UNIVERSITÉ DE GENÈVE

Analyse géomatique appliquée à l'évaluation des risques naturels et environnementaux à Djibouti

Structuration des données, production cartographique et diffusion WebGIS au sein de la Fondazione CIMA

Geneva, 28 August 2026

Author / student: Bruno Colavitto, PhD

Academic supervision (UniGe): Andrea De Bono, PhD

Host: CIMA Research Foundation

Company supervision: Eva Trasforini, PhD

Abstract

This dossier presents the work carried out during a two-month internship at the CIMA Research Foundation within the Risk Assessment and Loss Data Unit, as part of the project "Development of a National Multi-Risk Profile for the Republic of Djibouti and a Multi-Risk Early Warning System". The internship focused on the geospatial preparation, integration and dissemination of information required for hazard, exposure and vulnerability assessment.

The work applied several methods addressed during the Certificate in Geomatics, including raster and vector processing, coordinate reference systems, remote-sensing-derived data, spatial analysis, metadata, Google Earth Engine, Python scripting for QGIS and GeoServer layers publication. Priority hazard layers included riverine floods (both water-depth and floodable area footprint) and heavy rainfall. Exposure modelling covered population, residential buildings, healthcare and education facilities, and roads infrastructure. Census-based vulnerability was represented through the Population Vulnerability Index and the Coping and Warning Capacity Deficit.

A Local-Data-First approach was adopted, prioritizing national statistics and institutional information and complementing them with local consultations, remote sensing and international datasets when necessary. Particular attention was given to provenance, spatial resolution and the distinction between observed, derived and modelled information.

The internship resulted in contributions to Deliverables D2, D3 and D4, a set of harmonized geospatial layers, reproducible GEE workflow, and datasets being consolidated and progressively published through GeoServer and the WebGIS portal. The experience demonstrated that reliable geomatics products depend not only on spatial processing, but also on traceability, metadata, methodological consistency and explicit treatment of data limitations.

Résumé

Ce dossier présente le travail réalisé au cours d'un stage de deux mois à la CIMA Research Foundation, au sein de l'unité Risk Assessment and Loss Data, dans le cadre du projet « Development of a National Multi-Risk Profile for the Republic of Djibouti and a Multi-Risk Early Warning System ». Le stage avait pour objectif principal de préparer, harmoniser et diffuser l'information géospatiale nécessaire à l'analyse des aléas, de l'exposition et de la vulnérabilité.

Les travaux ont mobilisé plusieurs compétences développées durant le Certificat en géomatique : traitements raster et vectoriels, systèmes de coordonnées, télédétection, analyse spatiale, métadonnées, Google Earth Engine, scripts Python pour QGIS et publication des couches sur GeoServer. Les principaux aléas traités dans le cadre du stage comprennent les inondations fluviales (soit les cartes de profondeur de l'eau, soit les cartes délimitant les zones inondables) et les fortes précipitations. L'exposition a été représentée à travers la population, les bâtiments résidentiels, les infrastructures critiques (écoles, hôpitaux, réseau routier). La vulnérabilité géospatiale a été abordée au moyen du Population Vulnerability Index et du Coping and Warning Capacity Deficit.

L'approche *Local-Data-First* a privilégié les statistiques nationales, les données institutionnelles et les informations issues des consultations locales, complétées lorsque nécessaire par la télédétection et des sources internationales. Une attention particulière a été portée à la provenance, à la résolution spatiale et à la distinction entre données observées, dérivées et modélisées.

Le stage a contribué aux livrables D2, D3 et D4, à la production de couches géospatiales harmonisées, à des procédures reproductibles sous GEE, ainsi qu'à la structuration progressive et à la publication des données via GeoServer et le portail WebGIS. L'expérience a montré que la qualité d'un produit géomatique dépend autant du traitement spatial que de la traçabilité, des métadonnées et de la gestion explicite des limites des données.

Table of contents

List of figures

List of tables

1. Introduction

1.1. *Internship rationale*

Natural and environmental hazards can affect several dimensions of society at the same time, including population, housing, infrastructure, productive activities, public services and ecosystems. In this context, geospatial information provides an essential basis for understanding where people and assets are located, how they are distributed across the territory and how they may be affected by different hazards.

In the context of humanitarian research, and particularly for Disaster Risk Reduction initiatives, increasingly emphasize is being made in two aspects: locally-based multi-risk assessments and geospatial products being available for different users and stakeholders. For this reason, geomatics becomes a key scientific discipline that, by grouping several geospatial techniques and concepts facilitates exposure and risk assessments that can help decision-makers to improve risk knowledge and, therefore, make informed, data-based decisions to reduce future natural or environmental hazards impacts.

This dossier presents the work done during the two-month internship at the CIMA Research Foundation, within the Risk Assessment and Loss Data Unit. The internship was conducted in the framework of the project entitled "Development of a National Multi-Risk Profile for the Republic of Djibouti and a Multi-Risk Early Warning System". The main objective of the work was to produce a set of harmonized, georeferenced layers representing hazard, exposure and vulnerability information in Djibouti, as a key input to develop a national multi-risk profile.

The assessment followed a local-data-first approach, combining official national statistics and institutional datasets with locally collected data and international open source datasets. Several geospatial tools and methodologies were used to achieve the analysis such as: reprojection of data, geometry validation, attribute standardization, spatial allocation, rasterization, spatial overlay operations and documentation of sources and processing steps through a centralized layer catalogue. The internship contributed to the production of thematic maps and spatial outputs, designed for reporting but also for dissemination through the project's WebGIS infrastructure.

1.2. *The host organization*

The CIMA Foundation is an Italian research foundation working in the field of environmental risks, disaster prevention, climate resilience and risk-informed decision-making. Its activities combine scientific research, applied methodologies, technical assistance and cooperation with public institutions and international partners.

The internship was carried out within the Risk Assessment and Loss Data Unit. This unit contributes to the analysis of hazards, exposure, vulnerability and disaster losses, with the objective of improving the evidence available for risk assessment and disaster risk reduction. Within this institutional setting, geospatial information plays a central role in integrating demographic, infrastructural, environmental and socio-economic information.

The internship was supervised by Dr. Eva Trasforini. The work was conducted in collaboration with the international technical team, a team of three national experts and two main institutions involved in Disaster Risk Reduction in Djibouti. This collaborative structure was particularly important because the availability and quality of spatial data varied considerably between sectors.

1.3. The Djibouti multi-risk project

The project "Development of a National Multi-Risk Profile for the Republic of Djibouti and a Multi-Risk Early Warning System" aims to strengthen the national understanding of disaster risk and to support the development of information and warning capacities for multiple natural and environmental hazards.

The project requires a coherent representation of hazards, exposed elements, vulnerability characteristics and potential impacts. These components are interdependent. Hazard information cannot be translated into potential impacts without knowing which populations, buildings, infrastructures, productive systems and ecosystems are located in affected areas, and which conditions make these assets more vulnerable.

This dataset was conceived to be made available to the DRR users to support subsequent impact estimation and risk aggregation. Hazard products were produced through modelling or processing from open data sources (mostly ERA-5 derived), with part of the analysis made by colleagues of the Fondazione CIMA, remaining outside of the scope of this internship report. On the other hand exposure data collection was partially carried out by the local consultants team, with digitalisation, harmonisation or processing tasks remaining in charge of the student. For the vulnerability assessment, the 2024 RGPH-3 census data provided by the Republic of Djibouti remained as the main reference to produce social vulnerability indicators.

The resulting products are not the final completion of the national multi-risk assessment. They will subsequently be used for risk modelling following three different approaches based on hazard hierarchy established, hazard typology, the expertise of host organization and the exposure detail availability. Thus, probabilistic risk models will be combined with deterministic scenario approaches or simply qualitative descriptions of hazard dynamics and potential impacts. Once defined the potential impacts, and together with the three main components of risk properly assessed, the guidelines for an early warning system will be produced. In any case, the scope of the present internship and report are only limited to the main hazard, exposure and vulnerability layers production and publication.

2. Core theoretical concepts

2.1. *Risk components*

Risk assessment requires distinguishing between the processes that may cause harm and the people, assets or systems that may be affected. The definitions adopted in this dossier broadly follow the terminology established by UNDRR (2017) for disaster risk reduction.

Hazard

A hazard is a potentially damaging physical, environmental or human-induced phenomenon. In the context of Djibouti, several hazards were identified including floods, droughts, heatwaves, storms, earthquakes, sea-level-related processes and coastal pollution. However, not all of them were considered high priority, as a hazard is characterized not only by its presence, but also by attributes such as its location, intensity, duration, frequency, historical impacts and spatial extent.

During the internship, part of the hazard information was processed and treated by the student while part was analyzed externally (by colleagues of the host organization), therefore the associated tasks were limited to layers harmonization or map production. For this stage, the main objective in terms of hazard, whenever possible, is to prepare the spatial footprint of the hazard, in order to help assessing exposure correctly or processing raw data to obtain the hazard maps in terms of return periods.

Exposure

Exposure refers to the presence of people, buildings, infrastructure, productive activities, ecosystems or other valued elements in places that may be affected by a certain hazard. Exposure is therefore intrinsically spatial: it depends on both the location of the hazard and the location of the exposed elements.

The exposure assessment developed during the internship covered several categories: population, residential and non-residential buildings, transport infrastructure, health and education facilities, etc.

Exposure does not, by itself, indicate that damage will occur. It indicates that an element is located within an area potentially affected by a hazard. The severity of the resulting consequences also depends on the characteristics and vulnerability of the exposed element.

Vulnerability

Vulnerability encompasses the physical and socioeconomic conditions that influence susceptibility and the capacity to anticipate, cope with and recover from hazardous events (Blaikie et al., 2004; UNDRR, 2017). It may include physical, social, economic, environmental and institutional dimensions. Examples include building construction materials, demographic characteristics, access to water and health services, etc.

These elements can be grouped into 2 types of vulnerability: physical or socio-economic vulnerability. The former was not chosen here for geospatial representation, given that vulnerability/impact functions will be included in the risk calculation chain in further stages of the project. Contrarily, socio-economic indicators, obtained mainly from the national census report were treated as geospatial data, producing vulnerability outputs in map format.

Risk

Risk refers to the potential consequences resulting from the interaction between hazards, exposed elements and their vulnerability. At a conceptual level, risk can be represented as:

$$R = H \times E \times V$$

where [H] represents the hazard, [E] the exposure and [V] the vulnerability.

The relationship between these components depends on the available data, the spatial and temporal scale, the classification system, the assumptions adopted and the intended use of the results. For example, a population located in an area affected by flooding may have different potential consequences depending on the quality of housing, access to roads, proximity to health facilities and availability of early warning information.

This distinction between hazard, exposure and vulnerability is also consistent with the risk-knowledge framework adopted for multi-hazard early warning systems (UNDRR & CIMA Research Foundation, 2024)

2.2. Local-data-first approach

The assessment followed a local-data-first approach. This means that official national statistics, institutional datasets, local consultations and national expert knowledge were prioritized whenever they were available and sufficiently relevant.

The approach does not imply that local data are always complete or spatially detailed. Instead, it recognizes that data produced or validated within the country may better reflect national definitions, administrative structures, sectoral terminology and local conditions than generalized global datasets.

For example, for population or housing information or for vulnerability socio-indicators analysis, the RGPH-3 census (INSTAD, 2024) was treated as the main reference for official totals and main socio-economic attributes. Building footprints, settlement layers and other geospatial datasets were used to provide spatial detail and to support the disaggregation of census information rather than to replace official totals. Local consultants at the same time provided relevant information, not in a geospatial format, which was then converted for further include them into the risk calculation workflow.

Hence, the data used in the assessment is organized into four broad levels:

1. Official national data (e.g., the census data RGPH-3)
2. Regional and local evidence (e.g., derived from consultations)
3. National expert interpretation
4. Open-source and international datasets, where national geometries were unavailable.

This hierarchy was used to guide decisions when several sources described the same category. The dataset with the finest spatial resolution was not automatically considered the most authoritative. Authority, relevance, date, spatial accuracy, thematic consistency and institutional status were considered together.

2.3. WebGIS and online platform setting

A WebGIS provides a way to make spatial information accessible through a web interface without requiring every user to install specialized GIS software. It can support consultation, visualization, comparison and communication among technical teams, institutions and decision-makers.

In the context of the Djibouti project, the WebGIS provides access to:

- administrative reference layers;
- hazard layers;
- exposure layers;
- infrastructure datasets;
- thematic maps;
- metadata;
- selected vulnerability indicators;
- documentation and methodological notes.

The online interface is composed by two main functions: the dashboard and the WebGIS. For the purpose of this internship the work was focused on preparing the dataset and uploading them for the WebGIS publication, being out of the scope of the internship the dashboard final preparation (texts, organization, etc.).

Overall, the online interface is developed by the IT team of the host organization using an open and flexible framework running in Java. The online interface can be easily set and modified using a Django interface for the visual components and a GeoServer interface to treat and handle the geospatial layers. The latter is an open-source server used to publish geospatial data and maps through standardized web services. In the project architecture, it acts as an interface between the spatial database and the web mapping client. GeoServer helps to connect spatial data sources to the online interface, to publish vector and raster layers, defining workspaces and data stores and configuring styles, to provide online map services (WMS, WFS) and to manage layer metadata.

3. Data and methods

3.1. Methodological framework

The geospatial work carried out during the internship was designed to provide a consistent spatial framework connecting three components of disaster-risk assessment: hazard, exposure and vulnerability. The methodological approach followed the Local-Data-First principle adopted by the Djibouti multi-risk project.

The internship therefore involved three different levels of methodological intervention. First, some hazard products had already been generated through models developed by CIMA or external providers. In these cases, the main contribution consisted of harmonizing the datasets and preparing them for cartographic and WebGIS use. Instead, the heavy-rainfall analysis involved a dedicated geospatial workflow implemented during the internship to derive extreme daily precipitation scenarios for different return periods. On the other hand, exposure and social-vulnerability layers required more extensive data integration, including source comparison, geometry processing, statistical allocation and construction of derived indicators.

A common distinction was maintained throughout the workflow between source observations and project-derived products. Official statistics were retained as official quantities; consultation-derived information was treated as local or expert evidence; remote-sensing products were principally used as spatial representations; and results produced through allocation, classification or modelling were explicitly identified as modelled or derived datasets. This distinction was necessary to prevent spatial precision from being confused with statistical authority.

3.2. Data preparation, harmonization and quality control

The datasets used during the internship were highly heterogeneous. They included raster climate and hazard products, vector building footprints and infrastructure inventories, administrative statistics, historical cartography, consultation-derived locations, remotely derived datasets and tabular census information. Their spatial resolution, temporal reference, classification systems, attribute structure and coordinate reference systems therefore differed substantially.

