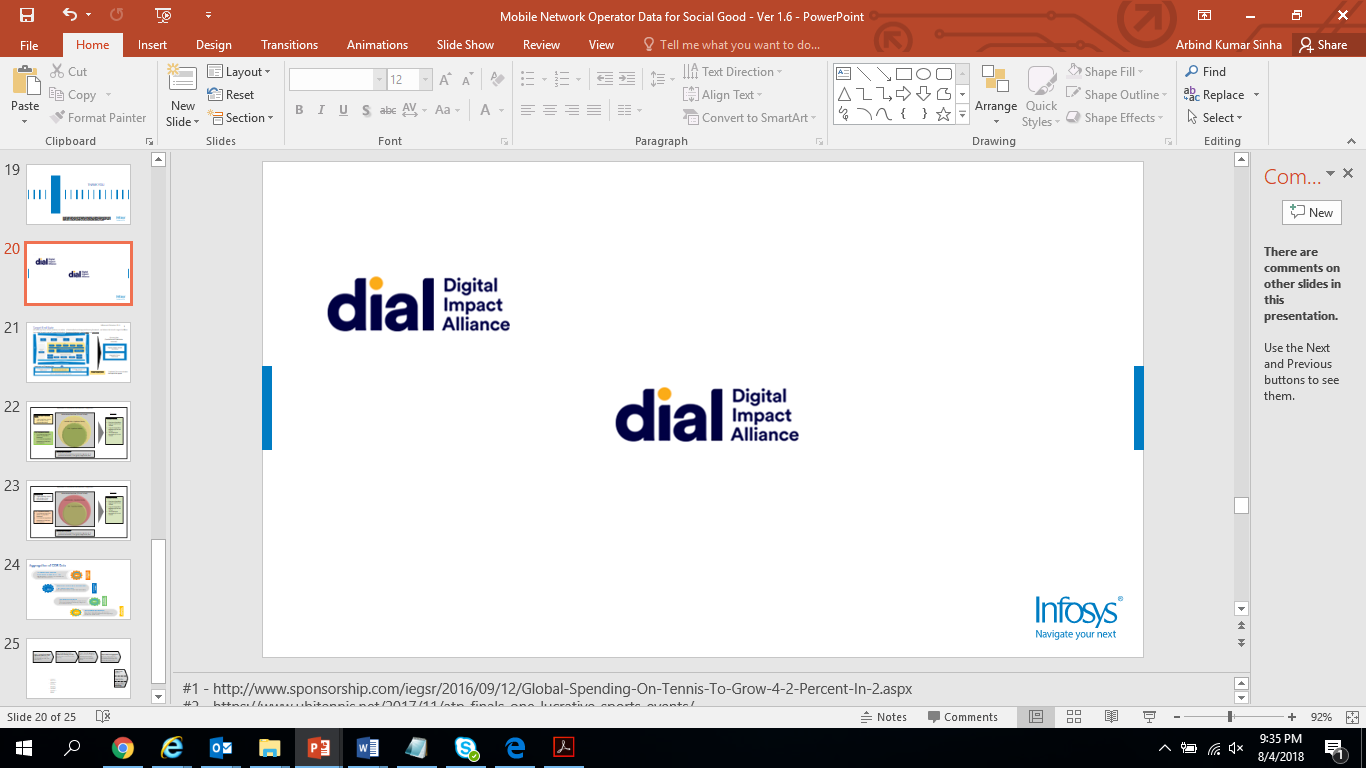
**Malawi Use Case**

***Technical Documentation on***

***Optimization Model***

**Version – 1 0**







Document Information and Revision History-

| **Version** | **Date** | **Author** | **Reviewer** | **Remarks** |
| --- | --- | --- | --- | --- |
| 1 0 | 27/03/2019 | Infosys | Infosys | Version-1 0 |
| 1 1 | 28/03/2019 | Infosys | Infosys | Version-1 1 |
|  |  |  |  |  |

**Table of Contents**

[1.0 Introduction 4](#_Toc4676518)

[2.0 Population Forecast 4](#_Toc4676519)

[2.1 Overview 4](#_Toc4676520)

[2.2 Approach 4](#_Toc4676521)

[2.3 Assumptions/Limitations 4](#_Toc4676522)

[2.4 Results Obtained 5](#_Toc4676523)

[3.0 TA Weight Calculation using Health Data 5](#_Toc4676524)

[3.1 Overview 5](#_Toc4676525)

[3.2 Approach 5](#_Toc4676526)

[3.3 Assumptions/Limitations 6](#_Toc4676527)

[3.4 Results Obtained 6](#_Toc4676528)

[4.0 Model Approach 6](#_Toc4676529)

[4.1 Problem Statement 6](#_Toc4676530)

[4.2 Assumptions 7](#_Toc4676531)

[4.3 Optimization Approach 7](#_Toc4676532)

[4.4 Datasets used 9](#_Toc4676533)

[4.5 Mathematical Formulation 10](#_Toc4676534)

[5.0 Reference 11](#_Toc4676535)

# 1.0 Introduction

The purpose of this document is to formulate an optimization problem for the allocation of 900 health posts within the next four years across all districts in Malawi.

