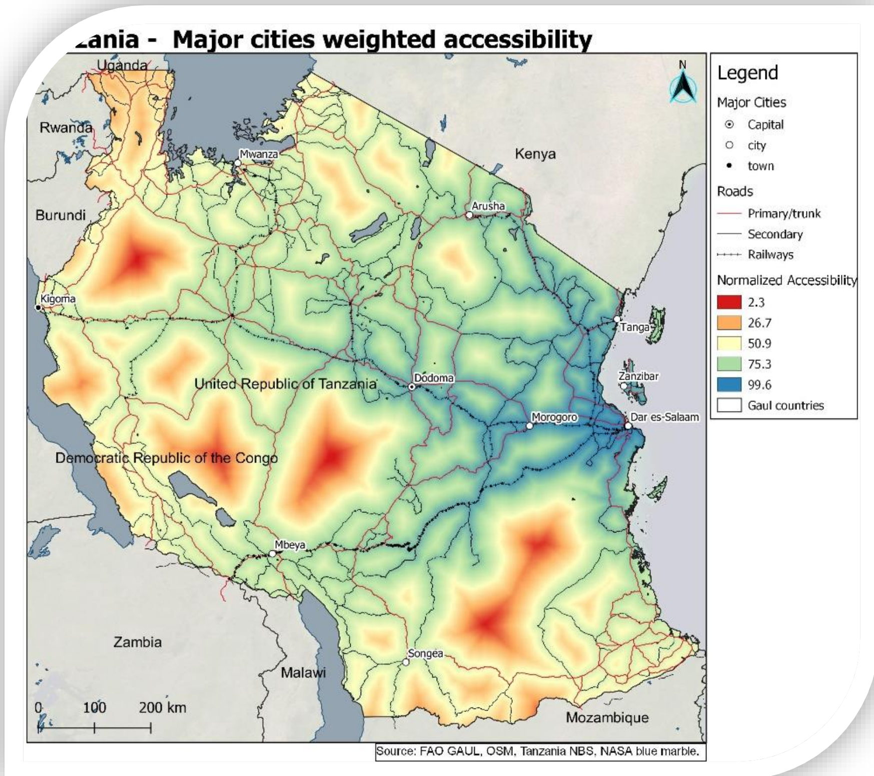
2021



GIS Multicriteria Decision Analysis - Tanzania

Crop storage location

NELSON RIBEIRO

Table of contents

Figures index

Introduction

This report documents a Geographical Information Systems - Multicriteria Decision Analysis (GIS-MCDA) aimed at the identification and definition of mobile storage (movable warehouses) sites in Tanzania, for major crop commodities, maize, cassava, and rice.

Tanzania GIS-MCDA follows previous analysis for Ghana and Kenya introducing few changes in the methodology:

1. Accessibility to cities, both in the country and region, is weighted using the city population size.
2. Commodities are not clustered or grouped – In previous exercises the location analysis was developed for root and tubers, cereal, or tree crops.
3. Top score maps - changed from a top percentage of the score to using the 99th percentile.
4. Final exclusive criteria access to finance - banks – increased the buffer size from 10 to 20km.

The modeling assumes a spatial or GIS-MCDA methodology using weighted factors (AscoughII et al., 2019; Boroushaki & Malczewski, 2010; Malczewski, 2006).

Modeling variables characterize supply, demand and infrastructure/accessibility, main logistical factors for warehousing facilities location. The criteria are:

1. Supply - Major crops - production areas.
2. Demand - Human population density, large cities.
3. Infrastructure - Transportation network (accessibility).

The main transportation network infrastructure is modeled as raster-based travel time/cost analysis (Mulrooney et al., 2017).

Accessibility/infrastructure travel time/cost to markets is computed for major maritime ports, national and regional large cities. Access to finance - distance to bank agencies – and major road proximity are complemented by access to IT access - mobile broadband coverage - to filter location score grids into recommend areas of interest.

The project is developed using open-source GIS software QGIS 3.10.12-A Coruña with open-data sources OpenStreetMap and FAO GIS platform, mobile broad band coverage data is derived from Collins Bartholomew's Mobile Coverage Explorer datasets.

This document is structured in 4 main sections: 1. Initial Specification; 2. Data Sources; 3. Vector Data Preprocessing/Editing; 4. Data Editing Geoprocessing, with Introduction and Closing Remarks. In Annexes there are: project flow diagram, algorithm diagrams and complementary map outputs.

GIS Multicriteria Decision Analysis

Spatial decision problems involve a set of geographically defined alternatives and multiple and sometimes opposing assessment criteria. Alternatives are commonly assessed by many intervenient (decision-makers, stakeholders, interest groups).

GIS multicriteria decision analysis GIS-MCDA consists of a method to convert and combine spatial data/geographical information and decision-makers criteria to attain evidence for a decision-making process. GIS capabilities are enhanced by MCDA procedures, techniques, and algorithms for structuring decision problems, to design, evaluate and prioritize alternatives.

Integration of GIS and MCDA provides a replicable model, improves communication between project participants or decision-makers, can offer a different perspective of problem and solution, helping to redefine initial specification and/or criteria.

GIS multicriteria analysis methods are usually presented in a three-stage hierarchy of: intelligence, design, and choice.

In the intelligence phase, data are acquired, processed, and exploratory data analysis is performed.

The design phase should entail the formal modeling/GIS interaction development of a solution set of spatial decision alternatives. The integration of decision analytical procedures and GIS functions is critical for supporting the design phase.

The choice phase involves selecting location alternatives from those available. Specific decision rules are used to evaluate and rank alternatives.

The three stages of decision making do not necessarily follow a linear path from intelligence, to design, and to choice.

In the case of the present exercise we would place it in the first stage of this hierarchy, the exploratory/intelligence phase.

From a critical standpoint it can be stated that, while quantitative data analysis and evidence gathering through GIS modeling certainly contributes to attaining evidence for decision-making processes, in reality, a complex set of socio-economic, political, cultural, ethno-anthropological aspects and power relations actually shape processes and govern decision-making.

Modeling is also as good as the input data, its quality and reliability support the extent to which conclusions can be trusted, and these are just as sound as the analysis conducted. From that prism, specification and objectives define the modeling assumptions and approximations and can always produce distinct answers (Kitchin, 2014b).

Data are both social and material do not just represent the world but can actively produce it, are not mere raw material of information and knowledge, do not exist independently of ideas, techniques, technologies, people and contexts that produce, process, manage, analyze and store it. Positionality is always present even when “data speaks for itself” (Kitchin, 2014a).

1. Initial Specification

The general data dimensions specified were the following:

1. Infrastructure dimension
    a. average lights.
    b. primary roads.
    c. railways.
    d. Airports.
    e. bridges.
    f. waterways.
2. IT access
    a. internet access.
    b. cellphone signal.
    c. broadband information.
3. Access to finance
    a. bank locations.
4. Market access
    a. cities and travelling time to cities.
    b. time to ports.
    c. port waiting times.
5. Population density
    a. Socioeconomic information maybe included if available at subnational level (poverty maps for example).
6. Production dimension
    a. Cereals and vegetables production.
    b. Crop calendars and plant phenology.

2. Data Sources

The following sources are used in the analysis:

1. ***Human Population Density 2020*** – WorldPop2020 - Estimated total number of people per grid-cell 1km. https://www.worldpop.org/geodata/summary?id=24777;
2. ***Mapspam Production*** – IFPRI's Spatial Production Allocation Model (SPAM) estimates of crop distribution within disaggregated units (Wood-Sichra et al., 2016). https://dataverse.harvard.edu/dataverse/harvestchoice
    - Production (mt);
3. ***OpenStreetMap* -** Map of the world built by volunteers and released with open-content license. Community mapping using wiki-style collaborative editing software. The data was downloaded in ESRI SHP file format for the selected countries available at: http://download.geofabrik.de/. (Ramm, 2019)
4. ***FAO*** – FAOStat, Rivers of Africa, Inland Waters of Africa, Airports, Ports;
5. ***Collins Bartholomew*** - Mobile Coverage Explorer raster data representation of the area covered by mobile cellular networks around the world. The dataset series is supplied as raster Data_MCE (operators) and Data_OCI (OpenCellID database).

3. Vector Data Preprocessing/Editing

This step is partially automated and includes data gathering/download, selection and edition by location and attribute, and the creation of a country database *vector geopakage*[1]:

1. ***OSM Road layer (gis_osm_roads_free_1.shp)*** – Selected by location and attribute to generate a major roads layer input. A comprehensive description of the features can be found in (Ramm, 2019). Lack of data on road network conservation, quality and speed limit for most of the network imposes a conservative approach.

    1.1 Attributes - 'motorway'; 'trunk'; 'primary', 'secondary'.

2. ***OSM Railways (gis_osm_railways_free_1.shp)*** – Selected by location.

3. ***OSM Point-of-Interest Layer (gis_osm_pois_a_free_1.shp)*** – Select by attributes to generate layer Banks.

    3.1 Attributes: 'bank'.

4. ***OSM Transport Layer (gis_osm_transport_free_1.shp)***

    3.1 'ferry_terminal'; 'railway_station'; 'railway_halt'.

5. ***OSM Places Layer (gis_osm_places_free_1.shp)*** - Selected by attribute to generate major human settlements layer.

    5.1 Attributes: 'city'; 'town'; 'national_capital'.

6. ***FAO Data - Ports; Airports; Secondary Airports*** - csv file formats - FAO http://rkp.review.fao.org/geonetwork – selected by location for the country.

7. ***FAO Major rivers*** - Rivers of Africa derived from the World Wildlife Fund's (WWF).

[1] http://www.geopackage.org/

8. ***FAO Inland Waters*** -

Data is edited extracted/clipped using country borders.

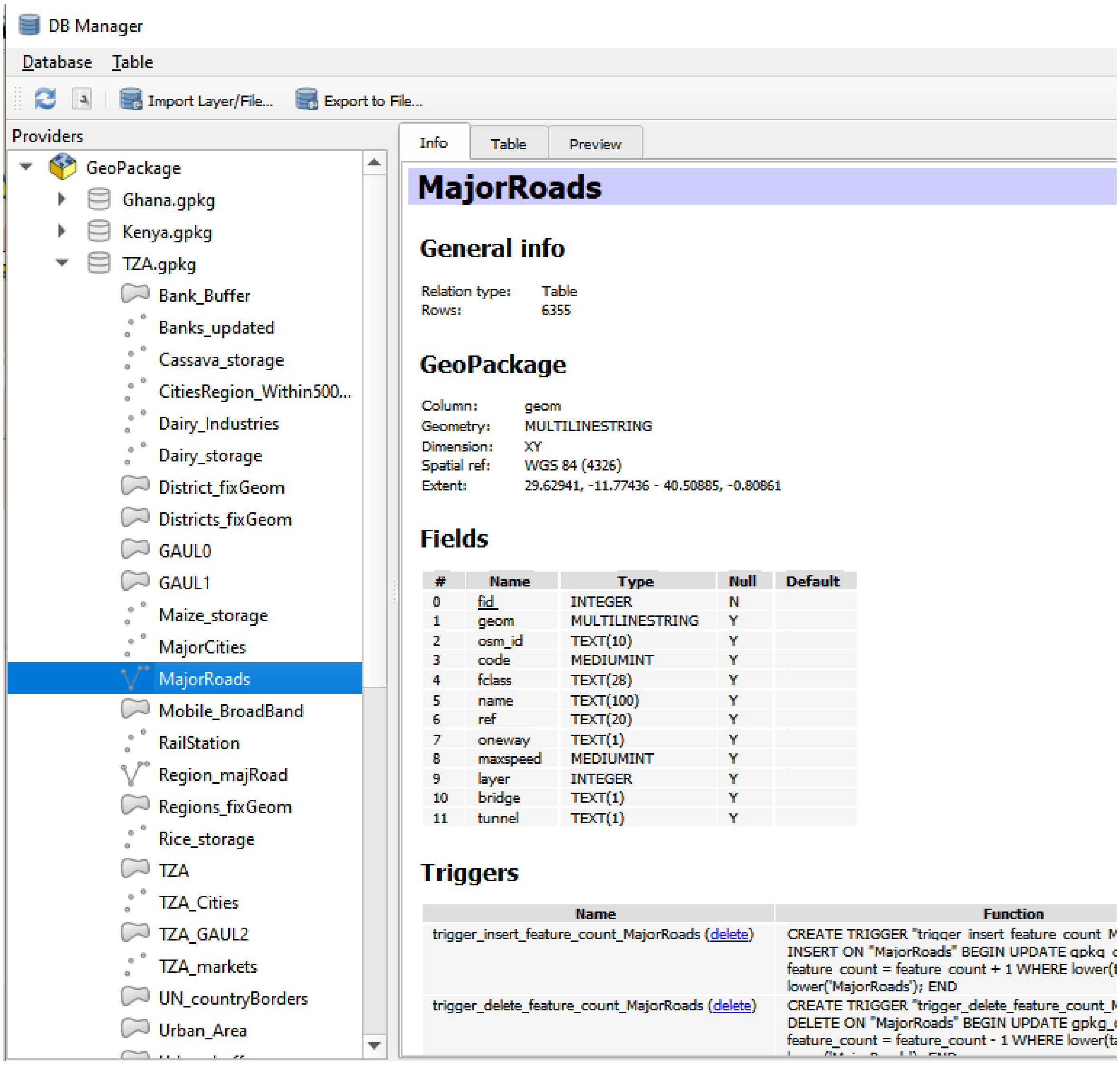


Figure 1 - Geopackage

Neighboring countries data were processed into country *geopackages* to compute major regional cities accessibility for:

- Kenya; Uganda; Rwanda; Burundi; Democratic Republic of the Congo; Zambia; Malawi; Mozambique.