All layers intended for spatial analysis were converted, where required, to the common project coordinate reference system WGS 84 / UTM Zone 38N (EPSG:32638). The projected coordinate system was preferred for operations involving area, distance, raster alignment and spatial overlay. Geographic versions in EPSG:4326 could subsequently be retained or generated where required for data exchange and WebGIS publication.

Data preparation included geometry validation, reprojection, clipping to the national area of interest, field-name standardization, harmonization of administrative identifiers, unit conversion, duplicate assessment and verification of missing or inconsistent attributes. Vector layers were visually compared with recent satellite imagery where positional accuracy or completeness could not be established. Raster products were checked for geographic coverage, spatial resolution, NoData values, units and consistency with the analytical interpretation documented by the team.

Metadata were managed through a centralized layer catalogue. For each processed layer, the catalogue records its source, reference period, spatial characteristics, principal processing operations, provenance, validation status, limitations and WebGIS publication status.

3.3. Hazard analysis

For this dossier, two hazard workflows were selected as they best illustrate the geomatics contribution of the internship: riverine/wadi flooding and heavy rainfall. They also represent important hazards within the national hierarchy made (report D2) and they introduce two different methodological approaches: an externally produced hydraulic product requiring spatial preparation to compute exposure (floodable area footprint); and an extreme 1-day rainfall return period-analysis workflow completely implemented during the internship.

Riverine flood footprint

The project flood assessment used Joint Research Centre (JRC) flood-hazard layers (Dottori et al., 2016) as the national-scale representation of flood extent and water depth. These products were subsequently hydrologically adjusted by the hydrology unit from Fondazione CIMA using discharge series generated through the Continuum model (Silvestro et al., 2013). The chosen dataset represents flood hazard through different return periods. 10-, 50- and 100-year return-period maps were selected from this dataset and are shown, for Djibouti-Ville, in Figure 1.

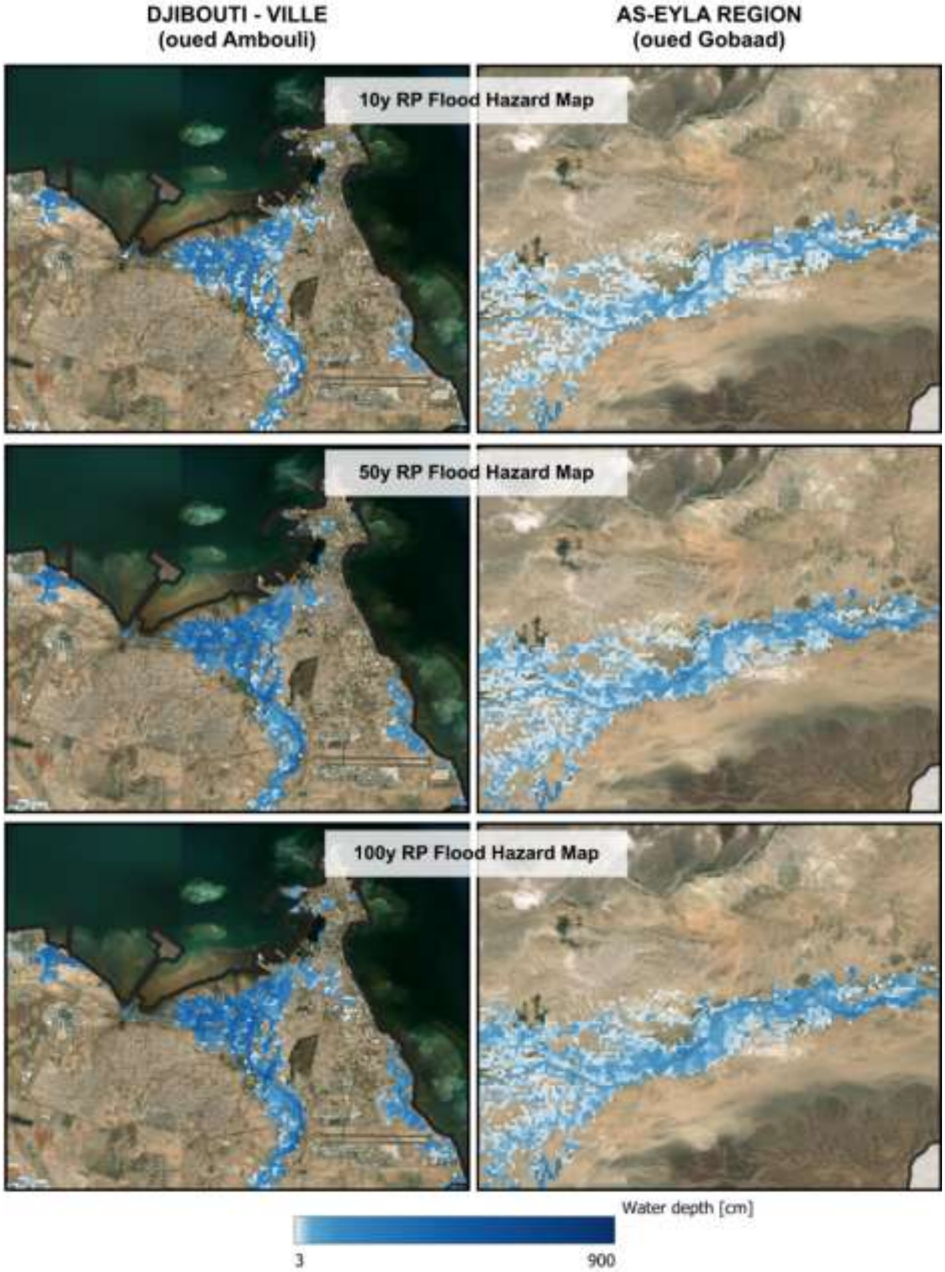

Figure 1. Evolution of riverine flood hazard maps for different RP in Djibouti-Ville and As-Eyla. Both locations show contrasting flood environments as mentioned in historical records and institutional/regional consultations. Increasing RP illustrate the progressive change in inundation extent and/or water depth under more severe flood conditions (Source: Annex A – D2 Figure 11).

The hydraulic and hydrological modelling itself is outside the scope of the internship. The geospatial task performed for the exposure workflow consisted instead of preparing a flood footprint suitable for intersection with population and asset layers.

The first geospatial procedure was to create from the original water-depth a polygon layer representing the floodable areas, required for exposure calculation (Figure 2). The original raster retains the information on flood intensity for subsequent physical-risk calculations. The raster was filtered by choosing a minimum water-depth threshold of 30 cm, applied to the 100y RP map to derive the Boolean flood footprint. NoData values were revised and the raster was aligned, keeping the 90 m grid to match the population and building exposure layers. Subsequently, this raster layer was converted to polygons, intersected to the administrative divisions (with the Union tool) to get the final footprint of floodable areas.

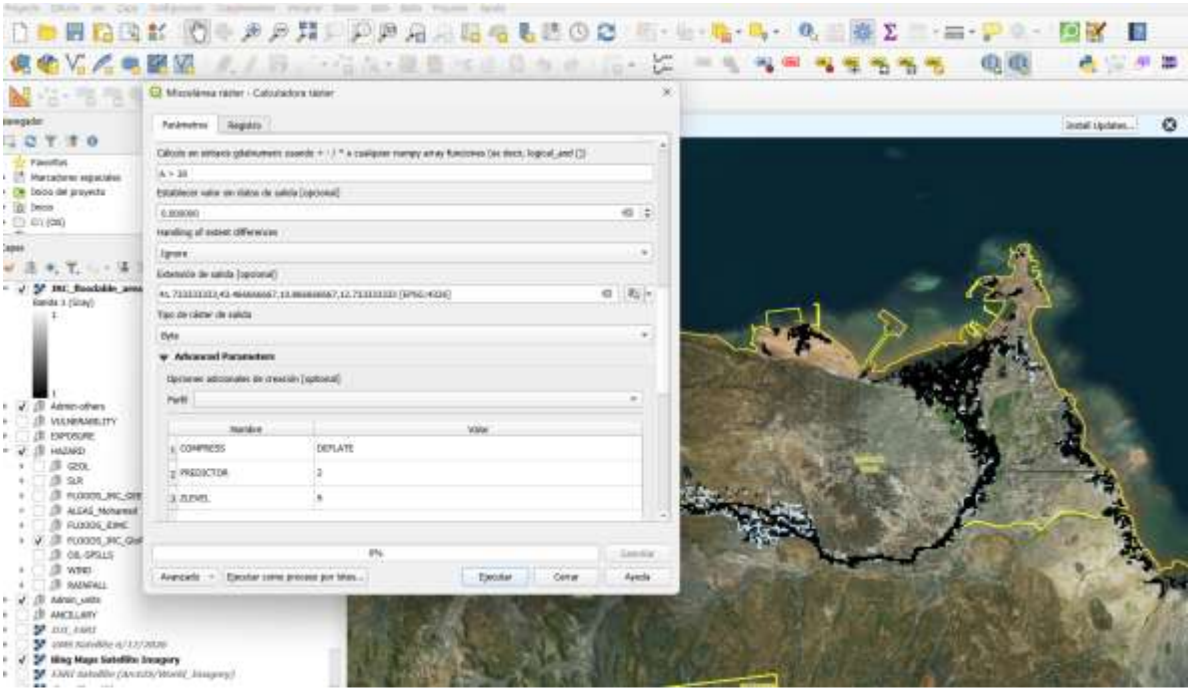

Figure 2. Screenshot from QGIS with the GDAL raster calculator tool used to derive the inundation footprint from the riverine flood hazard, by creating a Boolean raster, filtering out values lower than 30 cm.

Heavy rainfall

Heavy rainfall was the principal hazard for which a dedicated statistical and geospatial processing workflow was implemented during the internship. Daily ERA5 precipitation was analyzed for the period 1980–2019, providing 40 complete years of spatially consistent information across Djibouti.

Annual maximum 1-day precipitation (Rx1day) was first extracted for each raster cell. A Gumbel extreme-value distribution was then fitted independently to the annual maxima at each location using the method of moments. The fitted distributions were used to estimate 1-day rainfall return levels corresponding to 10-, 50- and 100-year return periods (RP10, RP50 and RP100). The RP50 and RP100 estimates involve progressively greater extrapolation beyond the 40-year record.

The workflow was implemented through Google Earth Engine (Annex E.1), allowing the meteorological archive to be filtered, processed and exported without downloading the complete source dataset locally. The resulting rasters were subsequently prepared for GIS integration, cartographic representation and WebGIS publication.

The three outputs represent rainfall intensity associated with different statistical return periods and should not be confused with the RP10, RP50 and RP100 riverine-flood scenarios. They characterize extreme 1-day precipitation, whereas flood return periods represent the hydrological and hydraulic response of the drainage system.

The main limitation is the length of the available series. The 50-year and especially the 100-year return levels require extrapolation beyond the 40-year analysis period and therefore have greater statistical uncertainty. A second important limitation is the ERA5 native spatial resolution (0.25° ≈ 27 km), which does not capture highly localized convective rainfall or fine-scale topographic effects.

Other relevant climatic hazards

Heatwaves and drought were also identified as critical hazards within the wider Djibouti multi-risk assessment. Their underlying climatic analyses were developed by CIMA's meteorological team and are documented in Deliverable D2. They were therefore not reproduced as methodological components of this internship. Where required for dissemination, the resulting raster products were treated as existing project layers and prepared for integration into the WebGIS environment.

3.4. Exposure analysis

Exposure assessment involved the production of different layers, among which some were selected for this internship report: spatial representations of the population, residential buildings and essential infrastructure that could subsequently be intersected with the hazard layers.

The source hierarchy followed the Local-Data-First approach. National census and institutional statistics were used to control quantities and official classifications; spatial layers supplied by the national GIS consultant were used where available; consultation-derived evidence was used to refine infrastructure; and global datasets and satellite imagery were used for spatial support, gap filling and quality control.

Administrative divisions

A common administrative divisions layer was required to associate statistical information with spatial geometries and to compare all subsequent layers consistently. Updated GIS boundaries corresponding exactly to the 2024 RGPH-3 census structure were not available during the assessment. Administrative units were therefore reconstructed from RGPH-3 maps and adjusted against existing national reference boundaries (Figure 3).

These polygons constitute project-derived geometries rather than a replacement for an official INSTAD administrative GIS dataset. This distinction was retained in their metadata and is important because population and vulnerability statistics are attached to these polygons.

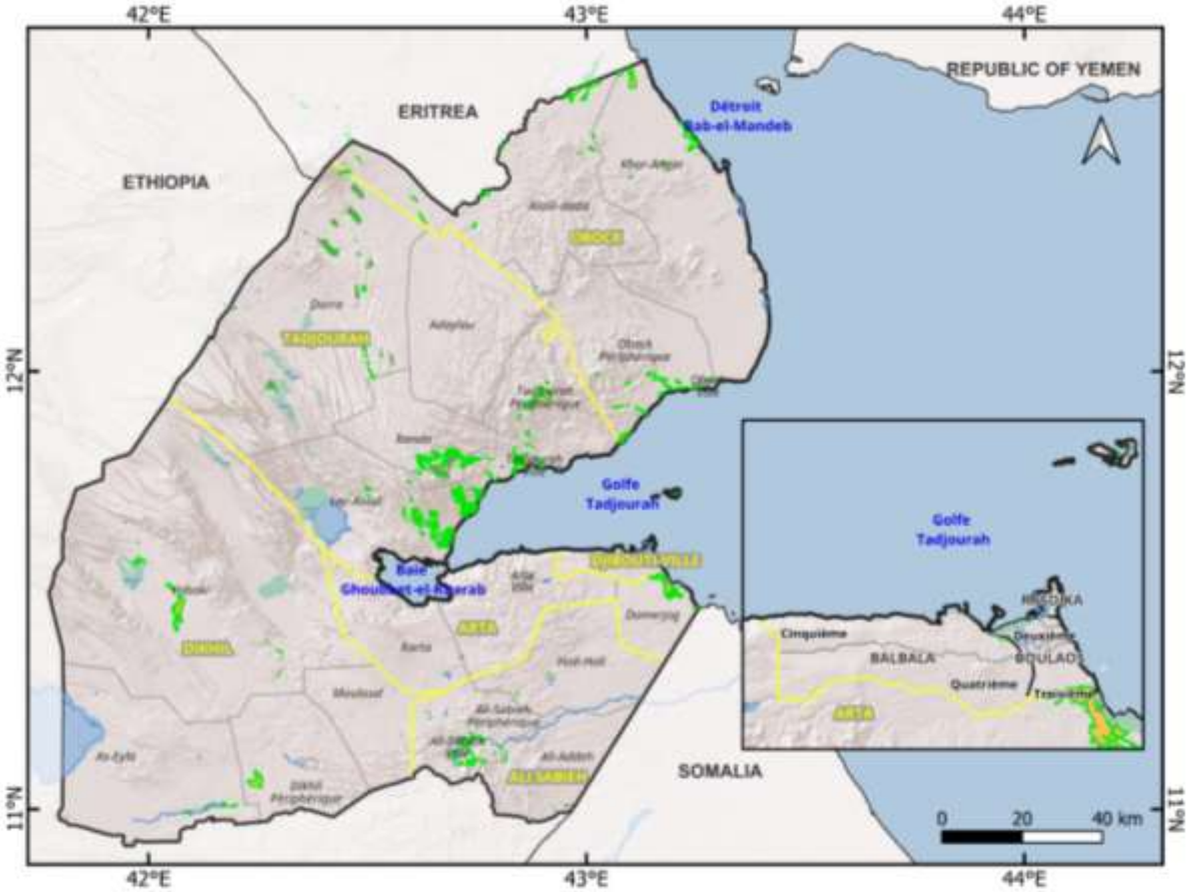

Figure 3. ***Administrative divisions used for exposure and vulnerability analysis.*** *The boundaries were digitalized from RGPH-3 census maps and adjusted to the national and regional reference boundaries available to the project. The layer is a project-derived spatial representation pending validation against an official INSTAD GIS dataset (Source: Annex B – D3 Figure 2).*

Residential buildings

Residential-building exposure was produced through a hybrid methodology combining spatially detailed building footprints with official census information (Figure 4). The principal spatial mask consisted of the building-footprint layer supplied by the national GIS consultant, which itself incorporated OpenBuildingMap-derived (Oostwegel et al., 2025) information and local expert interpretation. RGPH-3 was used as the principal reference for residential building quantities and construction-material distributions, while the Global Earthquake Model (GEM) dataset provided supplementary information such as number of storeys and replacement values where national information was not available.