This document will cover the population forecast, TA level weight calculation and approach adopted for optimization of health posts including the constraints / limitations involved as a part of optimization of 900 health posts.

# Population Forecast

## Overview

We are using Worldpop estimates for the years 2015 and 2020. Using these estimates, we are forecasting population for the next five years and adjusting it using long-term population movement.

## Approach

Following approach is adopted to forecast population for 2019-23.

* We have Worldpop population estimates for 2015 and 2020. Using these estimates, we calculate the population growth rate at national level per year as follows:

***Population Growth Rate = ln((2020 Worldpop population) / (2015 Worldpop population)) / t***

where t is time period, in our case t= 5

* This growth rate is used to project population for next year.
* This population count is adjusted according to the long term population movement (computed in phase 1) using:

Estimated Population count (1 + % change in Netflow at District level **\***)

## Assumptions/Limitations

* Population growth rate is considered same for all four years (2019-23)
* % change in Netflow is taken at TA level for Blantyre, Lilongwe and Mzuzu as the urbanization rate is high in these TAs.
* For all other TA’s we have used district level netflow.

## Results Obtained

Population forecasted for Year 2020-2023 *(Please refer Pop\_Estimate\_Adj\_NF\_Worldpop sheet).*



Census and Worldpop Population Comparison:

|  |  |  |  |
| --- | --- | --- | --- |
| Year | Population estimate using Worldpop | Census data | Difference \*\*\* |
| 2008 | 14,190,150 | 13,029,498 | 1,160,652 |
| 2018 | 18,811,171 | 17,563,749 | 1,247,422 |
| 2008-2018 Growth Rate | 32.56% | 34.80% |  |

*\*\*\* Shows our population estimates are closer to Census data.*

# TA Weight Calculation using Health Data

## Overview

Calculating weights at administrative unit level which will help us to prioritize health facility location. While calculating weight, disease burden of TA is important and hence it is considered in calculation of weight.

## Approach

Following approach is adopted to compute weight at TA level-

* For each health facility location, calculate the average number of people for each disease across all 24 months (2016-17).
* We have 27 District and Central Hospitals, these hospitals offer more services and therefore serve an entire district, so the disease burden is adjusted in the following way:

1. Calculate the population of the district and the percentage of this population that is in

each TA (so that the percentage across TA adds up to 100%).

2) Allocate the disease burden across the district in line with this percentage.

For example, if a TA contains 40% of a districts' population, add 40% of the District Hospital patients in each category to this TA's total disease burden

* Calculate number of people under each disease at TA level by aggregating the above calculated average across health posts.
* Calculate the % of people under each disease at TA level using 2016 and 2017 average estimated population for each TA.
* Calculate disease burden by taking weighted sum on all the % of people across each disease at TA level. These weights are assigned based on IHME (Institute for Health Metrics and Evaluation) ranking.
* Calculate rank for above calculated value across TA, which will be used as weights for the optimization model.

## Assumptions/Limitations

The calculation of TA weight approach assumes that district and central hospital provide the services to that district where the district /central hospital located. Therefore, this will also represent the disease burden for the district.

## Results Obtained

TA weights considering disease burden with disease ranking and without disease ranking.



# Model Approach

## Problem Statement

The Malawi Ministry of Health, in collaboration with the Digital Impact Alliance, is planning to roll out 900 health posts over next five years across all 28 districts of Malawi. These include both upgrades to existing facilities and the construction of new buildings to expand access, particularly in rural and remote areas, with an emphasis on the provision of primary health care.

As part of its draft Capital Investment Plan, the government of Malawi has drawn up proposed allocations of new facilities using the following 4 criteria:

1. Catchment Population
2. Distance to Nearest Existing Health Facilities
3. Facility Accessibility (high, medium, low)
4. Preferred year for work to take place as expressed by the District Health Monitoring Team

This report proposes to complement the above approach by incorporating population projections, as-well as short term and long-term population movement patterns, using these to optimize the placement of future health posts.

To approach this problem statement, certain assumptions are made as highlighted below.

