions. In absence of a dedicated dataset for Nepal, this remains an area to monitor (e.g. local NGOs or media sharing geotagged flood news).

Sources: OpenStreetMap/Geofabrik (roads) [1] [2] [3]; Copernicus DEM and SRTM (elevation) [4] [5] [6]; HydroSHEDS (rivers) [7] [9]; OpenWeatherMap and Open-Meteo (forecasts) [10] [11]; NASA GPM IMERG, CHIRPS (historical rain) [19] [13] [12]; UNOSAT/Copernicus EMS/ReliefWeb (flood extents) [14] [15]; Sentinel-1/2 and Sen1Floods11 (imagery) [16]; ICIMOD/NRCS PrakopAlert (crowd app) [17] [18]. Each dataset is publicly accessible (links above) and provides the variables needed for the FloodWatch prediction model.

---

[1] Geofabrik Download Server
https://download.geofabrik.de/asia/nepal.html

[2] License/Use Cases - OpenStreetMap Wiki
https://wiki.openstreetmap.org/wiki/License/Use_Cases

[3] OpenStreetMap Data in Layered GIS-Format
https://download.geofabrik.de/osm-data-in-gis-formats-free.pdf

[4] Copernicus Digital Elevation Model (DEM) - Registry of Open Data on AWS
https://registry.opendata.aws/copernicus-dem/

[5] USGS EROS Archive - Digital Elevation - Shuttle Radar Topography Mission (SRTM) 1 Arc-Second Global | U.S. Geological Survey
https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-shuttle-radar-topography-mission-srtm-1

[6] Dataset | ALOS@EORC
https://www.eorc.jaxa.jp/ALOS/en/dataset/aw3d30/aw3d30_e.htm

[7] HydroRIVERS
https://www.hydrosheds.org/products/hydrorivers

[8] [9] HydroBASINS
https://www.hydrosheds.org/products/hydrobasins

[10] Weather API - OpenWeatherMap
https://openweathermap.org/api

[11] ○ Free Open-Source Weather API | Open-Meteo.com
https://open-meteo.com/

[12] CHIRPS: Rainfall Estimates from Rain Gauge and Satellite Observations | Climate Hazards Center - UC Santa Barbara
https://www.chc.ucsb.edu/data/chirps

13 GPM IMERG Final Precipitation L3 1 day 0.1 degree x 0.1 degree V07 (GPM_3IMERGDF) at GES DISC - Catalog
https://catalog.data.gov/dataset/gpm-imerg-final-precipitation-l3-1-day-0-1-degree-x-0-1-degree-v07-gpm-3imergdf-at-ges-dis-f1669

14 unosat.org
https://unosat.org/static/unosat_filesystem/3993/UNOSAT_A3_Natural_Protrait_FL20240928NPL_CapitalCity_Kathmandu_03Oct2024.pdf

15 Home | Copernicus EMS On Demand Mapping
https://mapping.emergency.copernicus.eu/

16 Modified Sen1Floods11 Dataset for Change Detection
https://zenodo.org/records/7946594

17 18 Innovative 'PrakopAlert' mobile app launched to strengthen community resilience and disaster preparedness in Nepal – ICIMOD Servir HKH
https://servirhkh.yipl.net/update/innovative-prakopalert-mobile-app-launched-to-strengthen-community-resilience-and-disaster-preparedness-in-nepal/

19 IMERG: Integrated Multi-satellitE Retrievals for GPM | NASA Global Precipitation Measurement Mission
https://gpm.nasa.gov/data/imerg