Individual footprint detections were first filtered to reduce obvious classification noise. D3 reports the exclusion of footprints below 12 m^2 and subsequently applies additional neighbourhood/grid-level filters when creating the final residential model. The retained footprints were aggregated to a regular analysis grid. For each cell, building count and total residential footprint area were calculated. Cells containing only isolated detections or negligible built-up occupation were excluded according to the filtering criteria established in D3.

RGPH-3 prefecture-level construction-type proportions were subsequently allocated to the retained building stock. GEM (2023) storey information was used to estimate floor area where available, with a default single-storey assumption where no overlapping information existed. This generated a spatially explicit residential exposure surface while maintaining a clear distinction between remotely derived geometry, official census composition and supplementary modelled attributes.

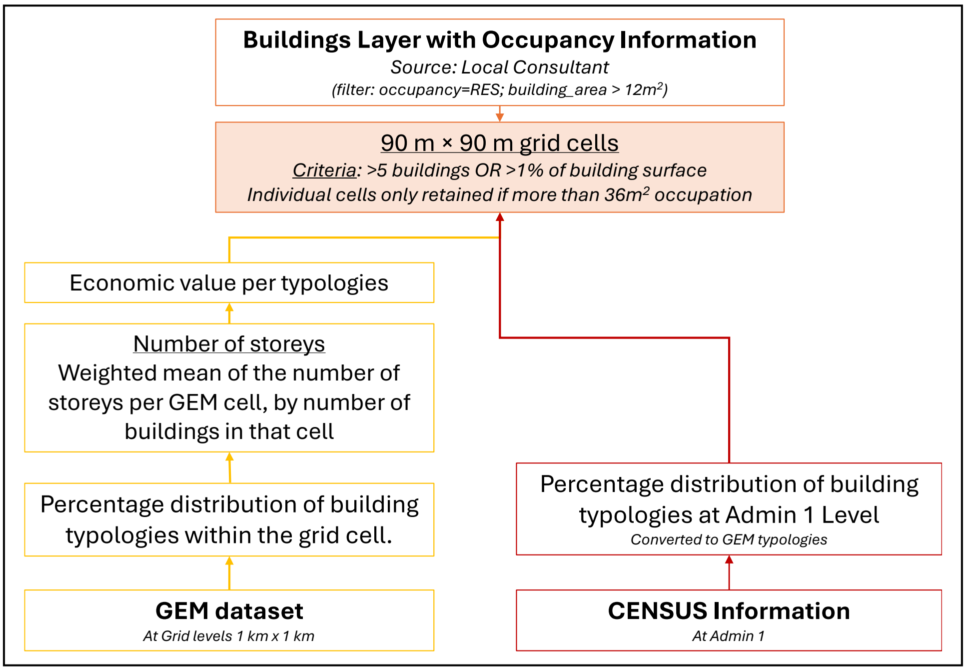

Figure 4. Summary of the methodological approach to create the building exposure model for this project.

Population layers

The 2024 RGPH-3 census constituted the authoritative reference for resident-population totals. Global gridded population datasets such as WorldPop (Bondarenko, 2025) or GHSL (Schiavina et al., 2023) were evaluated as spatial references but were not used to replace census control totals because substantial differences were identified at subnational level.

Population was spatially disaggregated from the census administrative units to the residential-building 90-m grid using a dasymetric approach. Within each sub-prefecture, commune or arrondissement, cells without retained residential buildings received no resident population. The census population was then proportionally distributed among residential cells according to estimated residential floor area, calculated from footprint area and number of storeys. Finally, the allocated values were rescaled so that their sum reproduced exactly the official RGPH-3 population total of the corresponding administrative unit.

The resulting layer is therefore hybrid in nature: population totals remain official census observations, whereas intra-administrative spatial distribution is modelled.

Healthcare facilities, schools and road network

Critical infrastructure was represented primarily through healthcare facilities, education facilities and the road network. The general integration rule was to prioritize official and locally supplied geometries, use open datasets to identify possible gaps or complete attributes, and verify questionable locations against satellite imagery.

Healthcare information combined Ministry of Health statistics, spatial information supplied by the national GIS consultant and OpenStreetMap. Since an updated official geospatial inventory was not available, the locally supplied layer was spatially checked and complemented using open-source information. Facility type and bed capacity were retained where available. The resulting exposure layer (Figure 5) contains 66 georeferenced establishments, while ministry statistics describe a broader total of 81 establishments; the difference reflects both source completeness and differences in institutional scope.

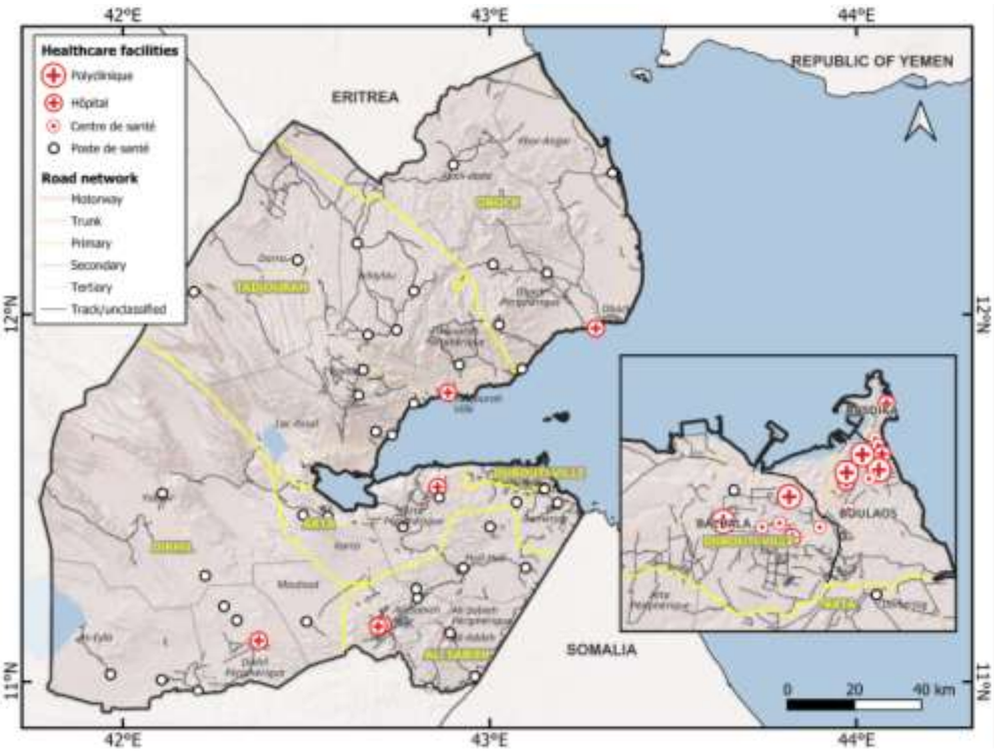

Figure 5. Healthcare facilities by service level. Road network is also shown as a reference for future accessibility studies (Source: Annex B – D3 Figure 14).

Education facilities were principally derived from georeferenced information made available by the Ministry of Education through the MENFOP web application, complemented by the national GIS consultant dataset and selected OpenStreetMap information for higher-education assets. Facility types and student numbers were harmonized to create a common spatial layer. The resulting layer (Figure 6) represents more than 400 establishments and approximately 160’000 students distributed among different education levels.

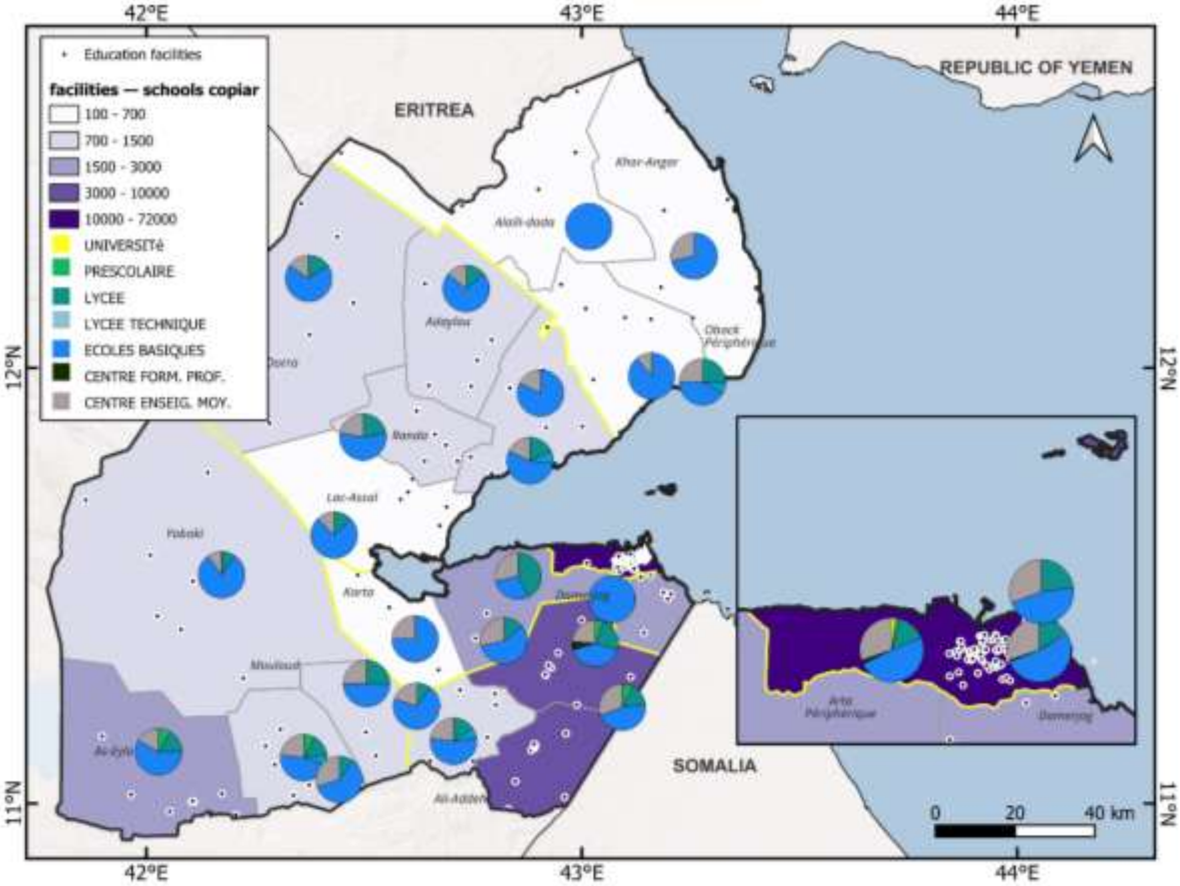

Figure 6. Education facilities as from the data of Ministry of Education (MENFOP). Points indicate the geospatial location, while administrative level aggregations and pie charts are indicative, respectively, of the number of students and the repartition of different types of education establishments (Source: Annex B – D3 Figure 15).

For roads, older institutional datasets were first compared with more recent information and satellite imagery. Because the available CERD road layer dated from 2004 and differed substantially from the current network, the improved OpenStreetMap-based layer supplied by the national GIS consultant was retained as the principal working geometry (Figure 7). The network was harmonized and filtered, considering only motorway, trunk/primary, secondary, tertiary and track/unclassified classes. It was visually checked against recent imagery and the cartography published by the *Agence Djiboutienne des Routes* (Annex B – D3 Figure 16).