4. Data Editing Geoprocessing

This section details the editing and geoprocessing steps.

4.1 Raster travel time surfaces – Cost/time distance calculation

The development of an algorithm (python script) partially automates the accessibility map data processing (travel time/cost surfaces).

The calculation of time/cost distance surfaces is based on some assumptions:

1. Ports and cities are points features, lakes (inland waters) represented by polygons, other infrastructure network layers consist of linear features. Alternatively, railway stations (points) could also be used as railway connectivity imposes a tunnel effect, time-distance compression between points (stations), with low accessibly in between.
2. River navigation is not computed for Tanzania.
3. Road travel time/cost is modeled for trucks, it is assumed that the accessibility is related to cargo freight vehicles, tertiary and local traffic roads are not included.
4. Lake and river navigation are treated as a surface (polygons) not taking into consideration navigation infrastructure (points), navigation is assumed to use small/medium cargo crafts.
5. The model travel time/cost from/to ports is defined for major Maritime ports only. Lake Ports in Lake Victoria, Lake Tanganyika or Lake Malawi can have some relevance. Major lake port, Mwanza port, accessibility[2] allows cargo to neighboring countries (trade partners).
6. Computation of travel time/cost to major cities weights population dimension. Due to this change accessibility can be considered both, as an infrastructure and a demand criterion.
7. Major islands accessibility is computed, ocean navigation input parameter value is the same as river and lake navigation.

[2] Lake Victoria Ports

The steps to produce accessibility maps (travel time surfaces) are: rasterization vector layers; creation of cost friction surface and; computation of a cumulative time/cost layer from/to points:

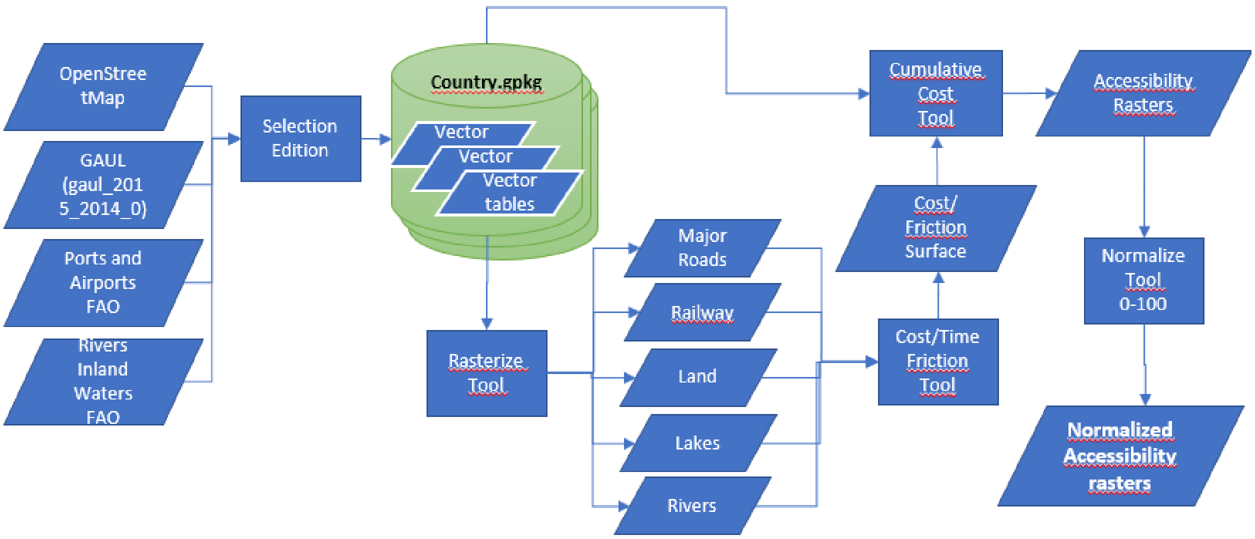


Figure 2 - Accessibility modeling flowchart

1. ***Rasterize* Tool** *(SAGA raster normalization tool)* – Communication/transportation network and surfaces conversion from vector to raster with 1km cell grid burning a value of average time (minutes) to travel 1km for the considered transportation mode ((a) land/walk, (b) major roads/vehicle, (c) railway/train, (d) river/navigation, (e) lakes/navigation, ocean(f).

   The assumed modeling values:

Land (a)	10
Major roads (b)	1
Railway (c)	0.6
River (d)	3
Lakes (e)	3
Ocean (f)	3

The rasterization step outputs individual 1km raster grids with the modelling value attribute for each cell. Modelling value – speed/time - parameter can be easily changed/adapted to a different specification.

2. ***Cost/Friction Tool*** *(GRASS r.series tool)* – A cost or friction surface or grid is obtained overlaying (a), (b), (c), (d) (e), (f), raster grids propagating the minimum cell values; The cost/friction surface output covers Tanzania and all bordering countries.
3. ***Cumulative Cost Tool (****GRASS r.cost tool)*– Service Area – The cumulative cost/accessibility maps are produced selecting a central point, or points, and defining service areas. This processing outputs grids for ports, country's major cities (>200k habitants) and regional major cities (>250k habitants) less than 500 km from the border. The rasters are clipped/masked using country boundaries for correct value classification and improved visualization.

   An accessibility/cumulative cost map was generated for each of the country's 10 largest cities and regional major cities less than 500 km from the Tanzanian borders.

   a. ***City Accessibility/CumulativeCost weighting*** *(GRASS r.series tool)* - Accessibility to cities is considered as both infrastructure and demand criteria, accessibility is weighted using the city population dimension as a measure of demand. This city pull factor layer consists of an arithmetic weighted sum of single city accessibility grids, using the city population dimension[3].

City	Population	Weight %
Dar es Salaam	4,364,541	0.579
Mwanza	706,453	0.094
Zanzibar	501,459	0.067
Arusha	416,442	0.055
Mbeya	385,279	0.051
Morogoro	305,840	0.041
Tanga	221,127	0.029
Kigoma	215,458	0.029
Dodoma	213,636	0.028
Songea	203,309	0.027

Table 1 - Population major cities Tanzania

[3] https://www.citypopulation.de

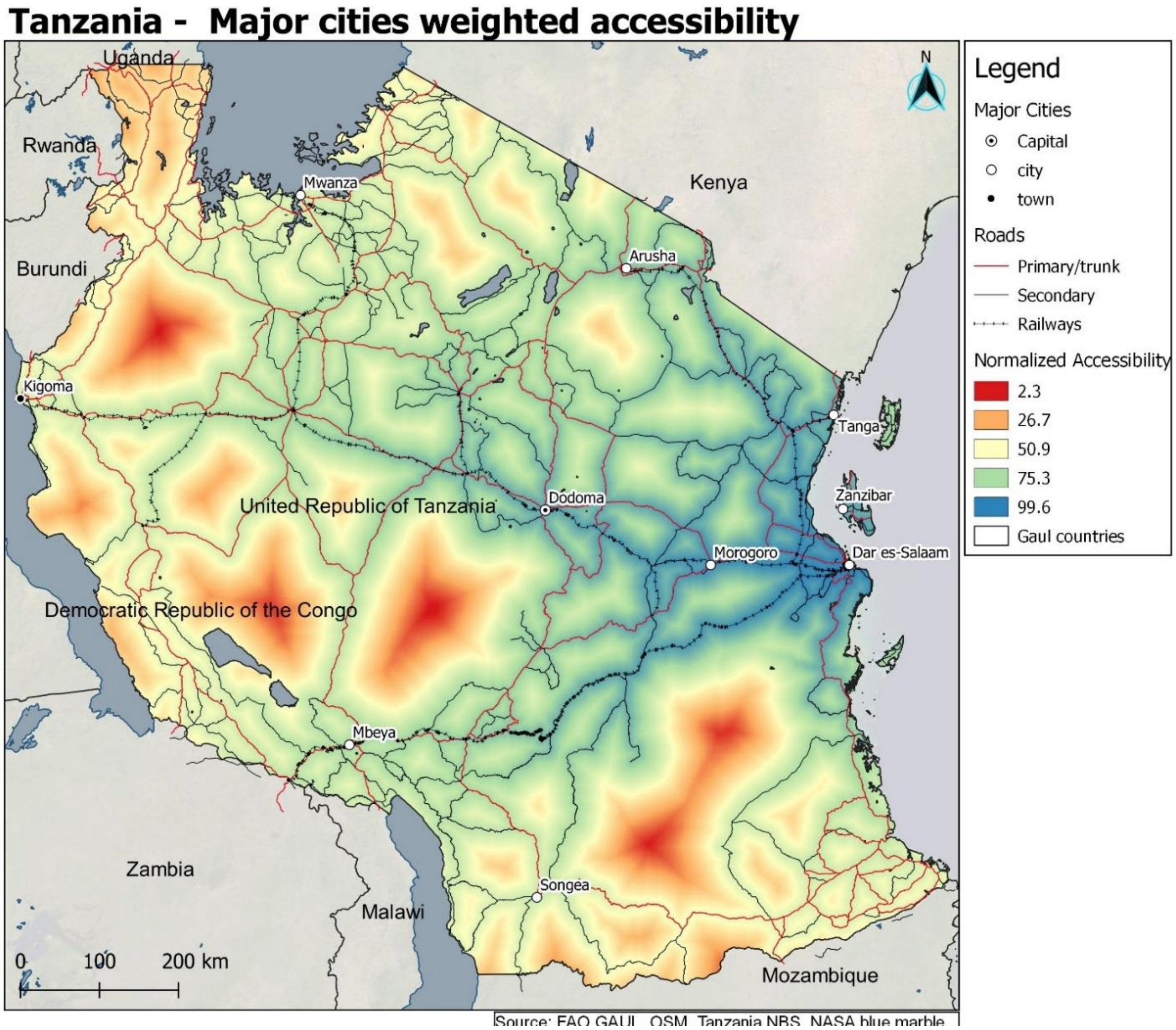


Figure 3 – Normalized weighted accessibility cities x population

Name	Population	Weigth %	Country
Nairobi	4,397,073	0.277	KEN
Kampala	2,094,000	0.132	UGA
Mombasa	1,208,333	0.076	KEN
Lilongwe	1,055,700	0.067	MWI
Kigali	859,332	0.054	RWA
Blantyre	830,100	0.052	MWI
Nampula	663,212	0.042	MOZ
Nakuru	570,674	0.036	KEN
Nansana	532,800	0.034	UGA
Bujumbura	497,166	0.031	BRI
Ruiru	490,120	0.031	KEN