## Assumptions

* All proposed new health posts are intended to serve a fixed catchment size ‘Cj’, with maximum catchment population of 12,500 and minimum catchment population of 1,200.
* Policy preferences are those outlined in the investment plan
* Populations clusters can be assigned to no more than one health-post
* Population clusters always opt for the nearest health post.
* Movement or shifting of existing health facilities are not allowed.
* Travel distance is considered a valid proxy of the cost and time to travel when optimizing the location of health facilities. Maximum travel distance is 6 KM.
* Capacity of a health post is considered based on the maximum number of patients that can be served based on its capacity constraint.
* All Health post provides similar level of services.
* There is a cap on # of health posts that can be built in a year.
* As in the investment plan, the cost to construct health posts across locations is assumed uniform across districts ($36,573), unless otherwise specified.
* The established facility will serve the calculated catchment area over the following years.

## Optimization Approach

In this approach, we may assume this problem statement to be a location-allocation problem – with minimizing the uncovered population given binding constraints on the number of facilities. To proceed with this approach, we chalk out the steps as discussed below.

Note: Step 1 and Step 2 are done at TA level.

Step1: Obtain the population densities

* We have Worldpop population estimates for 2015 and 2020. Using these estimates, we calculate the population growth rate at national level per year as follows:
* ***Population Growth Rate = ln((2020 Worldpop population) / (2015 Worldpop population)) / t***

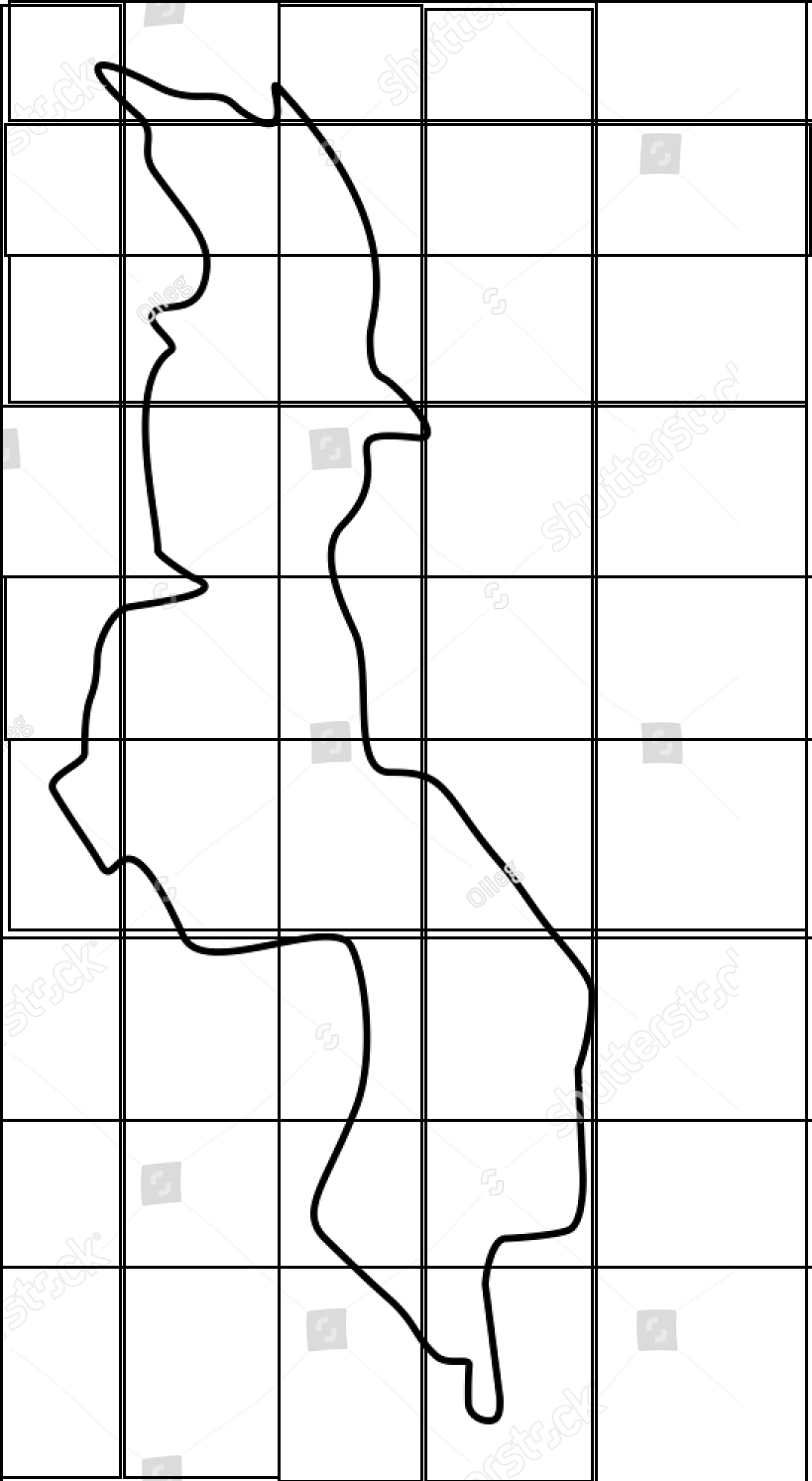
where t is time period, in our case t= 5