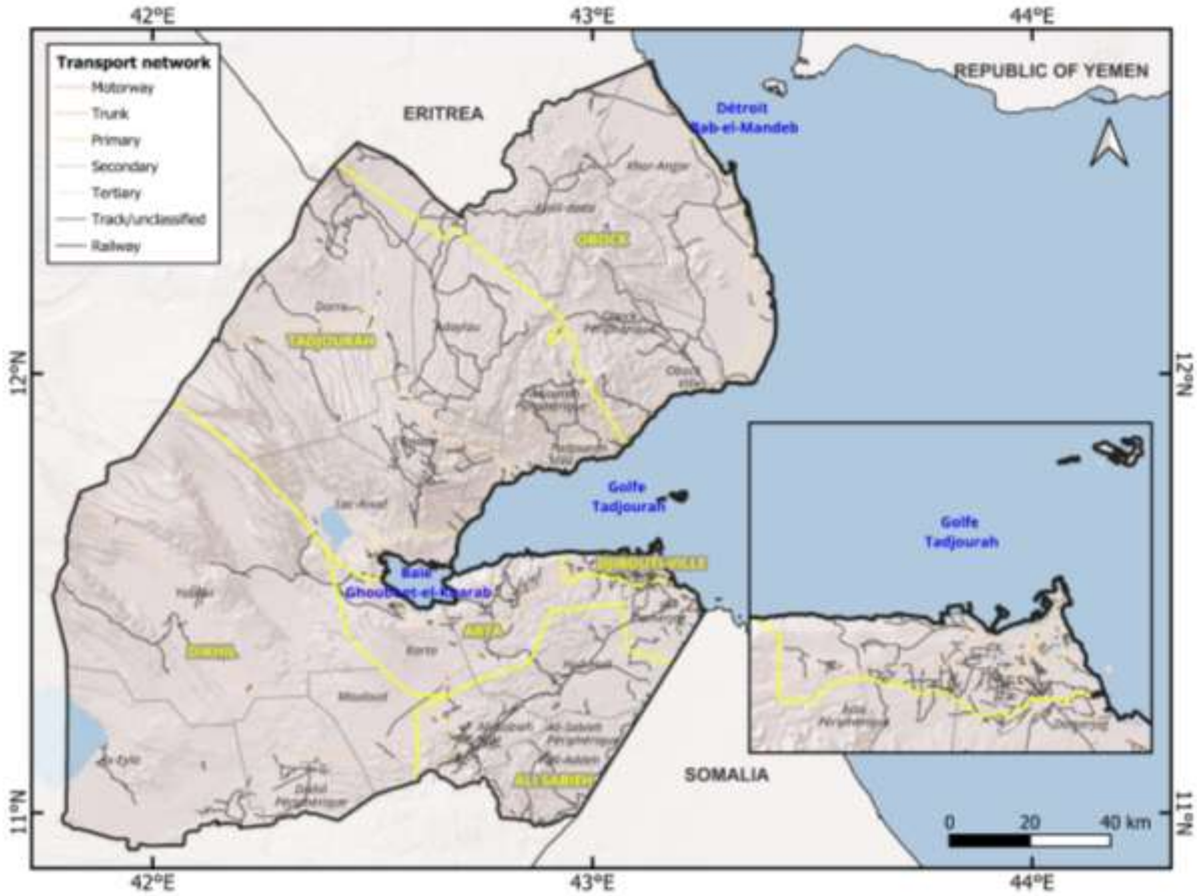

Figure 7. Djibouti transport network derived from local GIS consultant dataset and contrasted against satellite images and OpenStreetMaps (Source: Annex B – D3 Figure 17)

Other exposure layers

Additional exposure layers concerning agriculture (+pastoral systems and fisheries), critical infrastructure (such as antennas locations, power network, water sources) and key ecosystems (mangroves and coral reefs) are developed within Annex B – D3. They are not discussed further in this dossier, which focuses on the geospatial workflows most directly related to the objectives of the Certificate.

3.5. Vulnerability indicators

The vulnerability component presented in this dossier is restricted to the geospatial treatment of social-vulnerability indicators. Physical fragility and vulnerability functions developed for buildings are part of the wider project risk methodology but are not presented as a principal geospatial product of the internship.

Two census-based products were produced: the Population Vulnerability Index (PVI) and the Coping and Warning Capacity Deficit (CWC-D). Both were derived primarily from the 2024 RGPH-3 census.

A central methodological principle was to preserve the genuine observational resolution of each indicator. Regional statistics were not artificially downscaled to finer spatial units simply because finer administrative polygons were available. Where a regional value was represented on sub-prefecture polygons for mapping purposes, it remained explicitly identified as a regional observation. This avoided introducing false spatial precision into the vulnerability analysis.

Population Vulnerability Index

The PVI combines six indicators representing three dimensions of population vulnerability (Figure 8). Socioeconomic Fragility (SE) combines non-monetary poverty and BIT unemployment. Demographic Dependency (DD) combines the demographic dependency ratio and the proportion of population aged 65 years or older. Functional and Information Vulnerability (FI) combines disability prevalence and illiteracy.

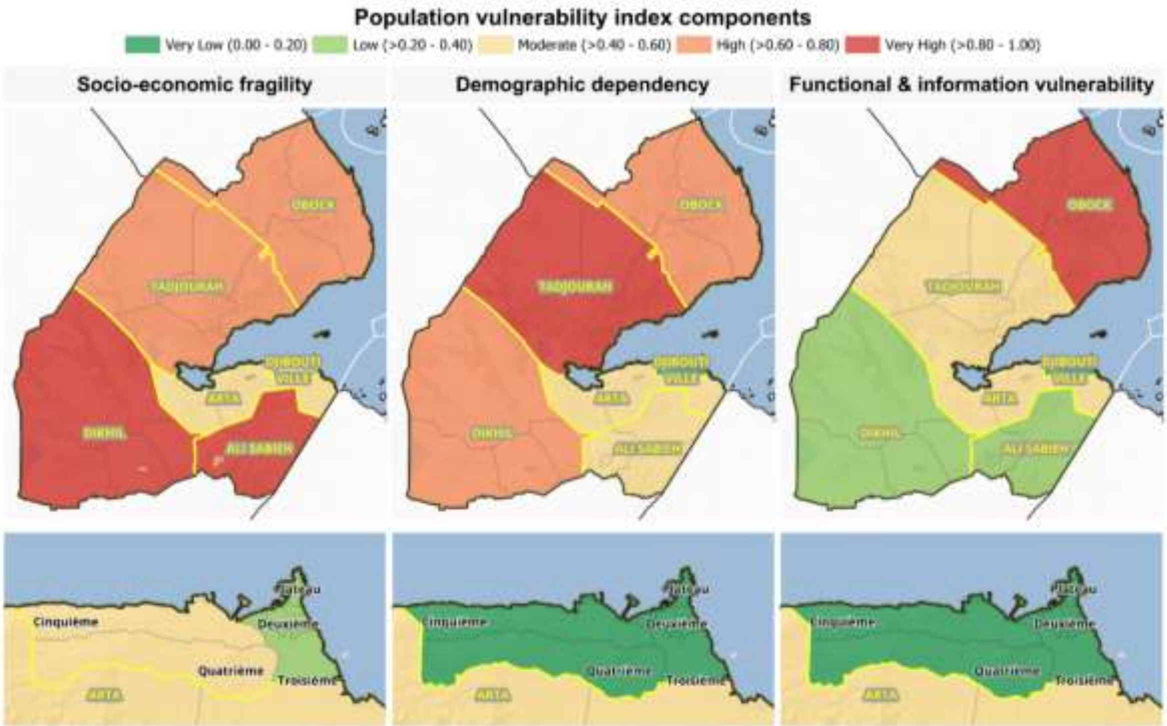

Figure 8. Population vulnerability index components. Left: Socioeconomic fragility combines non-monetary poverty incidence and BIT unemployment. Center: Demographic dependency component combines the demographic dependency ratio and the share of population aged 65 years and over. Right: Functional and information vulnerability combines disability prevalence among the population +5 years and illiteracy among the population +10 years (Source: Annex C – D4 Figure 8)

Each indicator was first oriented so that higher values represented greater vulnerability and was subsequently normalized using min-max scaling:

$$Inorm = \frac{I - I_{min}}{I_{max} - I_{min}}$$

Normalization was performed at the genuine observation scale. Non-monetary poverty was available for 29 prefecture/sub-prefecture/arrondissement units and was therefore normalized across those units. The other five indicators were available at the six-region level and were normalized across those six observations before being associated with their corresponding lower-level polygons for visualization.

The three components were calculated as:

$$SE = \frac{Poverty_{norm} - Unemployment_{norm}}{2}$$

$$DD = \frac{Dependency_{norm} - Pop > 65_{norm}}{2}$$

$$FI = \frac{Disability_{norm} - Illiteracy_{norm}}{2}$$

and the final index as:

$$SE = \frac{SE + DD + FI}{3}$$

This formulation produces an effective equal weight of one sixth for each individual indicator. The resulting index ranges from 0 to 1 and represents relative vulnerability within Djibouti. A score of zero therefore means the lowest observed relative value in the national dataset and not an absence of vulnerability.

For mapping, five fixed intervals were adopted: Very Low (0.00–0.20), Low (>0.20–0.40), Moderate (>0.40–0.60), High (>0.60–0.80) and Very High (>0.80–1.00). D4 Figure 8 illustrates the three PVI components and Figure 9 provides the final mapped index.

Coping and Warning Capacity Deficit

The CWC-D complements the PVI by representing household limitations in access to basic services and conventional communication channels (Figures 10 & 11 in Annex C – D4). Four RGPH-3 variables were used, all expressed as deficits so that higher values consistently represent less favourable conditions: households without access to improved drinking-water sources, households without electricity, households without internet access and households without a mobile telephone.

The normalized water and electricity deficits form the Basic Services Deficit:

$$BSD = \frac{WaterLack_{norm} - ElectricityLack_{norm}}{2}$$

while internet and mobile-phone deficits form the Warning-Channel Deficit:

$$WCD = \frac{InternetLack_{norm} - NoMobile_{norm}}{2}$$

The final index being:

$$CWC = \frac{BSD + WCD}{2}$$

This is equivalent to assigning a weight of 0.25 to each of the four census variables. As with the PVI, values describe relative differences between the six regions and should not be interpreted as absolute probabilities or as a complete representation of last-mile warning capacity.

The index does not directly incorporate mapped telecommunications towers, power lines or water points because physical proximity to infrastructure does not demonstrate household connection, operational status, signal quality or service availability. These layers can instead be used as contextual information.

PVI and CWC-D were deliberately retained as separate products. Population susceptibility and capacity constraints describe related but distinct mechanisms, and averaging both into a single synthetic indicator would conceal potentially important differences between regions.

3.6. Geospatial data integration and organization

Once harmonized, the hazard, exposure and vulnerability layers provide the spatial inputs required for subsequent risk calculations. The first integration step consists of intersecting hazard footprints with exposure layers in order to identify potentially affected people, buildings, or infrastructure.

For riverine flooding, for example, the flood footprint can be intersected directly with the residential-building and population grids and with vector infrastructure. The same general overlay principle can be applied to roads, with exposure expressed as affected length in kilometers.

Heavy rainfall requires a different interpretation since results are spatially continuous rather than discrete areas. Hence, the integration retains the original climatic metric (mm of rainfall) and relates it to relevant exposure, such as population, rather than converting hazard layers to a common binary classification.

The distinction between exposure and risk was therefore preserved throughout the workflow. Spatial overlap demonstrates that an element is located within a potentially affected area; it does not by itself demonstrate damage, disruption or loss. D2 (Annex A) explicitly identifies the intersection of hazard footprints with D3 (Annex B) exposure layers as the first step toward the subsequent Multi-Risk Atlas, while D4 (Annex C) adds the vulnerability information needed to translate this spatial coincidence into potential consequences.

For the purposes of the internship, the final layers were organized as a harmonized geospatial dataset. Consistency was maintained through a common analytical CRS, standardized layer naming and a centralized layer catalogue recording source, temporal reference, processing, provenance, quality-control status and main limitations.

3.7. Data publication: GeoServer and web platform

Selected final layers were published through GeoServer within the project workspace and made available to the WebGIS through WMS services. Publication involved creation or selection of the corresponding data stores, verification of layer names and coordinate reference systems, assignment of predefined cartographic styles and visual testing through the online interface. WFS publication was not required for the internship.

Publication to the web platform is managed through an existing Django interface developed by CIMA's IT team, which facilitates the connection between the WMS service of each layer and the WebGIS interface, using the corresponding layer names (e.g., *pop_grid* in Figure 9).

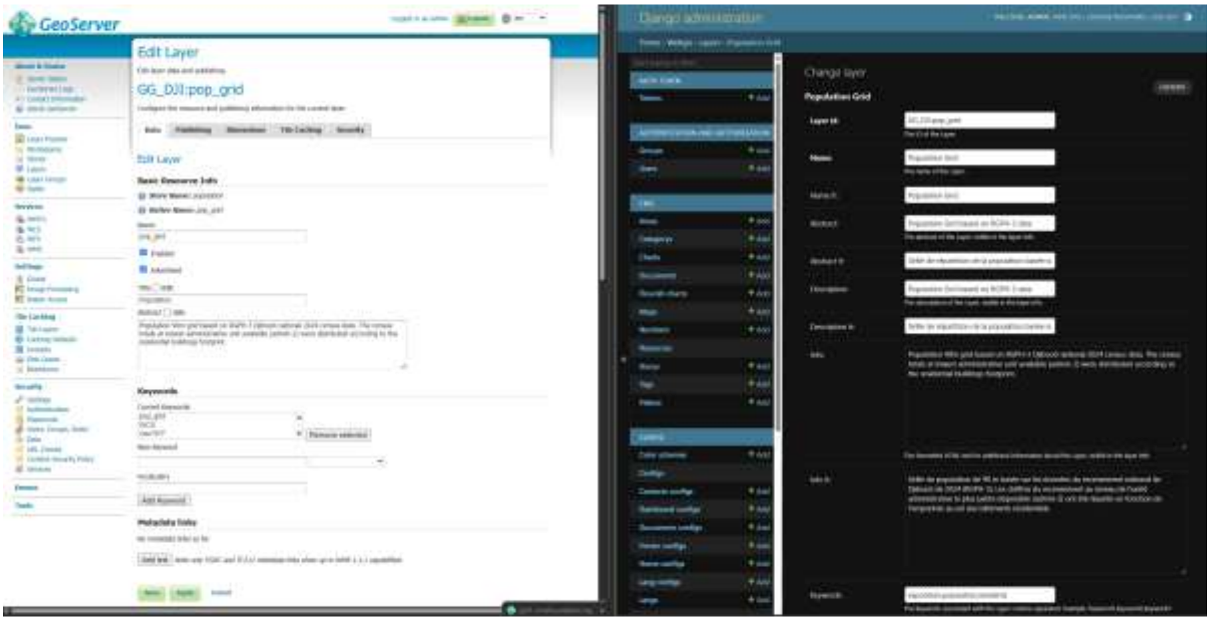

Figure 9. Screenshot of the GeoServer and Django interfaces Screenshot of the GeoServer and Django interfaces used to configure and publish raster and vector layers as WMS services.

4. Results

This section contains a selection of the results contained in the final products delivered that illustrates the main geospatial outputs of the internship. The complete thematic results remain documented in the project deliverables (Annexes D2-D4).

4.1. Geospatial dataset and WebGIS integration

Hence, the priority dataset comprises flood and extreme-rainfall hazard products, population and residential exposure, critical infrastructure and the PVI/CWC-D vulnerability indicators, as detailed in Annex D. The layers were prepared using compatible coordinate systems, common administrative units and progressively standardized naming and provenance information.