Eldoret	475,716	0.030	KEN
Bukavu	471,789	0.030	COD
Kira	462,900	0.029	UGA
Kisumu	397,957	0.025	KEN
Kikuyu	323,881	0.020	KEN
Kyengera	285,400	0.018	UGA
Thika	251,407	0.016	KEN

Table 2 - Population regional major cities

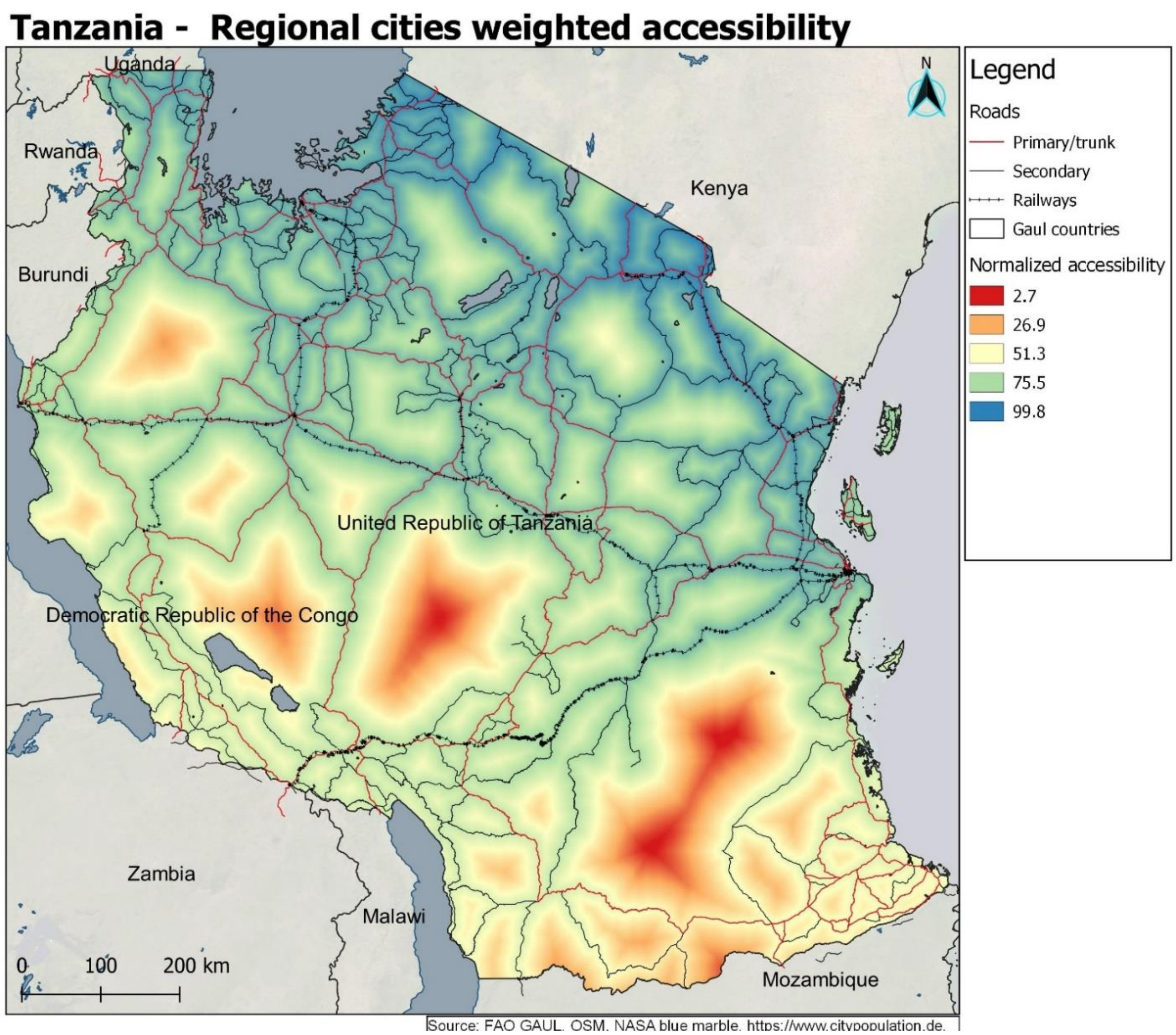


Figure 4 – Normalized weighted accessibility regional cities x population

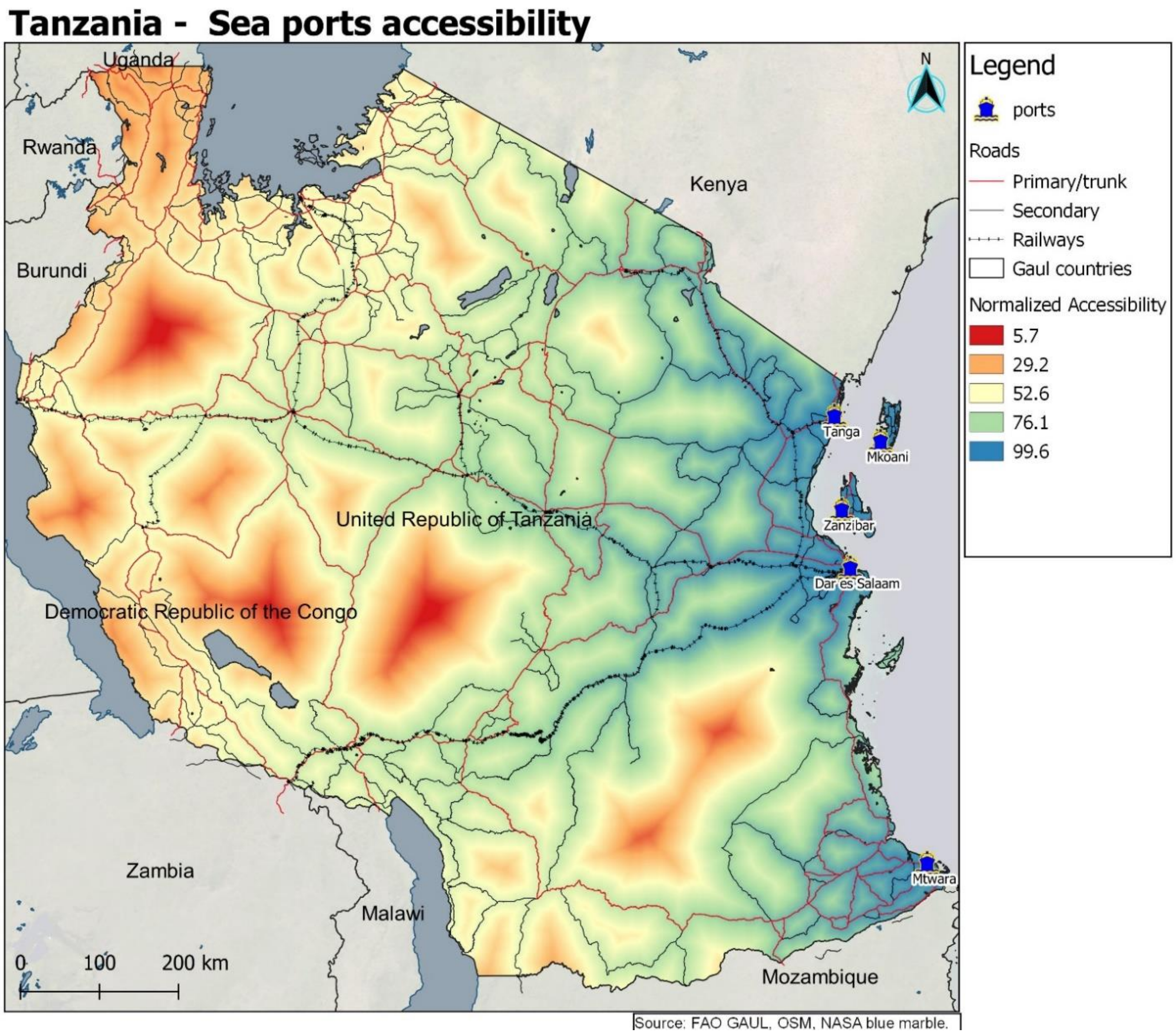


Figure 5 – Normalized accessibility to ports

Regional cities accessibility calculation implies accessibility modeling the entire region, Tanzania and neighboring countries, Kenya, Uganda, Rwanda, Burundi, Democratic Republic of the Congo, Zambia, Malawi, Mozambique. This geoprocessing step can be automated via algorithm/script.

Accessibility to ports is modeled for Maritime ports only. Large lakes have ports that allow cargo transportation namely Tanganyika, Malawi, and Victoria. Lake Victoria is the second largest freshwater lake in the world, Mwanza port[4] accessibility modeling can impact in optimal or

[4] Lake Victoria Ports

recommended storage locations definition. Victoria Lake catchment area includes Kenya, Uganda, Tanzania, Rwanda and Burundi, around 35 million habitants. The Lake zone is an important production zone for rice and cassava and exports find in the region (Rwanda, Uganda, Burundi and Kenya) relevant trade partners: 95% of Cassava exports, totaling 910.1 mt, goes to Burundi while Rice exports to the region represent 39.3% of the total, with 6,042.5 mt[5].

4. ***Normalization*** – To calculate a weighted sum location score the different criteria units are normalized scaling from 0 to 100. The lowest accessibility value (time or cost) corresponds to 100 (high accessibility).

4.2 Human Population Density

World human population density estimates for 2020 1km grid raster.

1. ***Editing*** – Clipped/masked using GAUL country borders to extract country grid.
2. ***Normalization (raster normalization)*** – Raster grids are normalized into common scale (0 to 100) to calculate a multicriteria location score (weighted sum).

In the case of population density, a high concentration of population in major cities, specially Dar es Salam, implies a frequency distribution extremely left skewed, a natural break/natural classes method of classification can allow an informed decision (expert value judgement) on thresholds of urban rural definition.

[5] World Integrated Trade Solution (WITS) | Data on Export, Import, Tariff, NTM (worldbank.org)

4.3 Production - Crops

It is assumed that the mobile warehouse location project focus is on food crops, where supply and demand disruptions impact on food systems with potential severe consequences. Cash and export crops being secondary might lead to a potential pitfall due to a possible conflict between market commodities and/or subsistence crops. Subsistence crops can be produced in large quantities without significative logistical storage/warehousing needs or supply/demand disruption potential.

Crops were selected using FAOStat data from 2014 to 2018 using production (tons) mean value.

Table 3 represents mean production value in metric tons for the last available 5 years statistics 2014-2018, FAO crop production items/code were aggregated to match MapSPAM coding/clustering. The top 13 crop and crop groups represented around 90% of the total production.

FAO code	Crop	SPAM	Mt
56	Maize	maiz	6081620.2
125	Cassava	cass	5315461.4
27	Rice	rice	4835635.2
122	Sweet potatoes	swpo	3760952.4
486	Bananas	bana	3432106.0
156	Sugar cane	sugc	2946068.8
a)	Other oil crops	ooil	2819773.4
b)	vegetables	vege	2623329.2
116	Potatoes	pota	1624441.8
c)	tropical fruits	trof	1283391.0
242	Groundnuts, with shell	grou	1260379.0
176	Beans, dry	bean	1168423.6

Table 3 - Mean crop production values 2014/2018

a) Other oil crops - 265 Castor oil seed, 267 Sunflower seed,280 Safflower seed,289 Sesame seed,358 Seed cotton, 339 Oilseeds nes.

b) vegetables - 358 Cabbages and other brassicas; 388 Tomatoes; 401 Chillies and peppers, green; 403 Onions, dry; 406 Garlic; 414 Beans, green; 417 Peas, green; 420 Vegetables, leguminous nes; 446 Maize, green; 463 Vegetables, fresh nes.

c) tropical fruits - 490 Oranges; 497 Lemons and limes; 567 Watermelons; 571 Mangoes, mangosteens, guavas; 574 Pineapples; 603 Fruit, tropical fresh nes.

The analysis, location score and final location maps, was applied for the top 3 crops individually, Maize, Cassava and Rice.