* This growth rate is used to project population for next year.
* This population count is adjusted according to the long term population movement (computed in phase 1) using:

Estimated Population count (1 + % change in Netflow at District level)

Step 2: Analogous estimation of population

* Overlay grid (0.5 sqkm) on Worldpop raster file and calculate the % of population in each grid.
* This % is then multiplied with estimated population from Step 1 to get distribution of estimated population.



Step 3: Extracting Uncovered Population Area using UNICEF shape file.

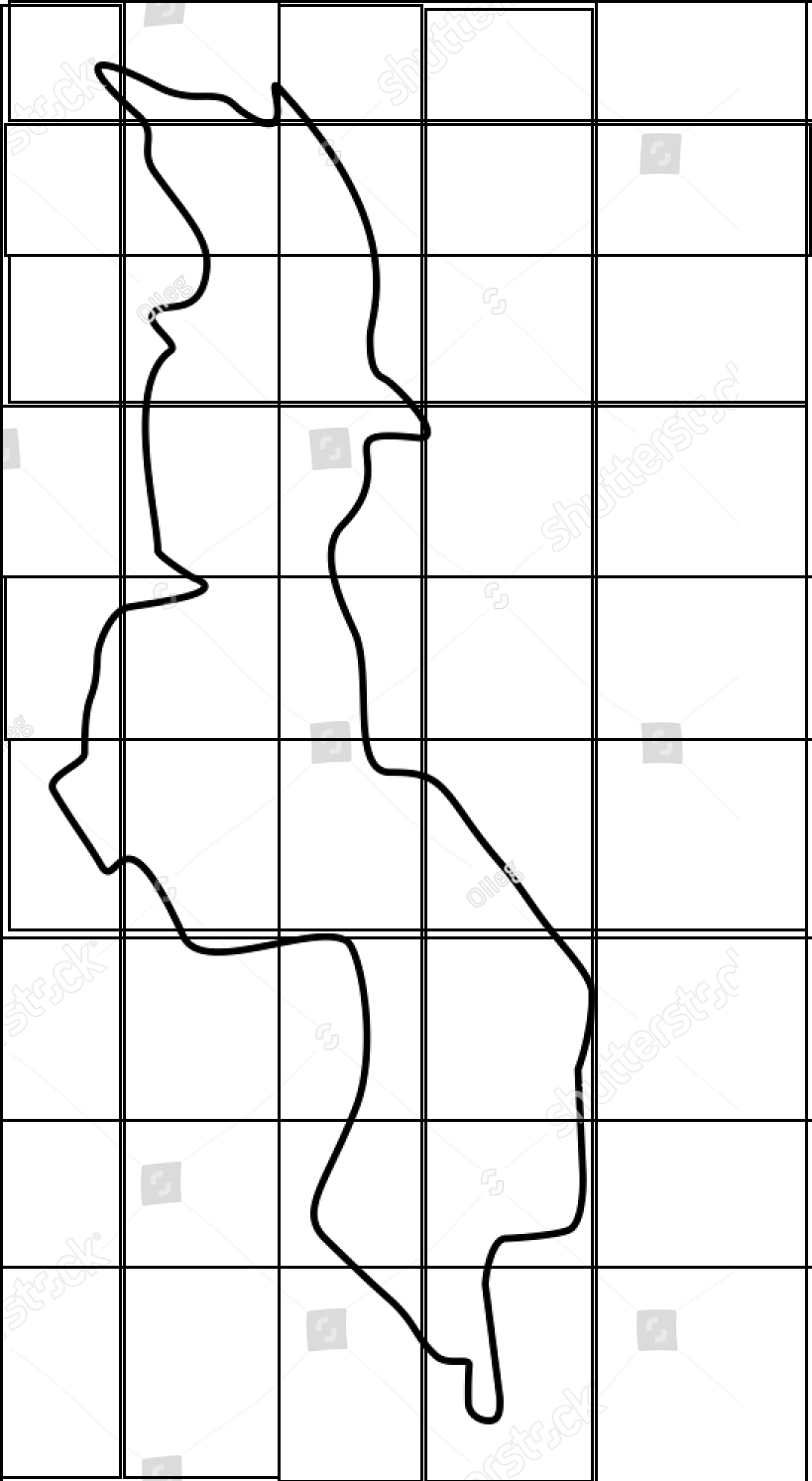
* UNICEF conducted an extensive survey of every operating clinic in the country and calculated the relevant catchment area. This catchment area constitutes the distance such that a patient would have to walk no more than 5km to reach a clinic. In its analysis, it accounted for the road networks, topography and potential for flooding in calculating the 5km catchment area for each clinic.