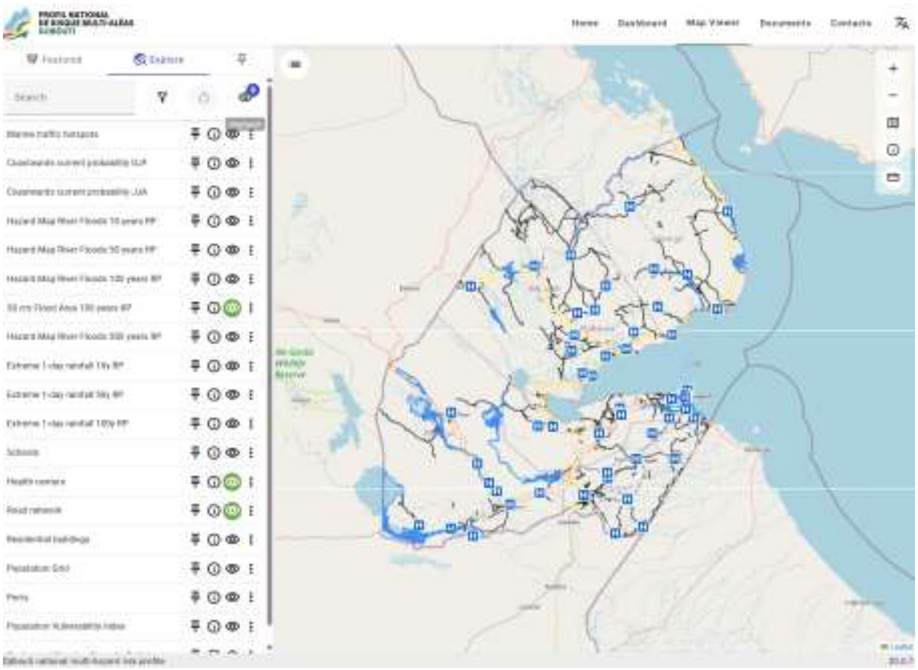

Figure 10. Screenshot from the WebGIS map viewer, with the layers uploaded. Shown is an example with the floodable area (30cm, 100y RP), the road network and hospitals.

The publication workflow provides the connection between spatial analysis and dissemination. Once loaded into GeoServer, the layers can be rendered through standardized web map services and accessed through the project WebGIS without requiring end users to manipulate the original GIS files (Figure 10 and Figure 11)

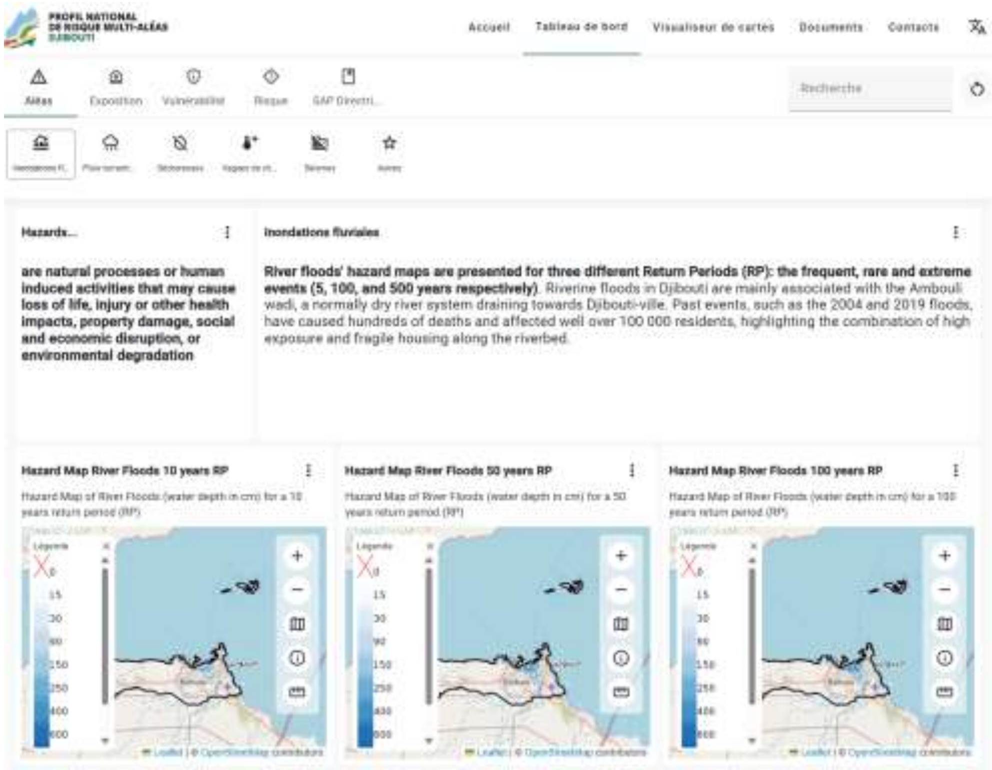

Figure 11. Screenshot of the dashboard (in progress), containing the flood hazard maps for different return periods.

The overall result of the internship can therefore be understood as a progression from heterogeneous source information toward a documented spatial information system: hazard layers were made analytically compatible with exposure; official statistics were translated into spatial exposure products without replacing their control totals; census indicators were transformed into geographically explicit vulnerability products; and the resulting layers were prepared for subsequent risk modelling and WebGIS dissemination.

4.2. Hazard data processing and harmonization

For riverine flooding, existing hydraulic outputs (Figure 1) were transformed into exposure-ready spatial footprints while retaining the original water-depth rasters for subsequent risk calculations. This conversion allowed population, buildings, and infrastructure to be spatially intersected with potentially inundated areas. Particularly, the 100 years RP water-depth map was filtered by 30 cm to create the floodable area footprint, that defines the geographical domain within which exposure calculations can subsequently be performed. The footprint was exported as Boolean raster and converted to a polygon layer, intersected (by Union procedure) by administrative unit to get a final flood exposure result (Section 4.4).

Heavy rainfall represents the main hazard analysis for which a geospatial processing routine was developed directly during the internship, using Google Earth Engine. Annual maximum daily precipitation derived from ERA5 for 1980–2019 was used to estimate pixel-wise Gumbel distributions and extreme 1-day rainfall return levels for RP10, RP50 and RP100. The resulting maps (Figure 12) show increasing rainfall intensity with return period, while maintaining a broadly consistent spatial pattern, with higher values principally in eastern and southeastern Djibouti. The workflow also produced a reproducible procedure that can be rerun when longer or updated meteorological series become available.

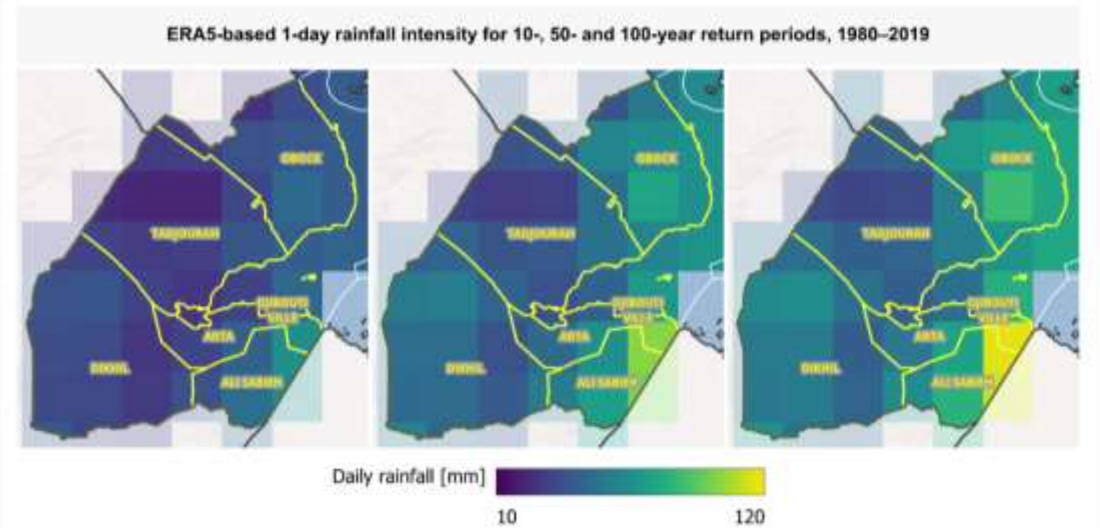

Figure 12. ERA5-based 1-day rainfall intensity for 10-, 50- and 100-year return periods, 1980–2019. Return levels were estimated from annual maximum daily precipitation using a pixel-wise Gumbel distribution fitted by the method of moments. The 50- and particularly the 100-year scenarios involve extrapolation beyond the 40-year analysis period and therefore carry greater statistical uncertainty (Source: Annex A – D2 Figure 13).

The GEE workflow (Annex E.1) demonstrated the use of cloud-based geoprocessing to query, filter and aggregate ERA5 rainfall information over Djibouti and to export spatial outputs for GIS processing and cartography. The geomatics result is therefore both the raster product and the reproducible processing chain used to derive it. In contrast to a map produced through a one-time manual operation, the GEE script can be rerun after changing the temporal period, spatial domain or aggregation statistic. This reproducibility represents one of the main technical outcomes of the hazard-processing component.

4.3. High-resolution exposure modelling

The residential-building and population layers constitute one of the main geospatial results of the internship because they transformed administrative statistics and remotely derived geometries into spatial exposure products suitable for risk analysis.

The final residential exposure layer was aggregated to two regular 90-m grids one containing information on the spatial value distribution of residential buildings and the other the total number of residential buildings per pixel. Figure 13 shows the zonal statistics aggregation at administrative sub-prefecture and communal levels, with the total stock value from residential buildings and, overlapped, the raster layer with the number of units per pixel.

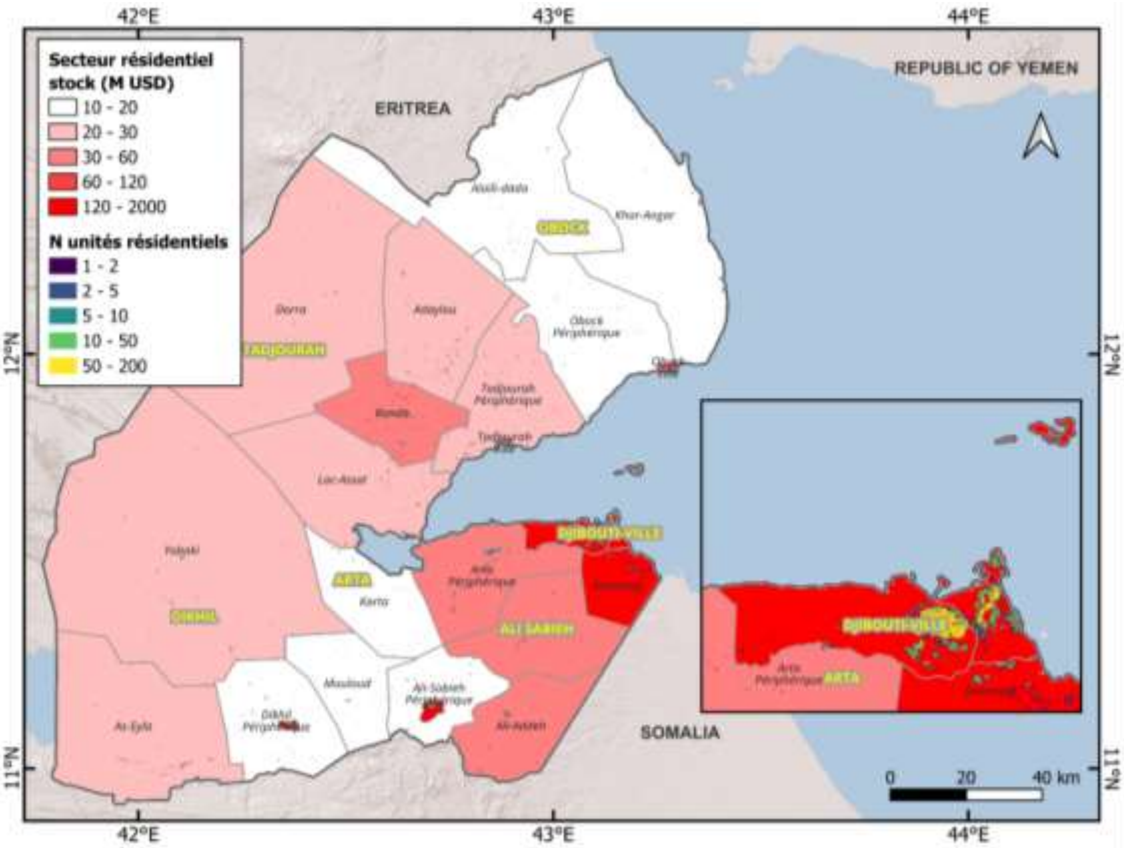

Figure 13. Total exposed asset value (in million USD) aggregated at sub-prefecture level and communal level (in Djibouti-Ville). Over imposed is the 90 m grid residential building layer produced. The number of residential units that fell within each cell are indicated (Source: Annex B – D3 Figure 7).

From a geomatics perspective, the most important result is the integration of information with different properties: individual remote-sensing-derived building geometries, census-based statistical distributions and supplementary model attributes. No individual source was sufficient by itself. The resulting layer is consequently more useful for national risk modelling but must also remain explicitly identified as a hybrid, project-derived product.

The same residential spatial framework was used to produce the population layer. The official RGPH-3 national population of 1,066,809 inhabitants was retained as the demographic control total, while intra-administrative population distribution was modelled using residential floor area. The output therefore preserved the official census totals while substantially increasing their spatial resolution.

The resulting population grid concentrates residents within areas where residential building evidence exists and reduces assignment to ports, industrial sites, airports and open land. This is particularly visible when compared with global gridded population products, which showed important local discrepancies in some parts of Djibouti-Ville and in regional settlements. Figure 14 and Figure 15 illustrate respectively the final population distribution - with the aggregation at administrative levels- and a comparison between the project-derived residential representation and global population products.