4.3.1 Production data processing

1. ***Editing*** – Clipped/masked SPAM datasets using GAUL country borders to extract a raster for each of the selected crops.

3. ***Normalization (raster normalization)*** – Raster grids are normalized into a 0 to 100 common scale to calculate location score (weighted sum).

4.4 Location Score / Multicriteria weighted sum

The location score is obtained by way of a simple arithmetic weighted sum (*GRASS r.series* tool) of the normalized/scaled grids, location score can theoretically vary from 0 to 100.

The assumed weighting for each of the criteria is as follows.

*("Crop Production" * 0.4) + ("Human Population Density" * 0.2) + ("Port Accessibility" * 0.2) + ("Major Cities Accessibility" * 0.1) + ("Regional Cities Accessibility" * 0.1)*

The final output is an approximately 1km (0.01 degree) raster grid with the location score cell value.

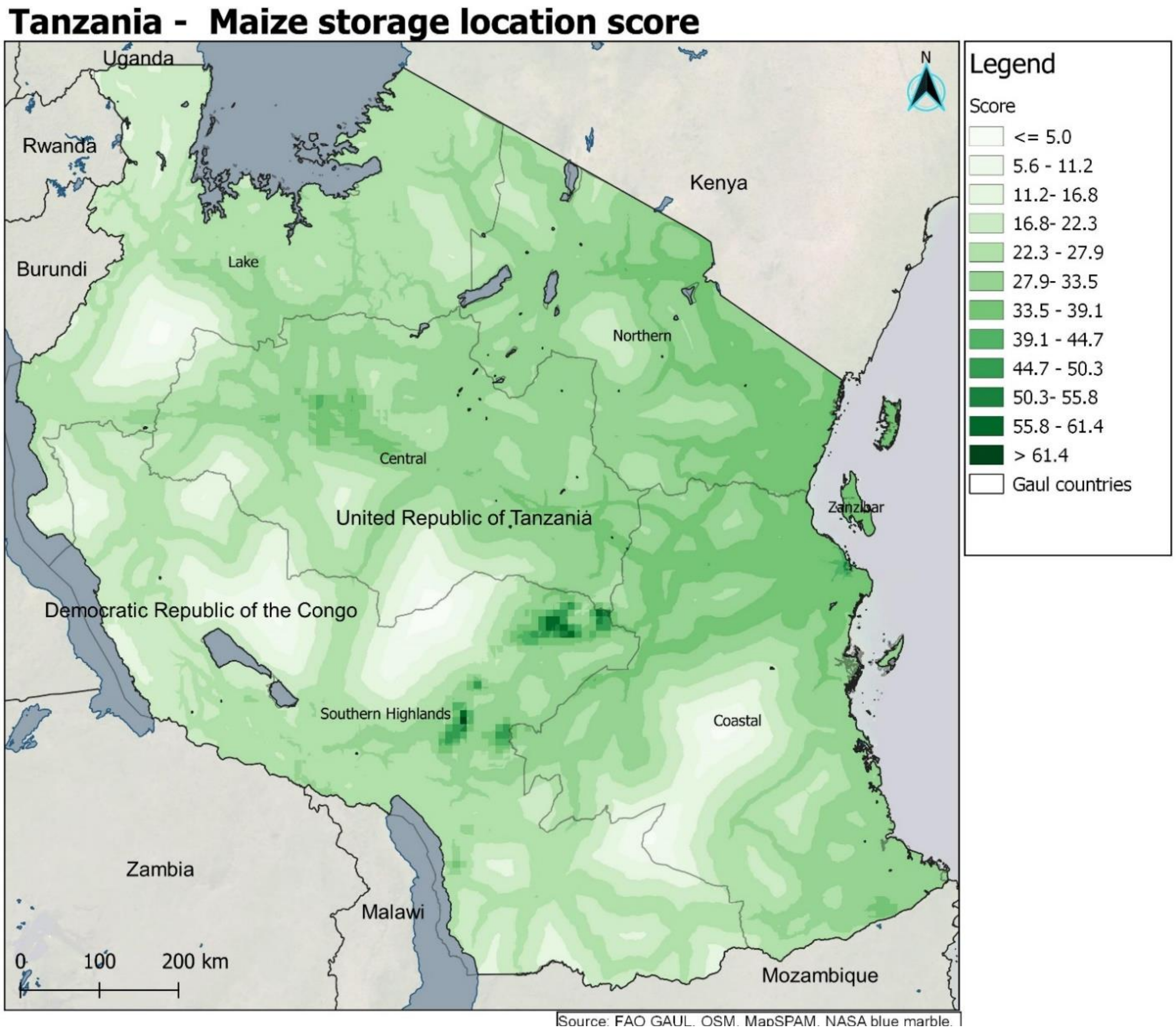


Figure 6 - Maize location score

Maize production shows a relatively disperse pattern. The highest location scores can be found in southern highlands and central regions, with some weigh in the coastal and northern production zones urban areas.

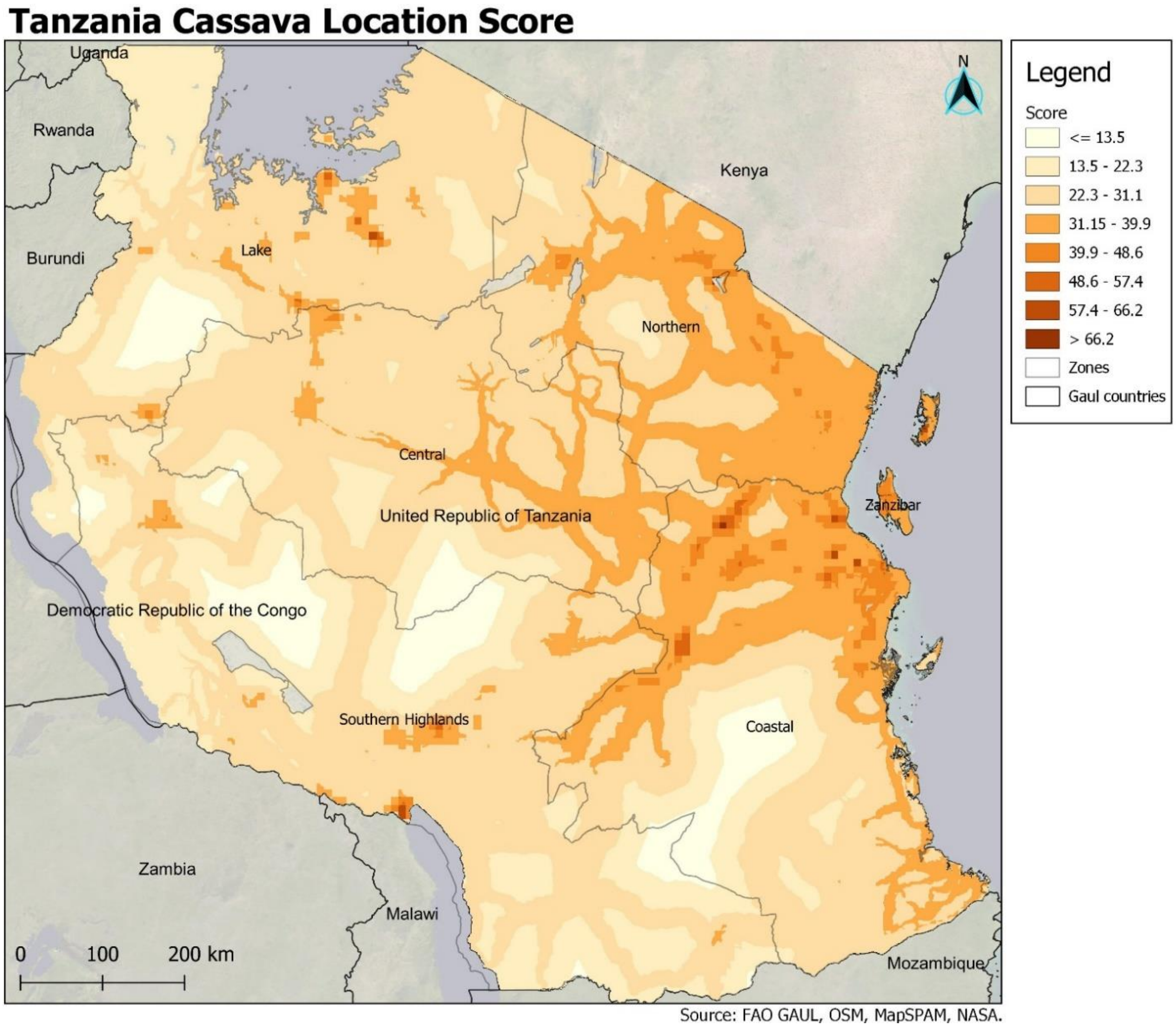


Figure 7 - Cassava location score

Cassava production is also spatially distributed and can be found in most of the country with high location score areas in all production zones. The northern half of the coastal zone is clearly emphasized displaying large areas in the mid and top score ranges.

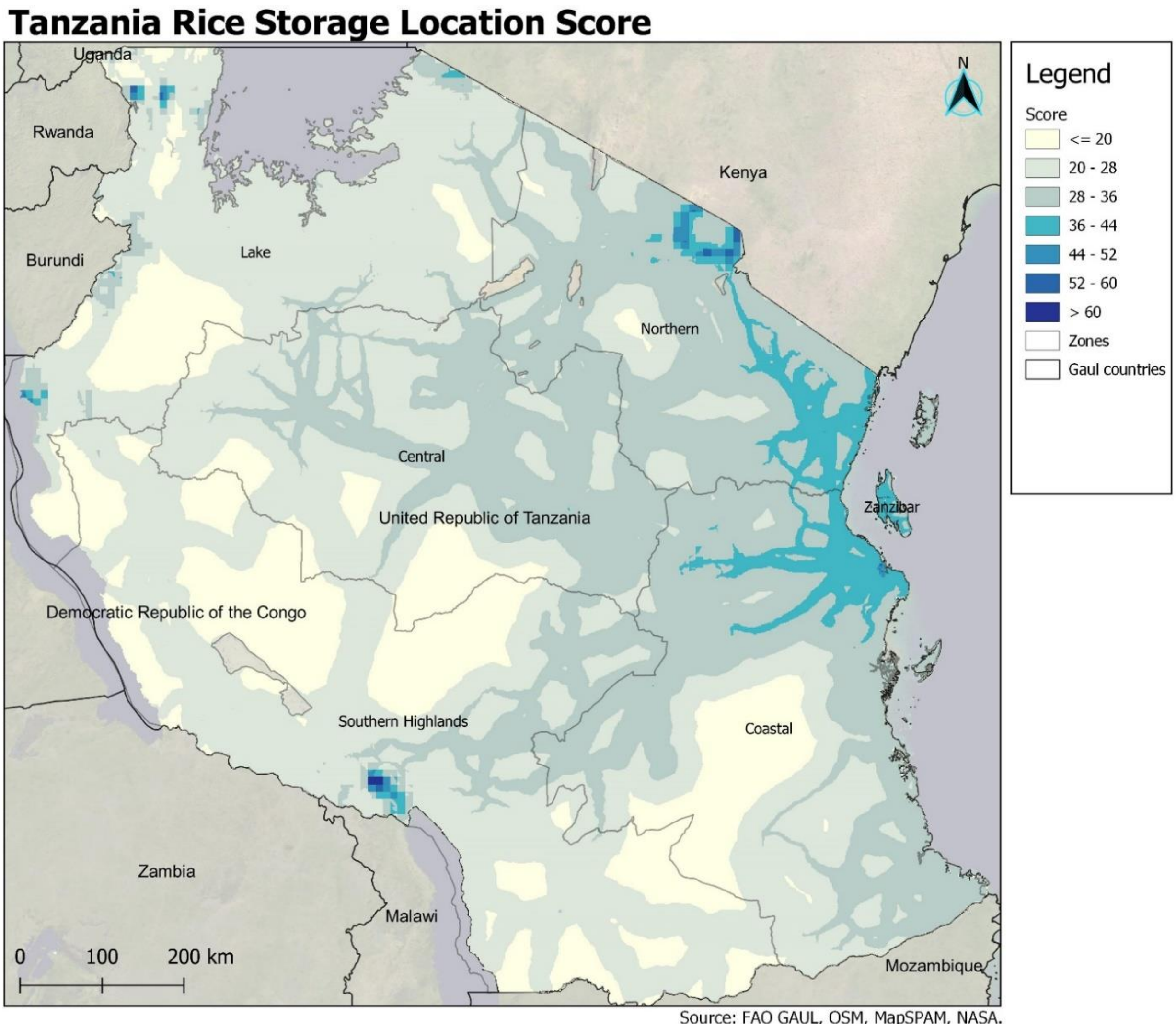


Figure 8 - Rice Location Score

Rice production spatial distribution is associated to ample water accessibility. Top score areas concentrate in Lake, Northern and Southern Highlands production zones. The northern zone also has higher accessibility to major Tanzanian and cross border population centers.