“Thanks to UNICEF Malawi country office for sharing this data with us”.

* Overlay this file over the Step 2 output.
* This gives us **uncovered population** (Popu).

Step 4: Formation of Clusters (For uncovered population)

* Identify centroid of each grid.
* Cluster grids (by measuring the shortest path between centroids) till you reach catchment population (Cj). This gives one cluster of grids.
* Continue the above till each grid is a part of one cluster.



Step 5: The centroid of grid in the cluster having maximum population is the **probable location for health posts**.

Step 6: Select a maximum distance beyond which a patient cannot travel to the health post - “D” called **Impedance** cut-off. The **catchment** area of a facility is defined such as either it covers the catchment population of Cj or reaches the impedance cut-off.

Step 7: We are building three models (i.e. Model-1, Model-2 and Model-3) for optimization of health posts

* Model-1 where weights are assigned as per uncovered population to each TA level.
* Model-2 where weights are assigned according to disease burden to each TA
* Model-3 where weights are assigned according to disease burden and disease ranking to each TA.

Step 8: We get ‘H’ locations out of all the probable locations based on weight assigned to that TA and catchment population ‘Cj’ in the cluster. ‘H’ is the maximum number of health post that can be built in each year.

Now repeat the above steps from 3 to 7 for next four years. We have to use earlier year shapefile which will give uncovered area for next year. If a potential candidate (health post location) for a given year turns out to be a favorable location, then for the next year the health post becomes a required candidate.

## Datasets used

* List of existing health facilities with location information.
* Facility burden (# of patients per facility) for each facility.
* For Disease ranking IHME DALY’s is used.
* Long-term movements (computed) for each administrative unit for each year.
* Worldpop 2015 and 2020 population.
* Shape file of UNICEF to know the uncovered population.

## Mathematical Formulation

*Objective function***:** Min **Popu - ∑∑ wj \* ∑ Pjjt \* ‘Cj’**

*Constraints:*

≤900 (900 is the maximum number of health posts that can be deployed across all administrative units over the 5 years)

≤ H (Threshold “H” is the maximum number of health facilities that can be deployed across all administrative units in the year “t”).

A population centroid “i” in administrative unit “j” must be mapped maximum to 1 existing health facility.

(summed over “x" existing facilities for a given “i” at adm unit “j” for year “t”) ≤ 1

For a given existing facility “x” in administrative unit “j” must be mapped maximum to 1 population centroid.

(summed over for “i" population centroids for a given “x” at adm unit “j” for year “t”) ≤ 1

If a centroid “i” in administrative unit “j” is getting mapped to a particular existing facility “x” and similarly if the existing health facility is getting mapped to the same population centroid “i” then candidate health post (proposed) should not be posted for that centroid.

=0; If the above case is not true, else =1

Here, we need to consider deploying a health post. Among the proposed health post candidates, select the location which is closest to the centroid “i”.

= (sum over all the centroids “i” for administrative unit “j” for the year “t”)

*Where:*

: Health post “p” in administrative unit “j” in the year “t”.

: Existing health post “x” in administrative unit “j” in the year “t”.

: distance “d” from existing health post “x” to population cluster centroid “i” in administrative unit “j” in the year “t”.

: distance “d” from candidate (proposed) health post “x” to population cluster centroid “i” in administrative unit “j” in the year “t”.

: is equal to 1 for the pair of centroid “i” and existing facility “x” which has shortest distance-

“” ≤ D (impedance distance). If no pair are between “D” then zero.

: is equal to 1 for the pair of centroid “i” and probable candidate facility “p” which has shortest distance- “” ≤ D (impedance distance). If no pair are between “D” then zero.

= 1, if there is requirement for new facility. Else 0.

# Reference

1. **(for basic approach):** Stummer, C., Doerner, K., Focke, A. et al. Health Care Management Science (2004) 7: 63. <https://doi.org/10.1023/B:HCMS.0000005399.23600.69>
2. Weight calculation- **IHME global burden of disease profile:  Malawi**