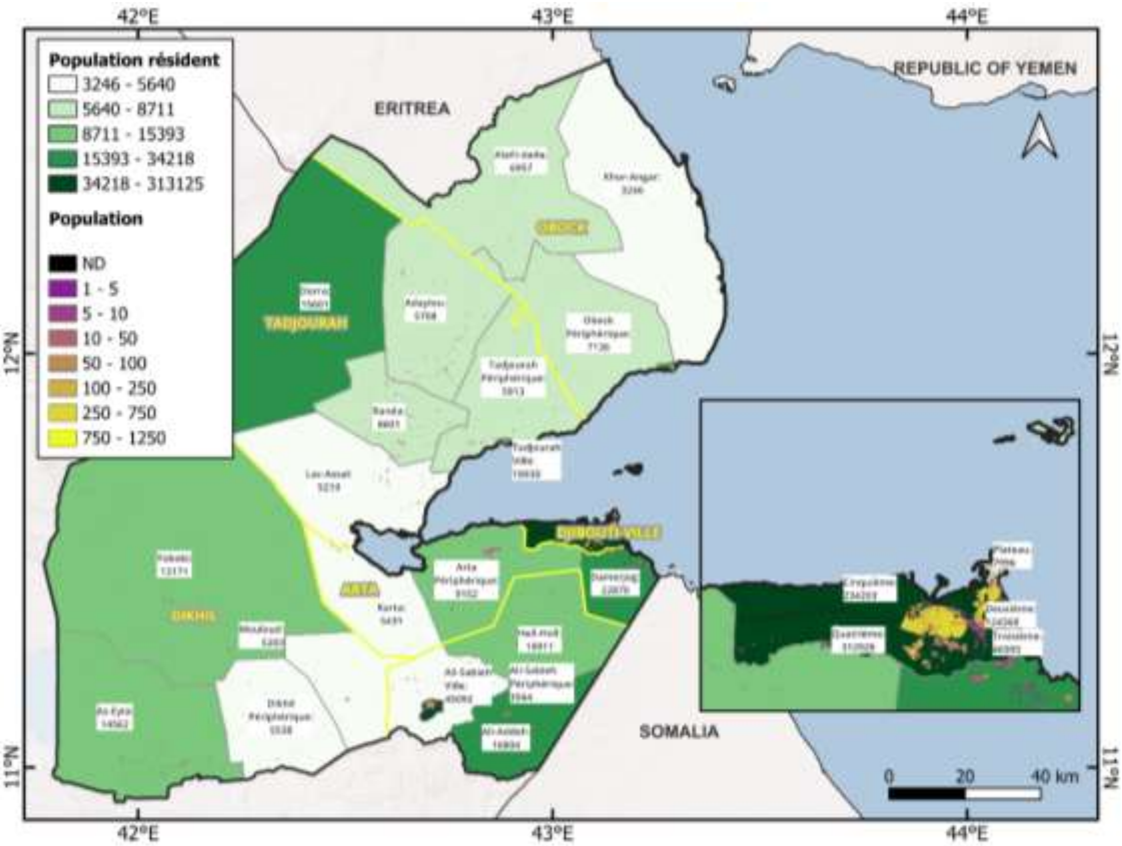

Figure 14. Modelled population distribution based on RGPH-3 administrative totals and the residential building mask. Population was allocated only to grid cells containing retained residential buildings, while the official census total was preserved within each administrative unit (Source: Annex B – D3 Figure 11).

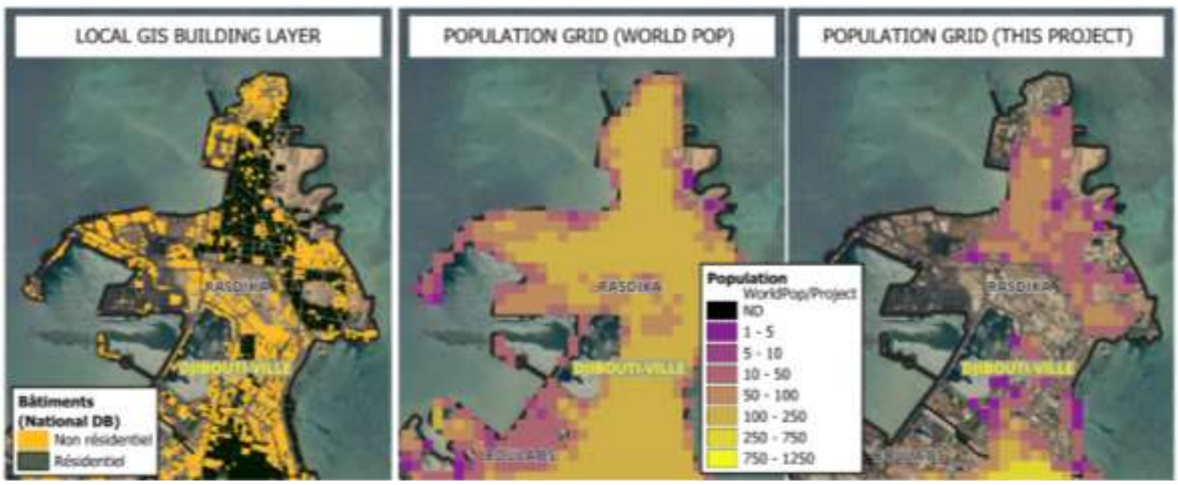

Figure 15. Comparison of residential footprints: polygon building footprint layer provided by the local GIS consultant, WorldPop population grid and the 90-m grid produced for this project. Should be noted that Rasdika region is a highly dense industrial and infrastructure area so WorldPop clearly overestimates the residential population in that region (Source: Annex B – D3 Figure 12).

This result illustrates the application of dasymetric mapping as a practical solution to a common geospatial problem: authoritative statistical information was available at administrative level, while the spatial distribution within those units was unknown. Rather than replacing the census with a global raster, the two types of information were combined according to their respective strengths.

The main limitation remains that the distribution within each census administrative unit is modelled rather than directly observed. The improved spatial detail should therefore not be interpreted as equivalent to individual geolocation of residents.

4.4. Flood exposure assessment

The floodable-area footprint derived from the 100-year return-period scenario, using a minimum water-depth threshold of 30 cm, was intersected with the main exposure layers developed during the internship. Population and residential-building grids were used to estimate exposed population, number of residential buildings and associated residential asset value, while point and line layers were used to identify exposed healthcare and education facilities and affected road length. Results were subsequently summarized by administrative unit (Table 1).

Table 1 shows that exposure is spatially concentrated. For the administrative units reported, approximately 51,615 people and 6,732 residential buildings fall within the RP100 floodable footprint, and *ca.* USD 253.8 million in residential asset value. Most of this exposure is concentrated in Djibouti-Ville (particularly, the Deuxième, Premier and Troisième arrondissements). Outside the capital, exposure is considerably lower but locally relevant: Yoboki presents the highest exposed population.

Table 1. Summary of flood exposure assessment. 100y RP flood, water-depth threshold considered = 30cm. Some administrative units are not shown since flood areas are not distributed unevenly in all units.

REGION	Name of unit	Admin level	Exp. pop	N exp buildings	Res. value exp. value [k US$]	N health units exp.	N edu. units exp.	Road length exp. [km]
ARTA	Damerjog	Sub-prefecture	50	7	276.8	0	0	0.25
DIKHIL	As-Eyla	Sub-prefecture	282	15	135.3	0	0	
	Yoboki	Sub-prefecture	299	34	721.7	1	2	10.7
	Dikhil Périphérique	Sub-prefecture	15	2	11.5	0	0	
	Mouloud	Sub-prefecture	25	4	26.4	0	0	
DJIBOUTI-VILLE	Plateau	Arrondissement	47	35	1140.9	0	0	1.7
	Premier	Arrondissement	14095	1984	71061.5	1	3	2.4
	Deuxième	Arrondissement	22619	1610	78075.5	0	6	5.5
	Troisième	Arrondissement	11007	2545	91096.0	0	4	4.0
	Quatrième	Arrondissement	150	32	414.7	0	0	
	Cinquième	Arrondissement	2783	437	10465.5	0	0	0.9
OBOCK	Alaïli-dada	Sub-prefecture	62	2	39.9	0	0	
	Obock Périphérique	Sub-prefecture	154	2	114.3	0	0	6.3
	Khor-Angar	Sub-prefecture	25	22	239.5	0	0	2.2
TADJOURAH	Dorra	Sub-prefecture	0	0	0.0	0	0	1.0
	Adaylou	Sub-prefecture	2	1	8.7	0	0	

In total, approximately 35 km of road network intersect the flood footprint. Two healthcare facilities are identified as exposed, while no education facilities intersect the footprint according to the harmonized education layer used in this assessment.

These results demonstrate the practical function of the harmonized geospatial dataset: a hydraulic hazard product generated independently of the exposure assessment can be transformed into an exposure-ready footprint and combined consistently with raster, point and line datasets. The resulting values represent potential exposure under the selected RP100 and 30 cm threshold scenario, rather than expected damage or observed losses. They are also dependent on the completeness and spatial accuracy of the underlying exposure layers, particularly the infrastructure inventories.

4.5. Census-based vulnerability mapping

The two main geospatial vulnerability products developed during the internship were the Population Vulnerability Index (PVI) and the Coping and Warning Capacity Deficit (CWC-D). Both were constructed from RGPH-3 information and mapped using the native administrative divisions.

The PVI resulted in a relative national differentiation of socioeconomic, demographic and functional/information vulnerability (Figure 16). Obock presents the highest regional PVI, followed by Tadjourah and Dikhil. The spatial pattern is not driven by the same variables everywhere: Obock is strongly influenced by functional and information vulnerability, Tadjourah by demographic dependency, and Dikhil by socioeconomic fragility. Alaïli-Dada and other units within Obock are recognized among the highest-scoring administrative areas.

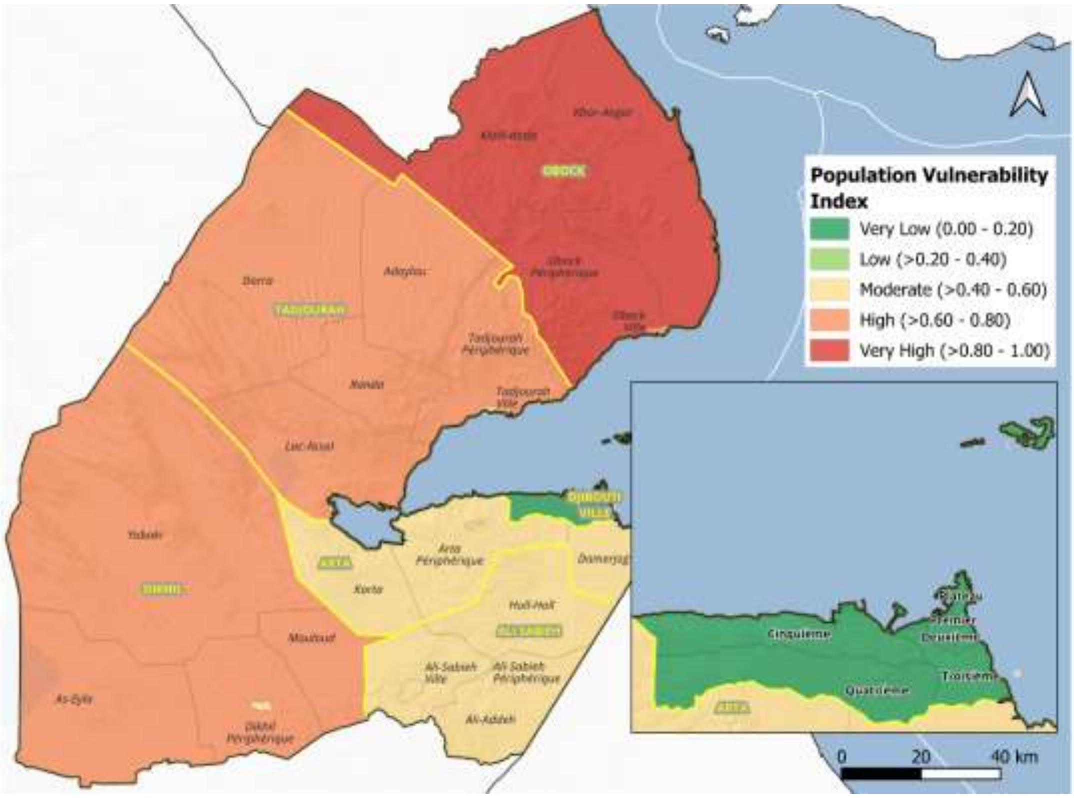

Figure 16. Population Vulnerability Index developed for this mission, based on RGPH-3 (2024) data. The index combines socioeconomic fragility, demographic dependency, and functional/information vulnerability. Non-monetary poverty is observed at prefecture/sub-prefecture/arrondissement level; the remaining indicators are published at regional level. The map is a project-derived mixed-resolution prioritization product and not an official RGPH-3 statistic (Source: Annex C – D4 Figure 9).

The principal geomatics result is not only the composite index but the way its spatial resolution was controlled. Several RGPH-3 indicators are published only at regional level, whereas non-monetary poverty is available at a finer administrative resolution. The mapping therefore preserves this mixed-resolution structure rather than generating artificial local variation through spatial downscaling. In Annex C – D4, is explicitly stated that regional variables were not dasymetrically redistributed because such a transformation would not have statistical meaning.

The CWC-D provides a complementary representation of household deficits in access to water, electricity, internet and mobile telephones. The highest combined deficits occur in Tadjourah and Obock, followed by Dikhil (Figure 17).

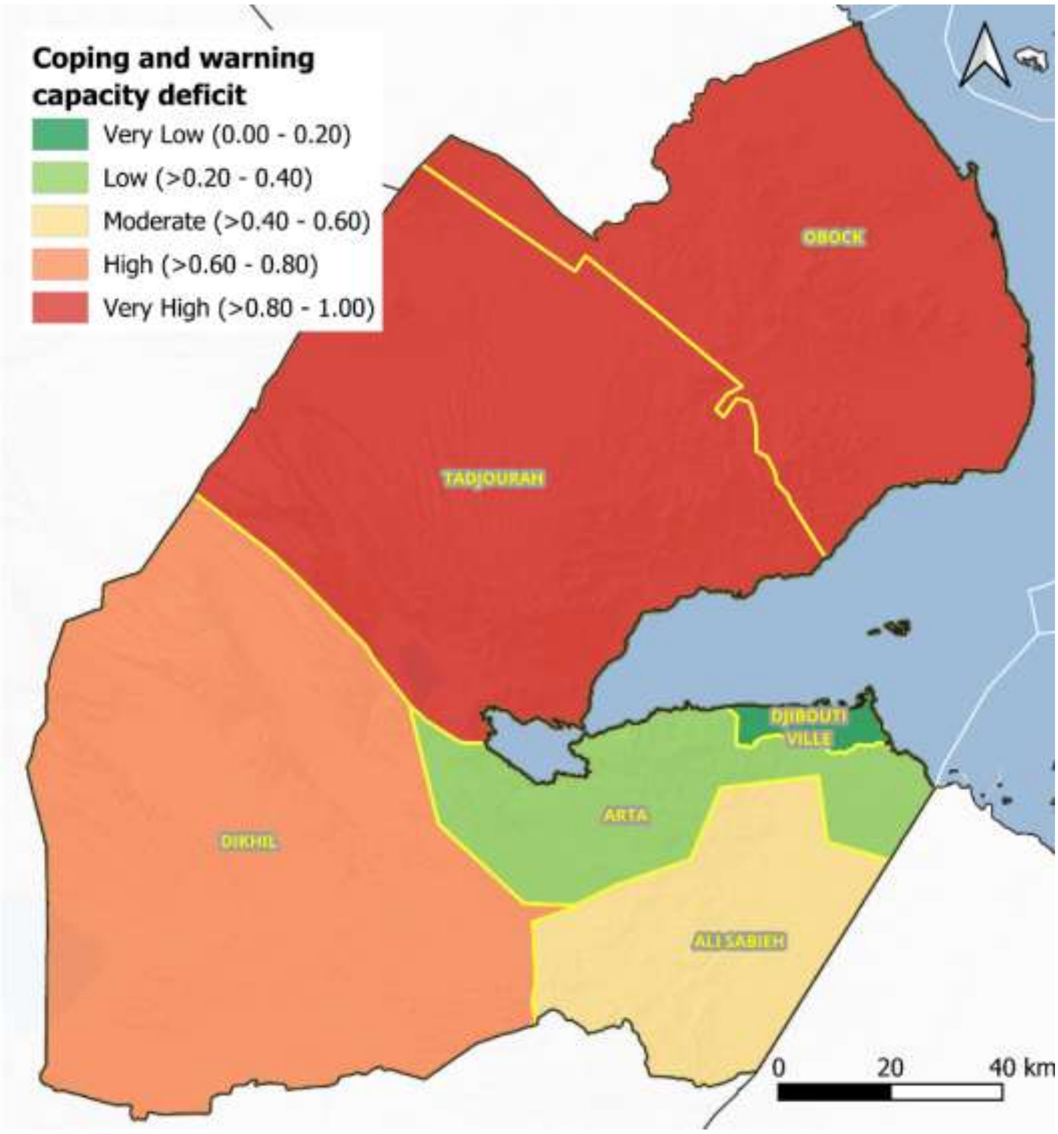

Figure 17. Coping and Warning capacity deficit based on RGPH-3 household access to water, electricity, internet and mobile telephones. Higher values indicate greater relative deficits (Source: Annex C – D4 Figure 12)

These two products also illustrate the limitations of composite cartography. A normalized score represents a relative position within the Djibouti dataset and not an absolute probability of impact. Likewise, mapping a regional indicator over a polygon does not mean that every household inside that polygon has the same condition. Metadata, legends and methodological notes are therefore required to prevent the visual precision of a map from overstating the statistical precision of the source.