For all analyzed commodities, coincident low score areas identify low accessibility, arid or hyper-arid climate areas with probably low agricultural production.

4.5 Final Location Top Score Map

Final location top score mapping methodology was further developed from mobile storage location MCDA for Kenya and Ghana. Top score grids defined using the top 99th percentile and the final location maps classified further into 3 equal classes intervals.

With minor and detailed exceptions, final location maps use the top class and add as exclusive criteria: access to finance; distance from major roads and; access to IT (mobile broadband connection).

4.5.1 Access to finance and distance to major roads

Access to finance and roads is defined as linear distance threshold using the following geoprocessing steps:

- *Buffering* – the tool computes a buffer area for all the features in an input layer using fixed or dynamic distance:
  - Banks - 20km (0.18 degree) buffer radius.
  - Major roads - 2km (0.018 degree) buffer radius.
- *Intersection* - extracts the overlapping portions of features in the Input and Overlay layers: Banks_Buffer and Roads_Buffer.
- *Dissolve* - Takes the intersection vector layer and combines the features into a new feature, a single polygon.
- *Clip Raster by Mask Layer* – The grids are extracted using the polygon (dissolved intersection of the banks and roads buffers).

4.5.2 Access to IT

Collins Bartholomew's GSMA Mobile Coverage Explorer database is the source for a derived raster, 1km resolution mobile broadband coverage - 1/0 (coverage/no coverage) - dataset. The grid results from the selection and aggregation of 3G (strong signal), 4G and 5G grids, from both operator submissions and OpenCell ID[6] cell tower locations datasets, by applying the following processing steps:

1) Data extraction strong signal coverage 3G strong signal (good indoor coverage) and 4G - MCE_GH3G_2020.tif;
2) Sum (OCI 3G + MCE 3G) and 4G (*r.series* tool cell statistics).
3) Convert raster to Boolean – coverage/no_coverage - (raster calculator).
4) *Clip Raster by mask layer* – clips data by countries borders.
5) Convert raster to vector (*polygonize* tool) - deleted 0 value polygons (no coverage).
6) Check polygons topology (*fix geometries* tool).
7) Create coverage polygon (*dissolve* tool).
8) *Clip Raster by Mask Layer* – top score grids are extracted using the coverage polygon as exclusive criteria.

4.5.3 Final Location Mapping

After cliping selected top score locations using exclusive criteria, a visual map inspection was performed to detect: logistics and industrial districts, port and railway station proximity, availability of space when in urban area and avoiding residential and turistic areas or districts.

Final location maps put in evidence the livestock production systems and the weight of the dairy herd density, recommended sites are pinpointed based on a detailed large scale analysis of OpenstreetMap, Google and Bing maps and satellite imagery.

Although it involves value judgement, this exercise step tests a return from the model abstraction to field reality, can be considered a first step into project operationalization and also allows reducing the number of proposed sites by aggredating close locations or areas.Large scale maps can be produced detailing each recommended location.

Two patterns appear evident, high productive areas or highly populated and high accessibility regions, with suggested locations in and around largest cities and specially Dar es Salam.

[6] OpenCellID is a collaborative community project that collects GPS position of user equipment to locate cell towers. As of October 2017, the database contained almost 36 million unique GSM Cell IDs. With more than 75.000 contributors and millions of daily measurements.

Large scale maps can be produced detailing each recommended location.

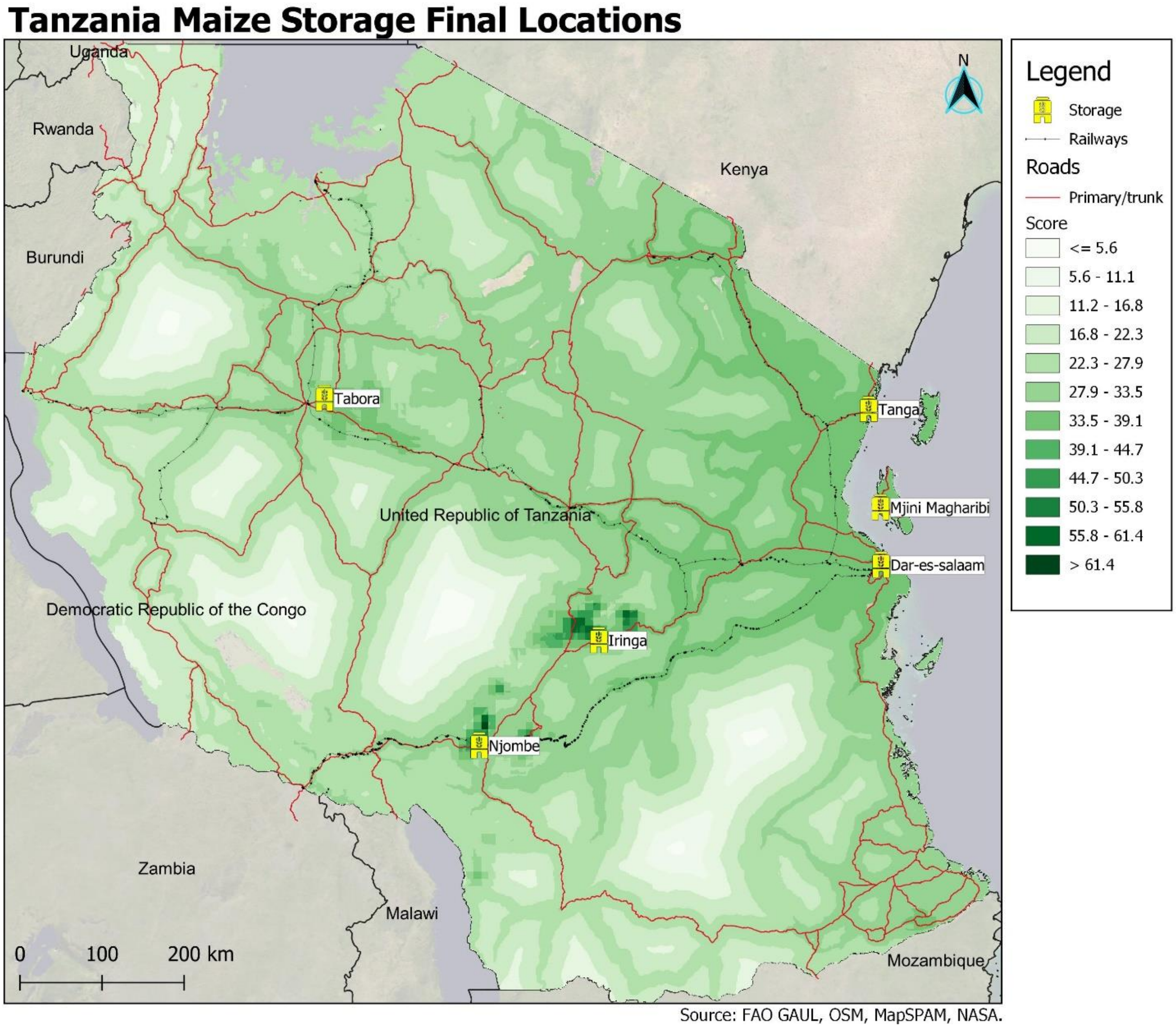


Figure 9 - Maize recommended storage sites

Region	District	Zones	Location
Tabora	Uyui DC	Central	T8 Rd, Mayombo
Dar-es-salaam	Ilala MC	Coastal	Nelson Mandela Rd, Tabata Mwananchi, logistics district
Tanga	Tanga CC	Northern	T13, Monbassa Rd, Tanga
Njombe	Wanging'ombe DC	Southern Highlands	T1 Rd, Wanging'ombe-Makambako
Iringa	Kilolo DC	Southern Highlands	T1, Iringa Rd, Ilula/Masukanzi

Mjini Magharibi	Mjini	Zanzibar	Zanzibar City

Table 4 - Maize storage sites

Tabora Uyui DC Storage proposed location serves Shinyanga, Geita and Tabora Regions with very extensive maize production, that have several top score areas with low infrastructure and financial services.

Iringa and Djombe Regions have also vast productive areas with low financial services and access to IT (cellular broadband coverage).

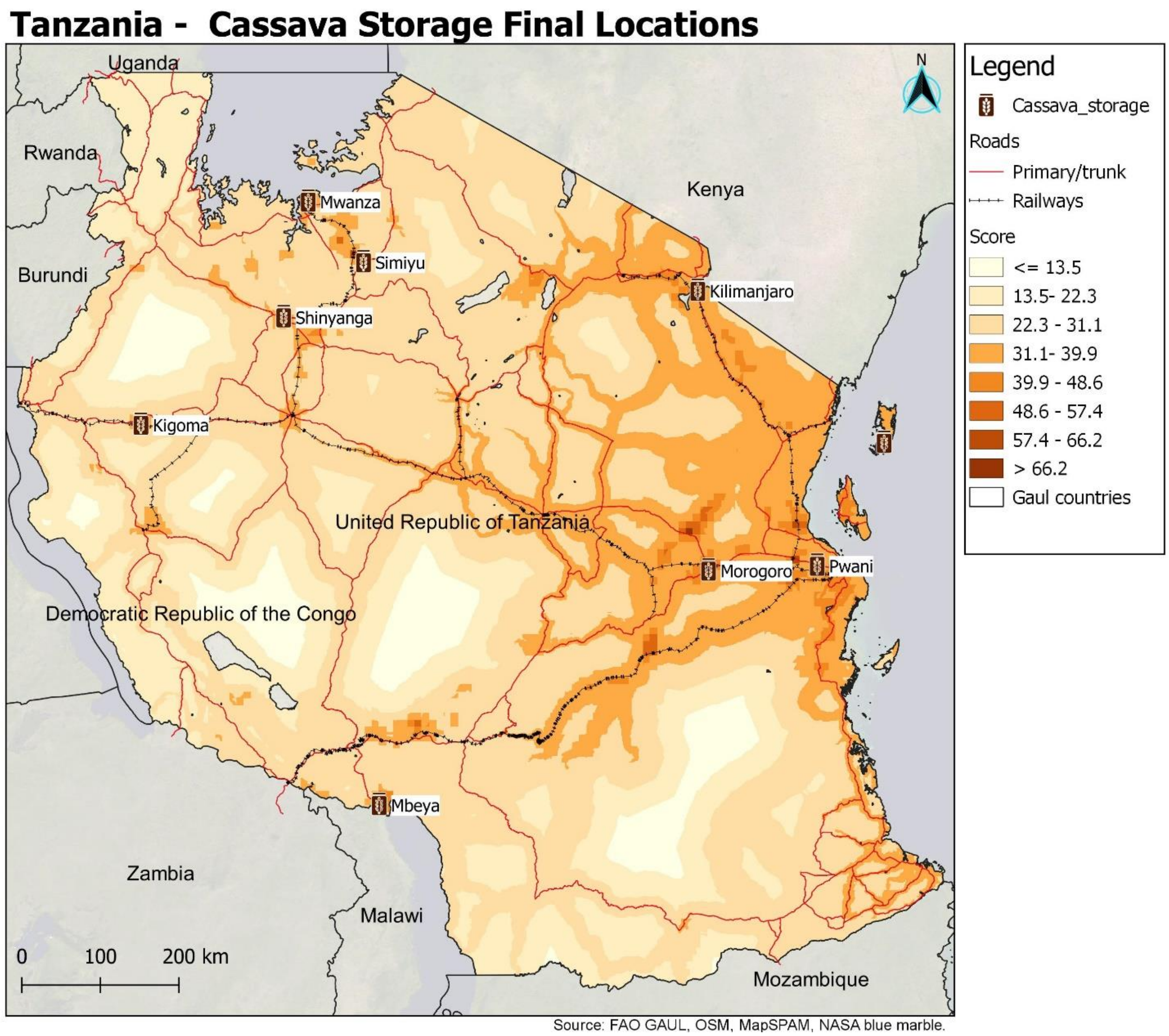


Figure 10 - Cassava recommended storage sites

Region	District	Zone	Location
Pwani	Kibaha DC	Coastal	T1, Morogoro Main Rd, Picha Ya Ndege
Morogoro	Morogoro MC	Coastal	Morogoro
Shinyanga	Kahama TC	Lake	T3, Rwanda Rd between Kahama and Kagongwa
Kigoma	Uvinza DC	Lake	T18-cental line (railway) Nguruka
Mwanza	Ilemela MC	Lake	Kahama
Simiyu	Maswa DC	Lake	T36 near Isulilo Village

Kilimanjaro	Mwanga DC	Northern	Kifaru, T2
Mbeya	Kyela DC	Southern Highlands	T28, Kyela
Kusini Pemba	Chake Chake	Zanzibar	Mkoani Chake-Chake Rd

Table 5 - Cassava storage sites

Cassava production has a more distributed pattern if compared to maize and rice. Due to poor financial services coverage and/or no broadband connectivity, some intensive production areas, like northern Morogoro Region (Mvomero), have no recommended sites.