5. Final products delivered

The internship contributed to a set of technical and geospatial products developed within the Djibouti multi-risk assessment. The complete project outputs are provided as annexes and are therefore not reproduced in detail in this dossier. The main text only discusses selected examples that illustrate the geomatics work undertaken.

The main formal deliverables are:

D2: Multi-Hazard Scenarios Definition (Annex A - CIMA Research Foundation, 2026a).

The internship contribution focused on the geospatial preparation of selected hazard products and included the development of the extreme-rainfall return-period analysis based on ERA5 data, producing RP10, RP50 and RP100 daily precipitation scenarios. Riverine-flood products were also analyzed to extract relevant floodable area footprints to compute exposure. All of them were subsequently prepared and harmonized for spatial integration and publication.

D3: Exposure Assessment (Annex B - CIMA Research Foundation, 2026b)

This report contains the main exposure products developed during the internship, including population, residential buildings, healthcare facilities, schools, roads, agriculture and water infrastructure, pastoral systems and fisheries, and ecosystems (mangroves and coral reefs). It constitutes the principal documentation of the spatial modelling and harmonization work carried out for the exposure component.

D4: Vulnerability Assessment (Annex C - CIMA Research Foundation, 2026c)

The geospatial contribution presented in this dossier focuses on the census-based Population Vulnerability Index (PVI) and Coping and Warning Capacity Deficit (CWC-D), produced from RGPH-3 indicators while preserving the observational resolution of the source data.

6. Discussion, recommendations and overall internship balance

This section critically assesses the geomatics work carried out during the internship, considering the methodological choices adopted, the principal data and analytical limitations, the quality and interoperability of the resulting products, and the contribution of the Certificate in Geomatics to the work performed within CIMA.

6.1. Methodological strengths

One of the main strengths of the work was the adoption of a Local-Data-First approach. Official national statistics and institutional information were retained as the principal reference whenever available, while global datasets, remote-sensing products and modelled information were used to provide spatial detail or fill specific gaps. This avoided automatically replacing nationally authoritative information with datasets that were spatially finer but potentially less representative of the Djiboutian context.

A second strength was the explicit combination of datasets according to their respective advantages. This was particularly evident in the population and residential-building models. RGPH-3 provided authoritative demographic and housing information, while remotely derived building footprints provided the spatial detail required for disaggregation. The resulting products therefore combined statistical consistency with a substantially finer spatial representation.

The methodology also sought to preserve the actual informational content of the source datasets. In the social-vulnerability assessment, for example, indicators available only at regional level were not artificially downscaled to generate visually more detailed but statistically unsupported spatial patterns. This distinction between cartographic resolution and observational resolution was considered essential to avoid false precision.

Another methodological strength was the adaptation of the spatial representation to the phenomenon being analyzed. Floods could be represented using inundation footprints and water-depth rasters, whereas climatic raster products need to preserve their continuous variables.

Finally, the use of reproducible computational workflows provided an important technical advantage. Google Earth Engine was used to process ERA5 precipitation and derive spatially explicit extreme-rainfall scenarios for 10-, 50- and 100-year return periods. The heavy-rainfall workflow also illustrates the integration of statistical modelling with raster geoprocessing: extreme-value parameters were estimated independently at pixel level and converted into spatial return-period raster layers suitable for GIS and WebGIS use. These procedures reduced dependence on manual processing and provide a basis for updating the analysis when new data become available.

6.2. *Main data and methodological limitations*

The main limitations of the assessment arise from the availability, spatial resolution, temporal consistency and completeness of the source information.

Several official datasets were available only in statistical or non-georeferenced form. This was particularly relevant for RGPH-3 socioeconomic indicators, where most variables required for vulnerability analysis were published only at regional level. Likewise, complete and updated official GIS inventories were not always available for buildings, healthcare facilities, infrastructure or administrative boundaries. Project-derived or open-source spatial datasets therefore had to be used as complementary sources.

Temporal heterogeneity was another limitation. Some spatial layers originated from historical inventories, while other datasets represented recent census conditions, satellite-derived products or contemporary open-source mapping. Although visual inspection and consultation evidence were used to assess their continued relevance, the resulting layers do not represent a single observation date.

Hazard products also differ substantially in their spatial and analytical characteristics. Flooding can be represented by event footprints and depth values, while climatic raster products such as extreme rainfall describe continuous variables with different temporal meanings. Consequently, these hazards cannot be combined directly without additional methodological assumptions. The internship products should therefore be understood as components of a future multi-risk calculation rather than as an integrated multi-risk index.

The rainfall return-period analysis is based on a 40-year ERA5 record. Consequently, RP50 and particularly RP100 involve statistical extrapolation beyond the observation period and carry greater uncertainty. Furthermore, the spatial resolution of ERA5 limits the representation of localized convective storms and topographically controlled precipitation.

Composite vulnerability indicators introduce additional methodological choices. The PVI and CWC-D depend on the selected indicators, their normalization and the weighting scheme. Equal weighting provides transparency and avoids introducing unsupported expert weights, but it does not imply that all vulnerability mechanisms contribute equally to disaster impacts. Similarly, fixed map classes simplify communication but inevitably influence the visual interpretation of continuous index values.

These limitations reinforce an important distinction maintained throughout the assessment: spatial overlap indicates potential exposure, not observed damage or risk. The calculation of consequences requires hazard-specific vulnerability or impact relationships and remains part of subsequent stages of the wider project.

6.3. *Interoperability, data quality and reproducibility*

A central lesson from the internship was that the quality of a geomatics product cannot be assessed exclusively through its spatial accuracy or cartographic appearance. Its usefulness also depends on whether the data can be understood, reproduced, updated and exchanged.

For this reason, provenance and metadata became important elements of the workflow. The project combines official statistics, institutional GIS information, consultant-derived datasets, remote-sensing products, global models and project-derived layers. These categories remain distinguishable in the layer catalogue and associated dataset, to let users understand the origin and level of uncertainty of each product.

A minimum common metadata structure includes source, reference year, coordinate reference system, processing status, methodological origin, validation status and principal limitations. A consistent naming convention and a clear distinction between source, processed and modelled datasets would further reduce ambiguity and facilitate future updates.

Quality assurance could be strengthened through greater automation. Geometry validation, CRS verification, duplicate detection, mandatory-field checks, raster coverage controls and comparison of administrative totals can all be incorporated progressively into scripted procedures. The GEE workflow developed during the internship illustrates the benefits of this approach by making important analytical steps reproducible rather than dependent on undocumented manual operations.

Version control and processing documentation are equally relevant. Where a layer is corrected, reclassified or updated, the relationship with the previous version should remain traceable. This is particularly important in an operational risk-information system where institutional datasets may improve over time.

Interoperability extends these principles to dissemination. Preparing the layers for GeoServer and WebGIS publication requires standardized coordinate systems, consistent names, appropriate styles and sufficient metadata for external users to interpret the products correctly. WMS publication makes it possible to disseminate cartographic information without requiring users to manipulate the original GIS datasets, while the analytical information required for future updates and risk calculations is preserved in the associated catalogue.

Many of these procedures are transferable to other disaster-risk projects. Reprojection, metadata, raster-vector integration, spatial allocation, automated quality control and WebGIS publication are generic geomatics workflows. However, other decisions remain strongly dependent on the local context. Replicability should therefore be understood as the transfer of a documented methodology rather than the mechanical reuse of identical parameters.

6.4. Overall internship balance

The internship provided an opportunity to apply concepts developed during the Certificate in Geomatics within an operational disaster-risk assessment. The contribution of the training was particularly evident in spatial-data management, coordinate reference systems, raster and vector processing, remote-sensing-derived information, spatial analysis, metadata, cloud processing with Google Earth Engine scripting, cartographic representation and WebGIS interoperability.

These competencies contributed directly to CIMA's Djibouti project. Raster and vector processing supported the preparation and harmonization of hazard footprints. Remote-sensing-derived building information and spatial-allocation methods were used to construct higher-resolution population and residential exposure products. Google Earth Engine supported the processing of meteorological raster time series and the production of extreme-rainfall return-period scenarios. Spatial-statistical concepts were applied to the construction and representation of the PVI and CWC-D, and metadata and interoperability principles supported the organization, documentation and preparation of layers for GeoServer and WebGIS publication.

The stage nevertheless showed that professional geomatics involves considerably more than the execution of GIS operations. A substantial part of the work consisted of determining which source should be considered authoritative, whether different datasets could legitimately be combined, which spatial representation was appropriate for a given phenomenon and where the available evidence did not support additional precision. These decisions required interaction with hazard specialists, national consultants and sectoral experts and demonstrated the importance of interdisciplinary work in DRR assessment.

The experience also highlighted the practical tension between analytical production and documentation. Project schedules often prioritize the rapid production of usable layers and maps, while long-term reproducibility requires metadata, structured file organization, processing records and quality-control procedures. The preparation of the layer catalogue and WebGIS made this issue particularly evident.

Overall, the Certificate provided a methodological framework that could be transferred directly to CIMA's operational needs, while the internship added an important professional dimension: working with incomplete and heterogeneous real-world information, evaluating uncertainty, documenting assumptions and adapting geomatics methods to the requirements of different users and disciplines.

Future improvements should focus on increasing automated quality-control procedures, introducing systematic version control, continuing institutional validation of the spatial datasets and maintaining a modular WebGIS architecture. These developments would improve both the reproducibility of the analysis and the capacity of the geospatial products created during the project to remain useful beyond the immediate assessment.

7. References

Blaikie, P., Cannon, T., Davis, I., & Wisner, B. (2004). At Risk: Natural Hazards, People's Vulnerability and Disasters (2nd ed.). London: Routledge.

Bondarenko M., Priyatikanto R., Tejedor-Garavito N., Zhang W., McKeen T., Cunningham A., Woods T., Hilton J., Cihan D., Nosatiuk B., Brinkhoff T., Tatem A., Sorichetta A.. (2025) *Constrained estimates of 2015- 2030 total number of people per grid square at a resolution of 3 arc (approximately 100m at the equator) R2024B version v1.* Global Demographic Data Project - Funded by The Bill and Melinda Gates Foundation (INV-045237). WorldPop - School of Geography and Environmental Science, University of Southampton. DOI:10.5258/SOTON/WP00803

CIMA Research Foundation (2026a). *D2 – Multi-hazard risk scenarios definition*. Development of a National Multi-Risk Profile for the Republic of Djibouti and a Multi-Risk Early Warning System. TIAEWS Project, Ref. UNDP-DJI-00180, August 2026.

CIMA Research Foundation (2026b). *D3 – Exposure layers*. Development of a National Multi-Risk Profile for the Republic of Djibouti and a Multi-Risk Early Warning System. TIAEWS Project, Ref. UNDP-DJI-00180, August 2026.

CIMA Research Foundation (2026c). *D4 – Vulnerability reports*. Development of a National Multi-Risk Profile for the Republic of Djibouti and a Multi-Risk Early Warning System. TIAEWS Project, Ref. UNDP-DJI-00180, August 2026.

Dottori, F., Salamon, P., Bianchi, A., Alfieri, L., Hirpa, F. A. & Feyen, L. (2016). Development and evaluation of a framework for global flood hazard mapping. Advances in Water Resources, 94, 87–102.

Global Earthquake Model Foundation (GEM) (2023). Global Exposure Model, version 2023.1.1. Global Earthquake Model Foundation.

Institut de la Statistique de Djibouti (INSTAD) (2024). 3ème édition du Recensement Général de la Population et de l'Habitat (RGPH-3). Djibouti.

Oostwegel, L. J. N., Schorlemmer, D., Lingner, L. & Evaz Zadeh, T. (2025). OpenBuildingMap. GFZ Data Services. https://doi.org/10.5880/GFZ.LKUT.2025.002

Schiavina, M., Freire, S., Carioli, A. & MacManus, K. (2023). GHS-POP R2023A – GHS population grid multitemporal (1975–2030). European Commission, Joint Research Centre. https://doi.org/10.2905/2FF68A52-5B5B-4A22-8F40-C41DA8332CFE

Silvestro, F., Gabellani, S., Delogu, F., Rudari, R. & Boni, G. (2013). Exploiting remote sensing land surface temperature in distributed hydrological modelling: the example of the Continuum model. Hydrology and Earth System Sciences, 17, 39–62. https://doi.org/10.5194/hess-17-39-2013

United Nations Office for Disaster Risk Reduction (UNDRR) (2017). Sendai Framework Terminology on Disaster Risk Reduction. United Nations Office for Disaster Risk Reduction.