The Lake zone includes most recommended locations and it must be mentioned that Burundi is a major destination for cassava exports, around 90.5% of the total.

Pwani Region (Kibaha DC) location is sited in Dar es Salam periphery (T1 Morogoro main road) in the area there are several top score options, in and around Dar es Salam. In this case the sites inside the urban area (>1500hab/km2) were excluded.

Kilimanjaro Mwanga DC Storage case can exemplify the final location method:

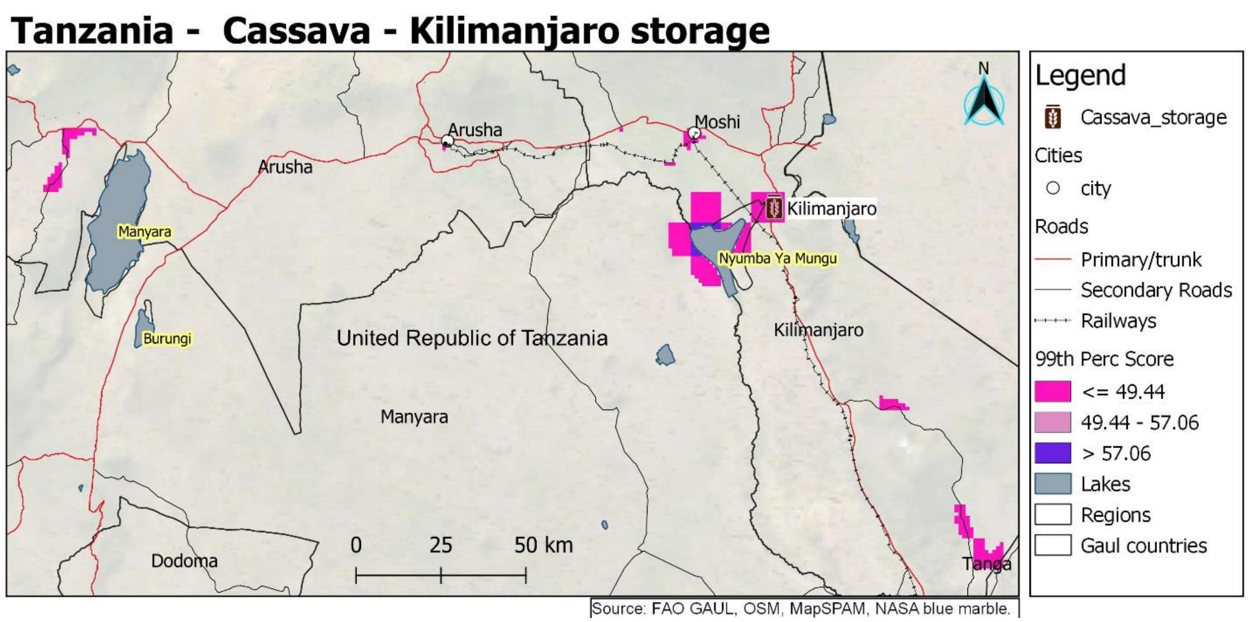


Figure 11 - Cassava Kilimanjaro storage site 99th percentile

Although the recommended storage location does not rank above the 57.06 score (in the 99th percentile top class) it aggregates several potential areas in the Northern Region and road south: Moshi and Kifaru, close to Nyumba ya Mungu reservoir (Moshi DC and Simanjiro DC), some above

57.06 score value. To the west, also including highly productive Karatu DC that stretches between Manyara and Eyasi Lakes and still Arusha city area, both in Arusha region.

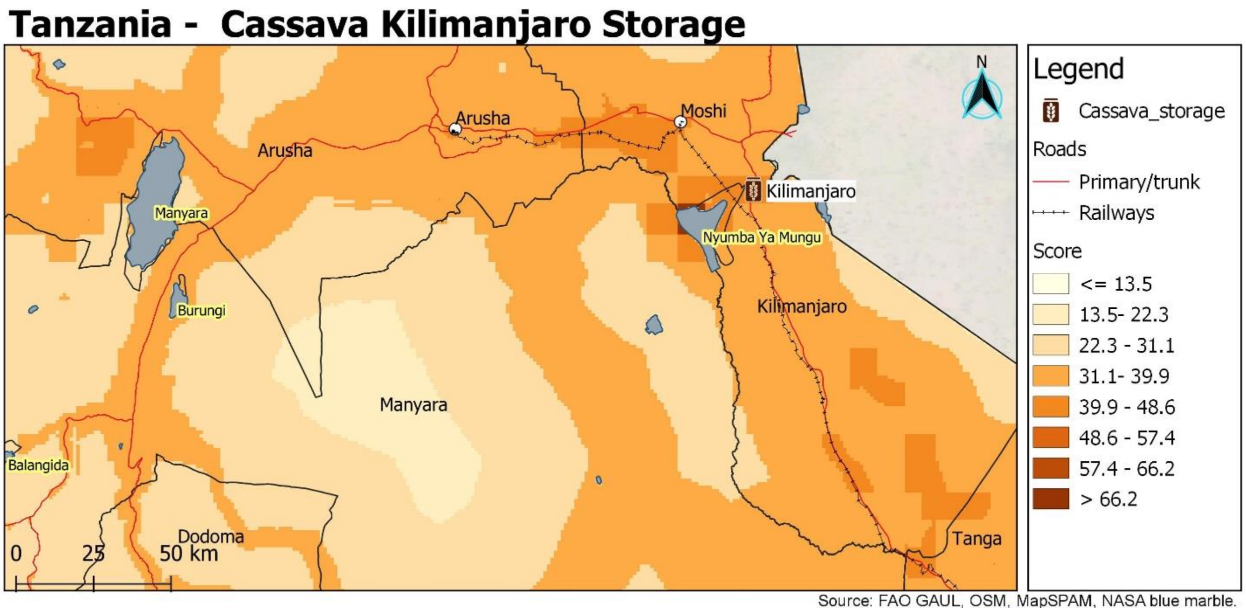


Figure 12 - Cassava Kilimanjaro location score

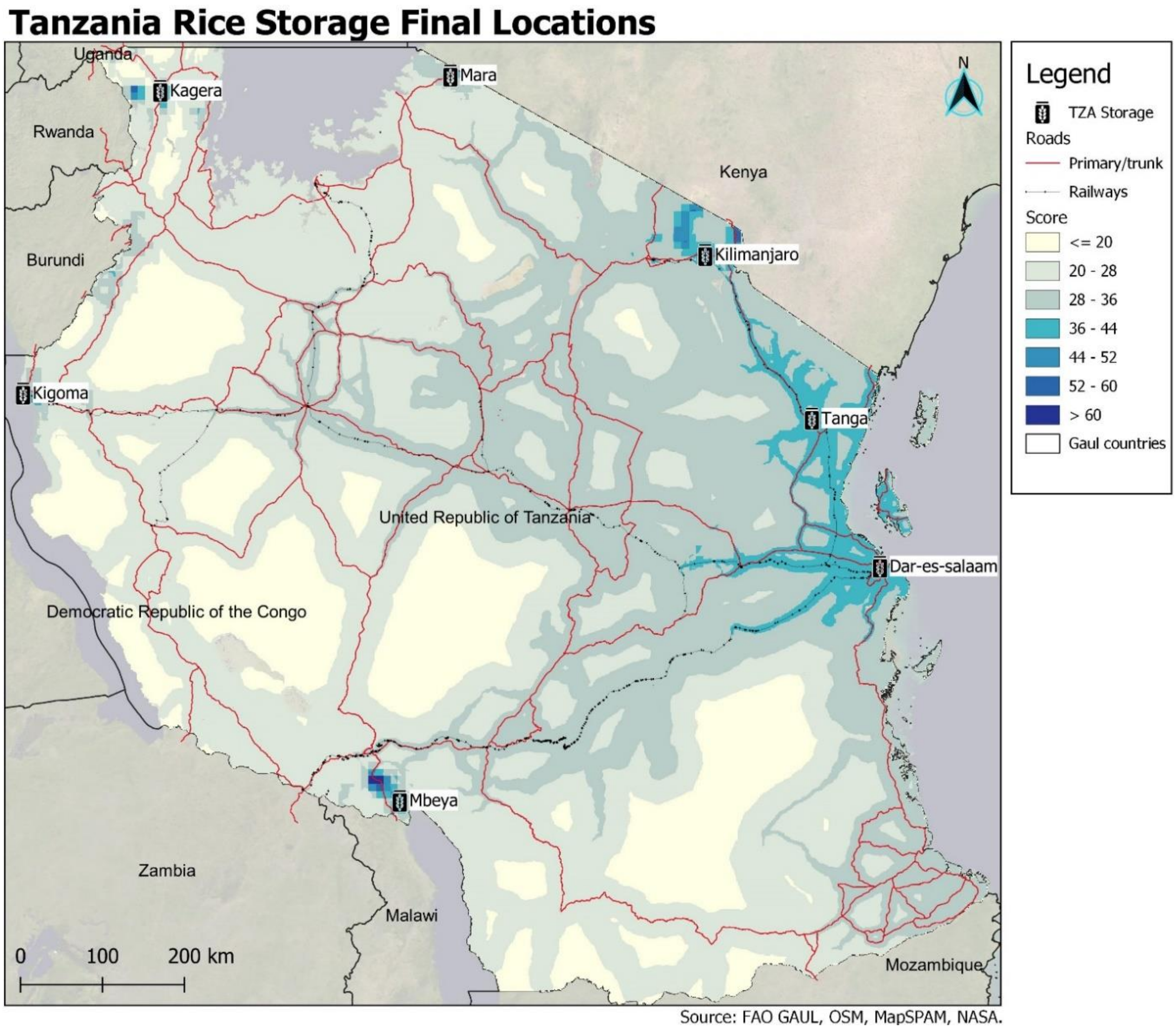


Figure 13 - Rice recommended storage sites

Final rice storage location map denotes a highly localized production in Lake, Northern and Southern Highlands zones.

Rwanda is a major export partner with around 64.5% of total rice exports.