United Nations Office for Disaster Risk Reduction (UNDRR) & CIMA Research Foundation (2024). Handbook on the Use of Risk Knowledge for Multi-Hazard Early Warning Systems. Geneva: UNDRR.

Annexes

Annex A – D2: Multi-Hazard Scenarios Definition

(attached pdf)

Annex B – D3: Exposure Assessment

(attached pdf)

Annex C – D4: Vulnerability Assessment

(attached pdf)

Annex D – Geospatial layer catalogue and WebGIS publication

Annex E – Google Earth Engine workflow for heavy-rainfall return-period analysis

Annex D – Geospatial layer catalogue and WebGIS publication

FILE NAME	Risk component	Type	Main source	Reference period	CRS / resolution	Processing	Main limitation
FLOOD_DEPTH_RP10	Hazard	Raster	JRC flood hazard + CIMA hydrological adjustment	RP10	EPSG:32638; 90 m	Hydraulic output harmonized for project use	Model produced outside internship;
FLOOD_DEPTH_RP50	Hazard	Raster	JRC + CIMA	RP50	EPSG:32638; 90m	Same as above	Same as above
FLOOD_DEPTH_RP100	Hazard	Raster	JRC + CIMA	RP100	EPSG:32638; 90 m	Same as above	Same as above
FLOOD_AREA_RP100	Hazard	Raster	Derived from flood-depth RP100	RP100	EPSG:32638; 90 m	Binary footprint derived from water-depth raster (>30cm)	
FLOOD_AREA_RP100_POLY	Hazard	Polygon	Derived from flood-depth RP100	RP100	EPSG:32638	Binary footprint divided by admin units	
RAINFALL_RX1DAY_RP10	Hazard	Raster	ERA5	1980–2019	EPSG:32638; ~27km	Annual Rx1day →pixel-wise Gumbel →RP10; GEE-derived	ERA5 spatial resolution
RAINFALL_RX1DAY_RP50	Hazard	Raster	ERA5	1980–2019	EPSG:32638; ~27km	Rx1day→Gumbel→RP50	Extrapolation beyond 40-year record
RAINFALL_RX1DAY_RP100	Hazard	Raster	ERA5	1980–2019	EPSG:32638; ~27km	Rx1day→Gumbel→RP100	Greater extrapolation uncertainty
POPULATION_GRID	Exposure	Raster	RGPH-3 + residential floor area	2024	EPSG:32638; 90 m	Dasymetric allocation preserving census totals; modelled spatial distribution	Intra-unit distribution modelled, not observed
EXP_POINTS.GPKG // RES_BUILDINGS	Exposure	Points	Local consultant (OpenBuildings) + RGPH-3 + GEM	Mainly 2024	EPSG:32638; regular point grid 90 m	Footprint filtering + census typologies + floor-area estimation	Hybrid/project-derived layer
EXP_POINTS.GPKG // HEALTHCARE	Exposure	Point	Ministry of Health + Local consultant + OSM	Mixed/latest available	EPSG:32638	Source harmonization	Completeness and update
EXP_POINTS.GPKG // EDUCATION	Exposure	Point	MENFOP + Local consultant + OSM	Latest available	EPSG:32638	Source/attributes/geometry harmonization	Completeness and update
ROAD_NETWORK_DJI	Exposure	Line	Local consultant + OSM + ADR validation	Current working representation	EPSG:32638	Harmonized functional classes; imagery/ADR QC	Local completeness/version dependent/ road types
VULNERABILITY_INDICES // PVI	Vulnerability	Polygon	RGPH-3	2024	EPSG:32638	Min-max normalization; six indicators; equal weights	Mixed observational resolutions; relative index
VULNERABILITY_INDICES // CWC-D	Vulnerability	Polygon	RGPH-3	2024	EPSG:32638	Water/electricity/internet/mobile deficits; equal weights	Regional-scale observations; relative index

Annex E – Google Earth Engine workflow for heavy-rainfall return-period analysis

Script developed for heavy-rainfall return-period analysis, used to derive annual maximum daily precipitation and the RP10, RP50 and RP100 extreme-rainfall scenarios.

```
// ============================================================
// DJIBOUTI - ERA5 DAILY RAINFALL HAZARD CHARACTERIZATION
// Period: 1980-2019
//
// Outputs:
//  1. Mean annual precipitation [mm/year]
//  2. Mean annual maximum 1-day rainfall - Rx1day [mm/day]
//  3. 90th percentile of annual Rx1day [mm/day]
//  4. RP10  = 10-year return-period daily rainfall [mm/day]
//  5. RP50  = 50-year return-period daily rainfall [mm/day]
//  6. RP100 = 100-year return-period daily rainfall[mm/day]
//
// Dataset: ECMWF/ERA5/DAILY,
// Resolution: 27.8 km
// ============================================================


// ============================================================
// 1. DJIBOUTI BOUNDARY
// ============================================================

var countries = ee.FeatureCollection('FAO/GAUL/2015/level0');

var djibouti = countries.filter(
  ee.Filter.eq('ADM0_NAME', 'Djibouti')
);

Map.centerObject(djibouti, 7);

Map.addLayer(
  djibouti.style({
    color: '000000',
    fillColor: '00000000',
    width: 2
  }),
  {},
  'Djibouti boundary'
);


// ============================================================
// 2. SETTINGS
// ============================================================

var START_YEAR = 1980;
var END_YEAR   = 2019;

var years = ee.List.sequence(START_YEAR, END_YEAR);


// ============================================================
// 3. LOAD ERA5 DAILY PRECIPITATION
// ============================================================

// ERA5 precipitation is in m
// Convert to mm

var rainfall = ee.ImageCollection('ECMWF/ERA5/DAILY')
  .filterDate(
    START_YEAR + '-01-01',
    (END_YEAR + 1) + '-01-01'
  )
  .select('total_precipitation')
  .map(function(image) {

    return image
```

```
      .multiply(1000)
      .rename('precip_mm')
      .copyProperties(
        image,
        ['system:time_start']
      );

  });

print('ERA5 daily rainfall collection:', rainfall);
print('Number of daily images:', rainfall.size());


// ===============================================================
// 4. MAIN PP INDICES
// ===============================================================

// ---------------------------------------------------------------
// 4.1 Mean annual precipitation
// ---------------------------------------------------------------

var annualPrecip = ee.ImageCollection.fromImages(

  years.map(function(year) {

    year = ee.Number(year);

    var start = ee.Date.fromYMD(year, 1, 1);
    var end = start.advance(1, 'year');

    return rainfall
      .filterDate(start, end)
      .sum()
      .rename('annual_precip_mm')
      .set('year', year)
      .set('system:time_start', start.millis());

  })

);

var meanAnnualPrecip = annualPrecip
  .mean()
  .clip(djibouti);


// ---------------------------------------------------------------
// 4.2. Maximum daily rainfall / year
// ---------------------------------------------------------------

var annualRx1day = ee.ImageCollection.fromImages(

  years.map(function(year) {

    year = ee.Number(year);

    var start = ee.Date.fromYMD(year, 1, 1);
    var end = start.advance(1, 'year');

    return rainfall
      .filterDate(start, end)
      .max()
      .rename('Rx1day_mm')
      .set('year', year)
      .set('system:time_start', start.millis());

  })

);

// Mean annual maximum daily rainfall

var meanRx1day = annualRx1day
  .mean()
  .clip(djibouti);
```

```
// 90th percentile of annual maximum daily rainfall

var p90Rx1day = annualRx1day
  .reduce(
    ee.Reducer.percentile([90])
  )
  .rename('Rx1day_P90_mm')
  .clip(djibouti);

// ------------------------------------------------------------
// 4.3. Print basic national statistics
// ------------------------------------------------------------

var statsRx1 = meanRx1day.reduceRegion({
  reducer: ee.Reducer.minMax()
    .combine({
      reducer2: ee.Reducer.mean(),
      sharedInputs: true
    }),
  geometry: djibouti.geometry(),
  scale: 27830,
  maxPixels: 1e9
});

print('Mean Rx1day statistics - Djibouti:', statsRx1);


// ============================================================
// 5. EXTREME-VALUE ANALYSIS
//     Gumbel distribution fitted to annual Rx1day maxima
//
//     Input:
//         annualRx1day
//
//     Period:
//         1980-2019 = 40 annual maximum values per grid cell
//
//     Outputs: 1-day extreme rainfall for RP 10, 50, 100 y
// ============================================================

// ------------------------------------------------------------
// 5.1 -- Mean and standard deviation of annual maxima
// ------------------------------------------------------------

var rxMean = annualRx1day
  .mean()
  .rename('Rx1day_mean');

var rxStd = annualRx1day
  .reduce(ee.Reducer.stdDev())
  .rename('Rx1day_std');

// ------------------------------------------------------------
// 5.2 -- Gumbel parameters using method of moments
//
// beta = sigma * sqrt(6) / pi
//
// mu = mean - gamma * beta
//
// gamma = Euler-Mascheroni constant = 0.5772156649
// ------------------------------------------------------------

var eulerGamma = 0.5772156649;

var beta = rxStd
  .multiply(Math.sqrt(6) / Math.PI)
  .rename('Gumbel_beta');

var mu = rxMean
  .subtract(beta.multiply(eulerGamma))
  .rename('Gumbel_mu');
```

```
// ------------------------------------------------------------
// 5.3 Function to calculate rainfall for return period T
//
// F = 1 - 1/T
//
// x(T) = mu - beta * ln[-ln(F)]
// ------------------------------------------------------------

function gumbelReturnLevel(T) {

  var F = 1 - (1 / T);

  var reducedVariate =
    -Math.log(-Math.log(F));

  return mu
    .add(beta.multiply(reducedVariate))
    .rename('RP' + T + '_mm')
    .clip(djibouti);
}

// ------------------------------------------------------------
// 5.4 Calculate return-period rainfall
// ------------------------------------------------------------

var RP10  = gumbelReturnLevel(10);
var RP50  = gumbelReturnLevel(50);
var RP100 = gumbelReturnLevel(100);

// ------------------------------------------------------------
// 5.5 Print national min / mean / max values / RP
// ------------------------------------------------------------

function printStats(image, name) {

  var stats = image.reduceRegion({
    reducer: ee.Reducer.minMax()
      .combine({
        reducer2: ee.Reducer.mean(),
        sharedInputs: true
      }),
    geometry: djibouti.geometry(),
    scale: 27830,
    maxPixels: 1e9
  });

  print(name, stats);
}

printStats(RP10,  '10-year rainfall return level [mm/day]');
printStats(RP50,  '50-year rainfall return level [mm/day]');
printStats(RP100, '100-year rainfall return level [mm/day]');


// ============================================================
// 6. VISUALIZATION
// ============================================================

var annualVis = {
  min: 50,
  max: 400,
  palette: [
    'fff7ec',
    'fee8c8',
    'fdd49e',
    'fdbb84',
    'fc8d59',
    'ef6548',
    'd7301f',
    '990000'
  ]
};

var rx1Vis = {
  min: 10,
```

```
  max: 60,
  palette: [
    'ffffcc',
    'c7e9b4',
    '7fcdbb',
    '41b6c4',
    '2c7fb8',
    '253494'
  ]
};

var rpVis = {
  min: 20,
  max: 80,
  palette: [
    'ffffcc',
    'c7e9b4',
    '7fcdbb',
    '41b6c4',
    '2c7fb8',
    '253494'
  ]
};


// Add maps

Map.addLayer(
  meanAnnualPrecip,
  annualVis,
  'Mean annual precipitation 1980-2019',
  false
);

Map.addLayer(
  meanRx1day,
  rx1Vis,
  'Mean annual Rx1day 1980-2019',
  true
);

Map.addLayer(
  p90Rx1day,
  rx1Vis,
  'P90 annual Rx1day 1980-2019',
  false
);

Map.addLayer(
  RP10,
  rpVis,
  'Rainfall RP10 - Gumbel',
  true
);

Map.addLayer(
  RP50,
  rpVis,
  'Rainfall RP50 - Gumbel',
  false
);

Map.addLayer(
  RP100,
  rpVis,
  'Rainfall RP100 - Gumbel',
  false
);


// ============================================================
// 7. EXPORT RASTERS TO GOOGLE DRIVE
// ============================================================

// Native ERA5 DAILY pixel size is approximately 27.8 km.
```

```
Export.image.toDrive({
  image: meanAnnualPrecip,
  description: 'DJI_ERA5_MeanAnnualPrecip_1980_2019',
  folder: 'ERA5_Djibouti',
  fileNamePrefix: 'DJI_ERA5_MeanAnnualPrecip_1980_2019',
  region: djibouti.geometry().bounds(),
  scale: 27830,
  maxPixels: 1e13,
  fileFormat: 'GeoTIFF'
});


Export.image.toDrive({
  image: meanRx1day,
  description: 'DJI_ERA5_Mean_Rx1day_1980_2019',
  folder: 'ERA5_Djibouti',
  fileNamePrefix: 'DJI_ERA5_Mean_Rx1day_1980_2019',
  region: djibouti.geometry().bounds(),
  scale: 27830,
  maxPixels: 1e13,
  fileFormat: 'GeoTIFF'
});


Export.image.toDrive({
  image: p90Rx1day,
  description: 'DJI_ERA5_P90_Rx1day_1980_2019',
  folder: 'ERA5_Djibouti',
  fileNamePrefix: 'DJI_ERA5_P90_Rx1day_1980_2019',
  region: djibouti.geometry().bounds(),
  scale: 27830,
  maxPixels: 1e13,
  fileFormat: 'GeoTIFF'
});


Export.image.toDrive({
  image: RP10,
  description: 'DJI_ERA5_Rainfall_RP10_1980_2019',
  folder: 'ERA5_Djibouti',
  fileNamePrefix: 'DJI_ERA5_Rainfall_RP10_1980_2019',
  region: djibouti.geometry().bounds(),
  scale: 27830,
  maxPixels: 1e13,
  fileFormat: 'GeoTIFF'
});


Export.image.toDrive({
  image: RP50,
  description: 'DJI_ERA5_Rainfall_RP50_1980_2019',
  folder: 'ERA5_Djibouti',
  fileNamePrefix: 'DJI_ERA5_Rainfall_RP50_1980_2019',
  region: djibouti.geometry().bounds(),
  scale: 27830,
  maxPixels: 1e13,
  fileFormat: 'GeoTIFF'
});


Export.image.toDrive({
  image: RP100,
  description: 'DJI_ERA5_Rainfall_RP100_1980_2019',
  folder: 'ERA5_Djibouti',
  fileNamePrefix: 'DJI_ERA5_Rainfall_RP100_1980_2019',
  region: djibouti.geometry().bounds(),
  scale: 27830,
  maxPixels: 1e13,
  fileFormat: 'GeoTIFF'
});
```