Region	Districts	Zones	Location
Dar-es-salaam	Ubungo MC	Coastal	Nelson Mandela Rd, Tabata industrial/Logistics
Kigoma	Kigoma Municipal-Ujiji	Lake	Kigoma
Kagera	Karagwe DC	Lake	T38 Rd, Kayanga-Kishoju

Mara	Tarime TC	Lake	T4 rd, Magena
Kilimanjaro	Hai DC	Northern	T2 Rd, Kwasadala-Moshi
Tanga	Korogwe TC	Northern	T2 Rd, Korogwe
Mbeya	Busokelo DC	Southern Highlands	R602, Kabula Mpuguti - Ipinda

Table 6 - Rice storage sites

Conclusions

The modeling exercise used a GIS-MCDA methodology using weighted factors to identify and recommend mobile storage (movable warehouses) location sites in Tanzania, for major crop commodities, maize, cassava, and rice.

Some minor changes were introduced to previous mobile storage location analysis for Kenya and Ghana:

1. Accessibility to cities is weighted using the city population size.
2. Commodities are not clustered.

The modeling variables characterize supply, demand and infrastructure/accessibility, main logistical factors for warehousing facilities location.

Quality and reliability of base input data support the extent to which conclusions can be trusted, and these are just as sound as the analysis conducted. Also, as with any modeling exercise, specification and objectives define the assumptions and approximations, an informed review of initial specification can always produce distinct answers.

Location score maps shows a relatively disperse pattern in maize production. The highest location scores can be found in southern highlands and central regions, with some weigh in the coastal and northern production zones urban areas.

Cassava production is also spatially distributed and can be found in most of the country with high location score areas in all production zones. The northern half of the coastal zone is clearly emphasized displaying large areas in the mid and top score ranges.

Rice production spatial distribution is associated to ample water accessibility. Top score areas concentrate in Lake, Northern and Southern Highlands production zones. The northern zone also has higher accessibility to major Tanzanian and cross border population centers.

For all commodities, two clearly identified location patterns emerge associated to “opposing assessment criteria”, market (urban areas and accessibility) and production zones. Coincident very low score areas recognize low accessibility, arid or hyper-arid climate areas with probably very low agricultural production.

Closing Remarks

This closing remarks discuss limitations of the findings, changes in methodology and additional location criteria, closing with a few points on future directions.

As with any modelling exercise, assumptions in criteria selection and weighting determine the output. GIS data capture, analysis, processing, and representation involve a high degree of generalization and derived, calculated, and classified values, that translates into consequent loss of information and detail. There are obvious pitfalls in modeling abstraction reductionism that must be taken into consideration when in a decision-making or operationalization phase.

Some of the assumptions in travel time/cost (accessibility) modeling must be emphasized:

A. The defined speed/cost input parameters for each transportation mode.
B. Road speed is computed for large trucks and as single value for considered road classes: 'motorway', 'truck', 'primary' and 'secondary'[7].
C. River navigation is not modeled for Tanzania.
D. Accessibility to ports is modeled for Maritime ports only. Lake ports can have relevance considering that there are identified export markets in neighboring countries that can be reached through Lake Victoria, Lake Tanganyika, or Lake Malawi. Lake Victoria is the second largest freshwater lake in the world, Mwanza port[8] is the major infrastructure and its accessibility modeling can impact storage recommended sites. Victoria Lake catchment area includes Kenya, Uganda, Tanzania, Rwanda, and Burundi, around 35 million habitants. In the Lake production zone rice and cassava are grown and exports find in the region (Rwanda, Uganda, Burundi and Kenya) relevant trade partners: 95% of Cassava exports, totaling 910.1 mt, goes to Burundi while Rice exports to the region represent 39.3% of the total, with 6,042.5 mt[9].

---

[7] See (Ramm, 2019) for details on road classification.
[8] Lake Victoria Ports
[9] World Integrated Trade Solution (WITS) | Data on Export, Import, Tariff, NTM (worldbank.org)

Some **assumptions and pitfalls on demand factors** must also be marked.

A. Regional cities accessibility – Treats markets and regional trade as uniform, does not consider the impact of possible conflicts.
B. Population density – A high concentration of population in metropolitan areas implies a frequency distribution extremely left skewed. A classification method rather than normalization can allow an informed decision on the definition on urban/rural settlement typology thresholds, better weighting the factor.

On the **supply side, crop/production,** some of the **assumptions** can also produce downsides:

- Focus on food crops - Subsistence crops can be produced in large quantities without significative logistical requirements.

It can be analyzed which commodities are traded with bordering countries to conduct a location analysis for those.

Data normalization and final exclusive criteria can also have some downsides.

Distinct criteria aggregation requires data normalization/scaling with consequent loss of data.

For the supply criteria - production - most crops show some degree of spatial concentration. In one hand, top production areas are very localized, possibly resulting from intensive farming agrobusinesses which typically will have good logistics in place. On the other hand, small to mid-size family farms in productive land, cover much larger areas and will be represented on the middle range. Although this farmer group could benefit most from mobile storage facilities, they're location is not valued. Natural classes and natural break manual classification based on an informed decision, while surely introducing bias, can have into consideration an envisioned outcome.

The final location maps geoprocessing involves a high degree of generalization. It is based on the definition of a threshold linear distance to/from banks and roads, with several physical and human geographical singularities potentially impacting on results accuracy and reliability. There are also assumptions involved in the estimation of the mobile broadband coverage maps grid. It must be noted that, some highly productive areas are not recommended locations because lack good cellular coverage, are distant to banks or major roads, but in some cases, by a very short distance.

Part of the geoprocessing was semi-automated through the development of algorithms (python scrips), namely:

- Preprocessing, extraction and editing of vector layers (*AuxData* script);
- Computation of accessibility maps cumulative travel time/cost from ports (*TravelCostTimeSurface script*);
- Production aggregation geoprocessing (*ProdData script*).

Scripts can also be developed for other geoprocessing steps:

- Regional cities accessibility - data gathering, data selection, data rasterization, creation of country time/cost layers and combining country's cost layers into a regional grid;
- Final top score map - Sequence of geoprocessing functions for IT access, road buffer distance and mobile broadband connectivity.

Full automation of the geoprocessing is not realistic as distinct geographies impose demand, supply, and infrastructure specific characteristics: population distribution and settlement typology, agroecological zones and farming systems, crop commodities and seasonality, infrastructural quality and development, navigation of inland waters and rivers.

Some of the possible project developments include the methodological improvement of current features or the introduction of new criteria:

1. Infrastructure and accessibility dimension:
    a. Railways are modeled as linear phenomena; railway connectivity nevertheless imposes a tunnel effect with time-distance compression between points (stations) but low accessibly in between; Motorways can also be modeled using access points.
    b. Road network modeling uses a single value for heavy truck transportation irrespective of road class (motorway, truck, primary, secondary) or road speed limits.

2. Socioeconomic data – Other subnational scale level data can be included.
3. Production dimension:
    a. Time dimension - Seasonal criteria can be included.
    b. Plant phenology.

The model has flexibility to accommodate new developments or improvements or can be closed with the present configuration to be automated and to process a larger number of countries and regions. Different scenarios and outputs can be explored using the location criteria weighting (weighted sum) as a *what if scenario* tool.

Bibliography

AscoughII, J. C., Rector, H. D., Hoag, D. L., McMaster, G. S., Vandenberg, B. C., Shaffer, M. J., Weltz, M. A., & Ahjua, L. R. (2019). Multicriteria Spatial Decision Support Systems: Overview, Applications, and Future Research Directions. *Extended Education – an International Perspective*, 175–180. https://doi.org/10.2307/j.ctvdf0hzj.10

Boroushaki, S., & Malczewski, J. (2010). Measuring consensus for collaborative decision-making: A GIS-based approach. *Computers, Environment and Urban Systems*, *34*(4), 322–332. https://doi.org/10.1016/j.compenvurbsys.2010.02.006

Kitchin, R. (2014a). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 205395171452848. https://doi.org/10.1177/2053951714528481

Kitchin, R. (2014b). *The Data Revolution: BIG data, open data, data infrastructures & their consequences*. SAGE.

Malczewski, J. (2006). GIS-based multicriteria decision analysis: A survey of the literature. *International Journal of Geographical Information Science*, *20*(7), 703–726. https://doi.org/10.1080/13658810600661508

Mulrooney, T., Beratan, K., McGinn, C., & Branch, B. (2017). A comparison of raster-based travel time surfaces against vector-based network calculations as applied in the study of rural food deserts. *Applied Geography*, *78*, 12–21. https://doi.org/10.1016/j.apgeog.2016.10.006

Ramm, F. (2019). *OpenStreetMap Data in Layered GIS Format V7* (p. 21). Geofabrick.

Wood-Sichra, U., Joglekar, A. B., & You, L. (2016). Spatial Production Allocation Model (SPAM) 2005: Technical Documentation. In *Harvest Choice*. http://mapspam.info

Annex. 1 – Pilot Case Ghana/Kenya Flow Diagram

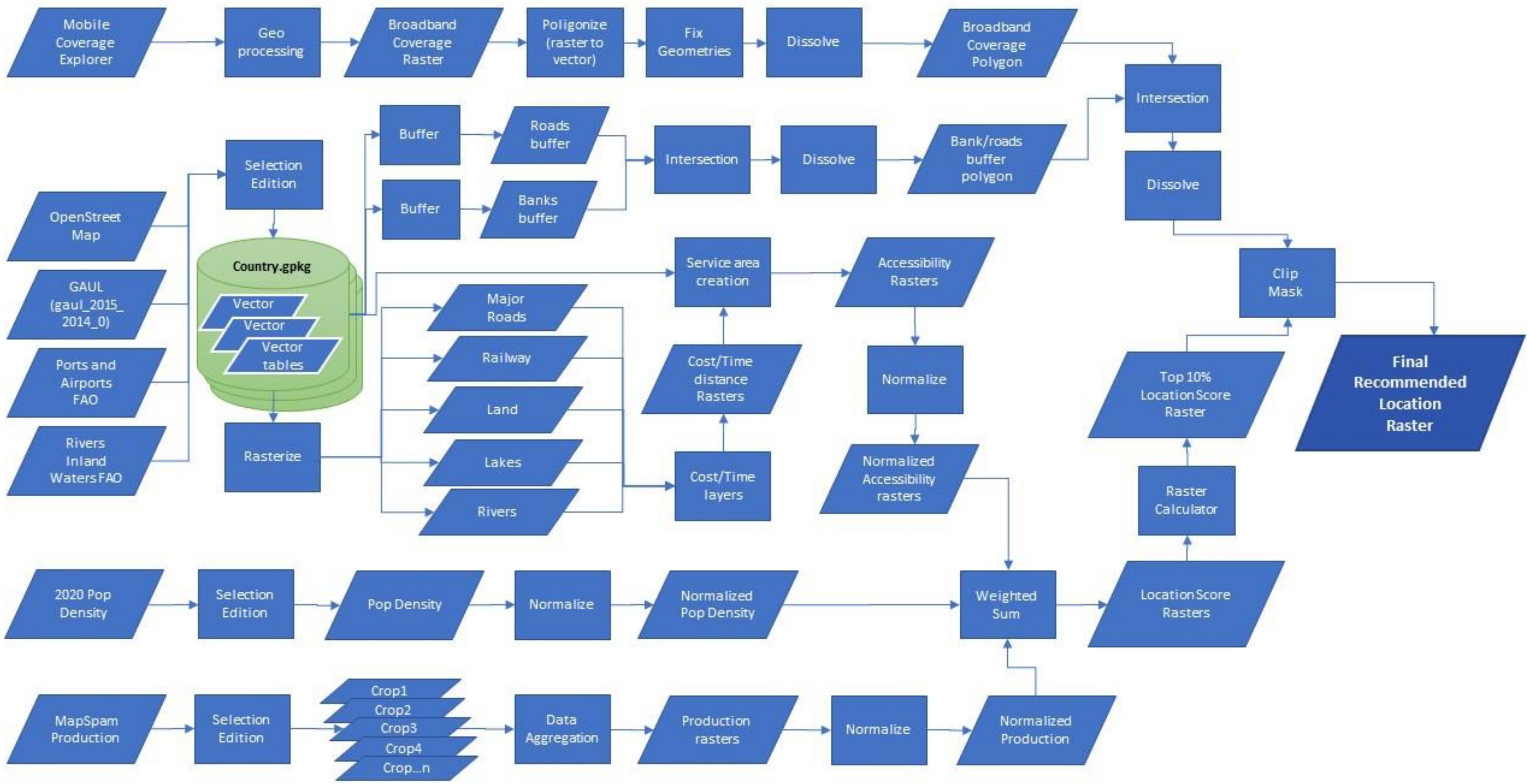


Figure 14 - Project Flow Diagram

Annex 2 – Algorithm diagrams

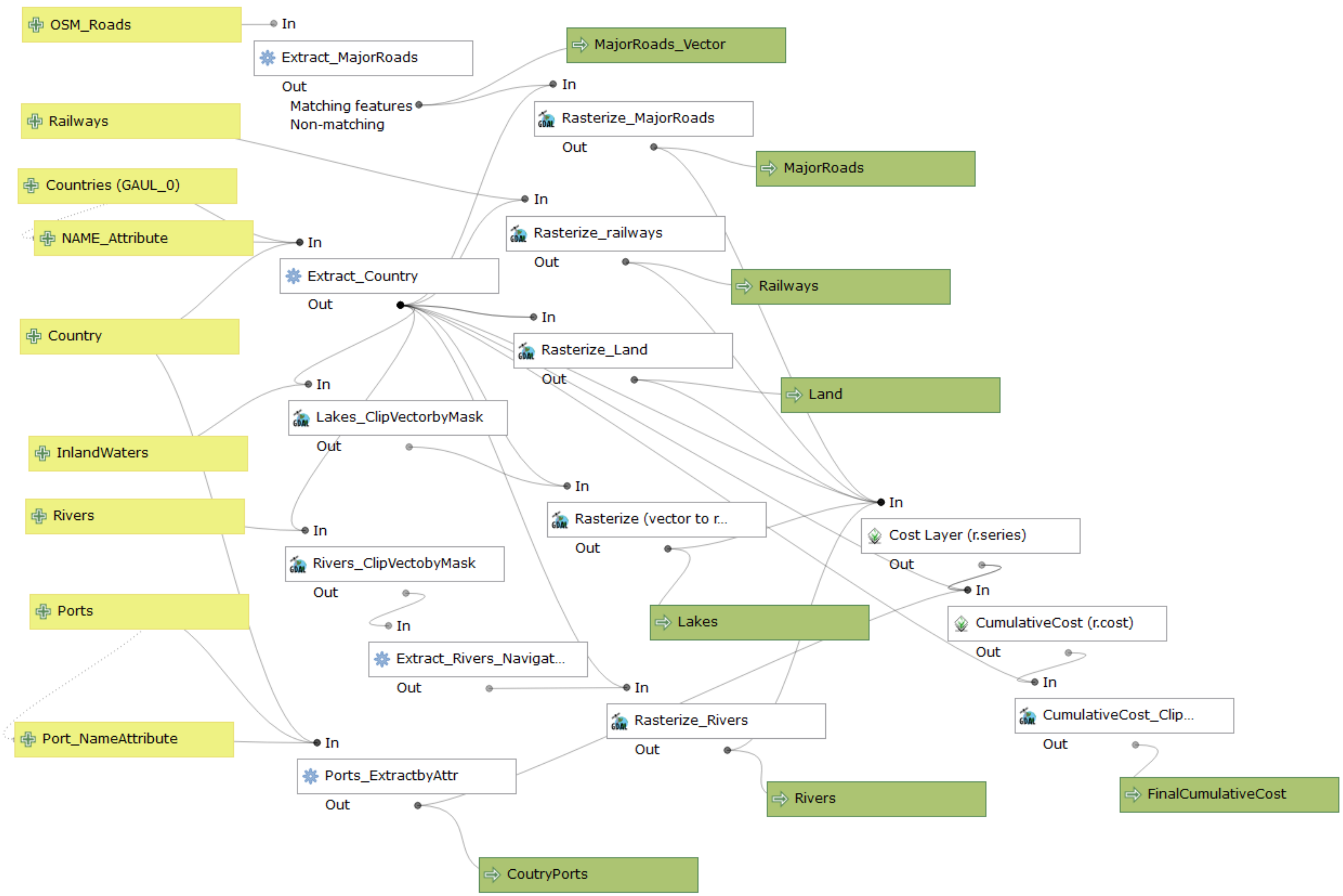


Figure 15 - TravelCostTimeSurface algorithm model diagram

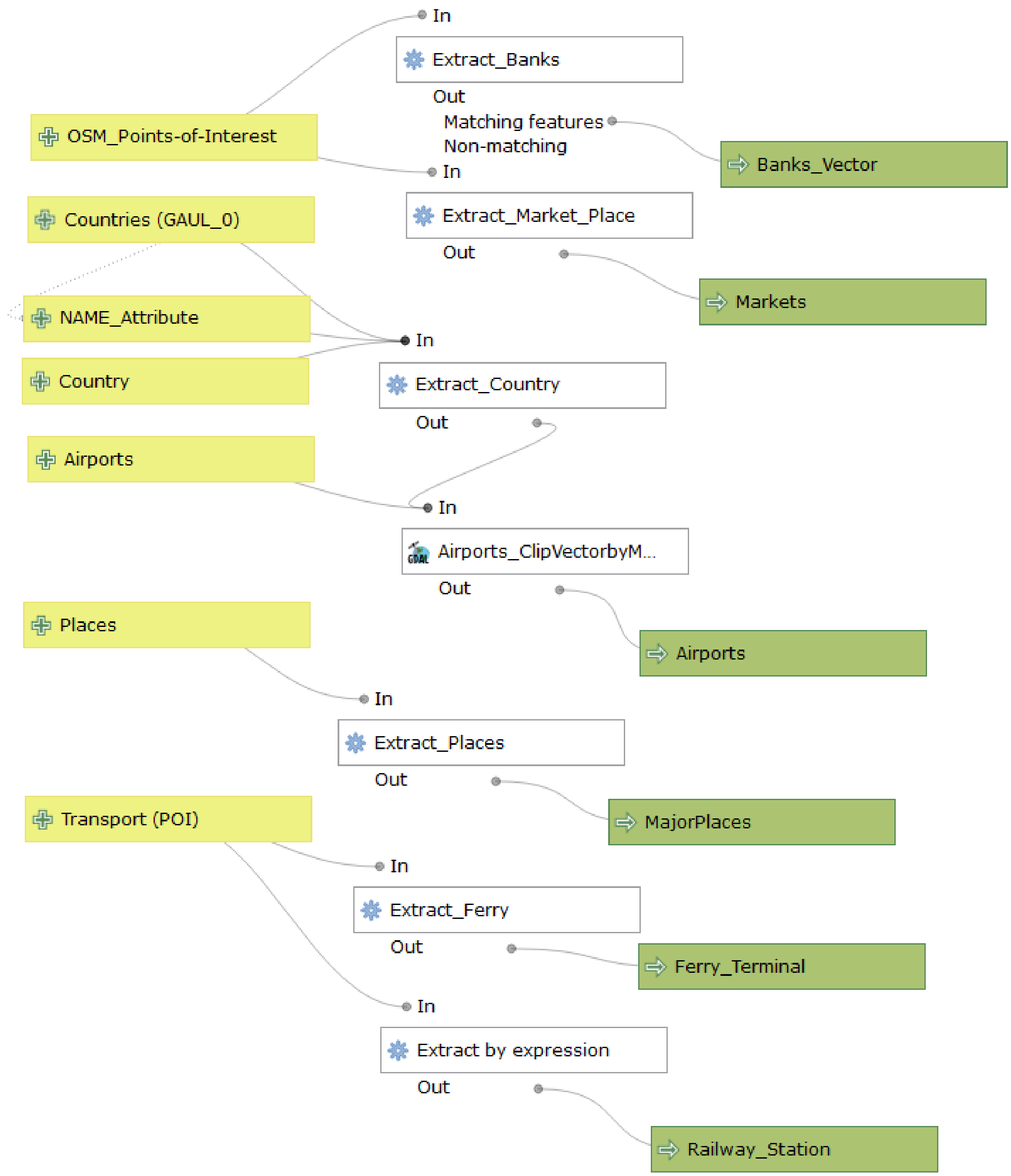


Figure 16 - AuxData algorithm model diagram

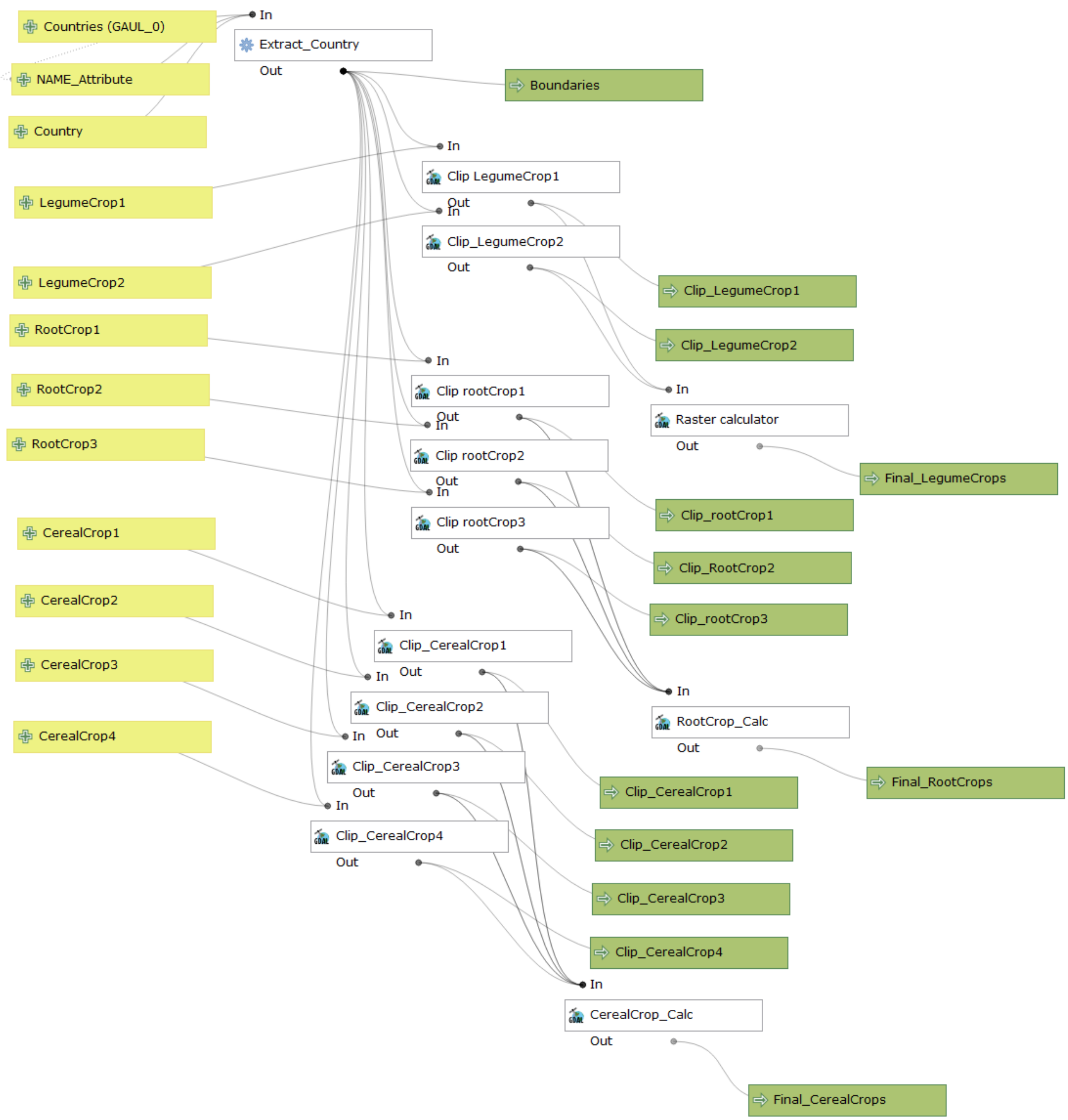


Figure 17 - ProdData algorithm model diagram